World Applied Sciences Journal 33 (5): 715-721, 2015
ISSN 1818-4952

DOI: 10.5829/idosi.wasj.2015.33.05.30

Signal Processing Algorithm for Brain Computer Interface - A Review

A. Sharmila

School of Electrical Engineering, VIT University, Vellore-14, Tamilnadu, India

Abstract: Brain–computer interface (BCI) aim at providing a non-muscular channel for sending commands to the external world using the electroencephalographic activity or other Electrophysiological measures of the brain function. An essential factor in the successful operation of BCI systems is the methods used to process the brain signals. In the BCI Literature, comprehensive survey of the signal processing techniques has been used.

Key words: Brain computer Interface · Electroencephalograph (EEG) · Signal processing

INTRODUCTION

The ultimate purpose of a direct brain–computer interface (BCI) is to allow an individual with severe motor disabilities to have effective control over devices such as computers, speech synthesizers, assistive appliances and neural prostheses. Such an interface would increase an individual's independence, leading to an improved quality of life and reduced social costs.

A BCI system detects the presence of specific patterns in a person's ongoing brain activity that relates to the person's intention to initiate control. The BCI system translates these patterns into meaningful control commands. To detect these patterns, various signal processing algorithms are employed.

Signal processing forms an important part of a BCI design, since it is needed in extracting the meaningful information from the brain signal. This paper summarizes the results of a comprehensive survey of different signal processing schemes that have been used in BCI systems. It specifically focuses on the following signal processing components of a BCI: the preprocessing referred as the signal enhancement, feature extraction, feature selection/dimensionality reduction, feature classification and post-processing blocks. To address all related BCI research, we include all the approaches that use standard scalp-recorded EEG as well as those that use epidural, subdural or intracortical recordings.

Characteristics of Eeg: Electroencephalograph (EEG) was first recorded by Berger in 1929 by externally attaching several electrodes on the human skull [1]. Such signals generally deliver in indirect way information about physiological functions, which are related to the brain. Possible applications using such signals are very numerous. They are for example integrated in the design of new technological devices with embedded intelligence and allow for Brain-Computer-Interfaces. There is also an important demand, in the medical domain, for automatic signal interpretation systems [2]. BCI is composed of signal collection and processing, pattern identification and control systems [3].

There are five major brain waves distinguished by their different frequency ranges [4]: Delta waves lie within the range of 0.5 to 4 Hz, Theta waves lie within the range of 4 to 7 Hz, with an amplitude usually greater than 20 μ V, Alpha with a rate of change lies between 8 and 13 Hz, with 30-50 μV amplitude, Beta, the rate of change lies between 13 and 30 Hz and usually has a low voltage between 5-30 μV. Beta is the brain wave usually associated with active thinking, active attention, focus on the outside world or solving concrete problems and finally the Gamma waves which lie within the range of 35Hz and up. It is thought that this band reflects the mechanism of consciousness. Theta, alpha and beta frequencies are used in our work to classify the mental tasks.

The primary focus of this paper is to summarize feature extraction methods that have been used in EEG pattern recognition applications. Therefore, techniques using implanted devices or EEG biofeedback will only be mentioned briefly.

Corresponding Author: A. Sharmila, School of Electrical Engineering, VIT University, Vellore-14, Tamilnadu, India.

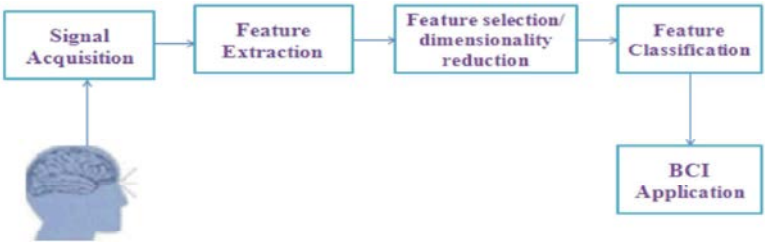

Fig. 1: Block diagram of BCI

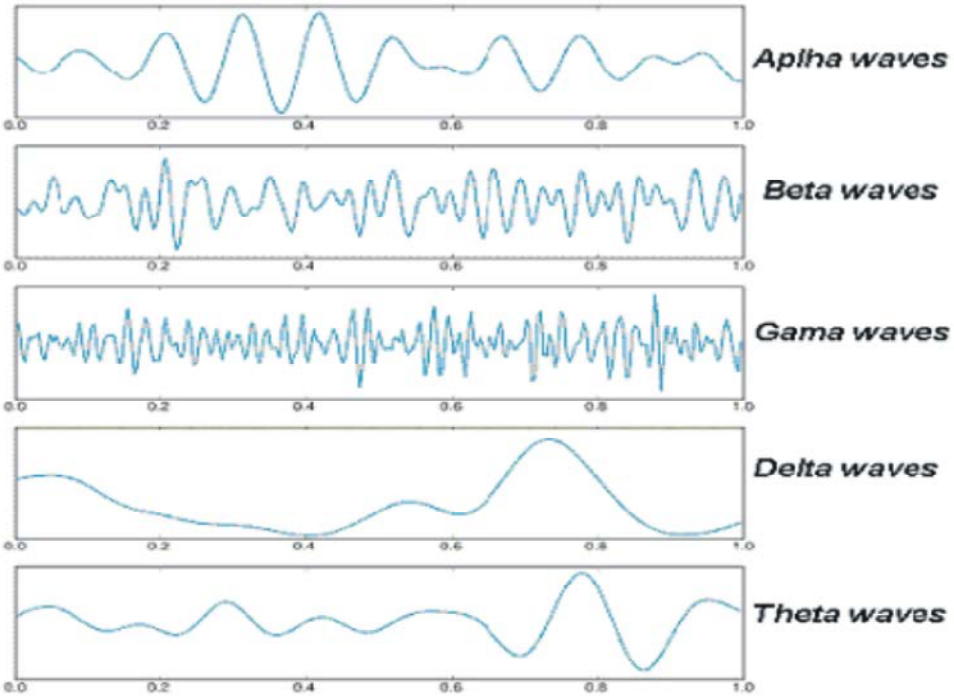

Fig. 2: Brain wave types

Implanted Techniques: Invasive techniques permit signals to be recorded more directly from their sources. For example, Kennedy, Bakay and Moore [1, 2, 3, 4] have developed and experienced a method involving a neurotrophic electrode implanted in the human neocortex which detects signals and transmits them through the skull/scalp to a receiver which allows control of computer mouse function after signal processing have been performed. Also, this scheme has been successful in allowing human subjects with locked-in syndrome to communicate with the outside world.

EEG Techniques: Although implanted electrodes afford cleaner signals for a BCI, the invasive techniques hold noticeable drawbacks. Exterior, Electroencephalograph (EEG), electrodes have been researched widely as a non-invasive BCI signal acquisition technique. While, because of the inherently low signal-to-noise ratio of EEG and the effects of spatial blurring due to the volume conductor model of the head, accurate, reliable and reproducible signals are hard to create/detect.

EEG based BCI research can be classified into two broad categories:

- Subject-training
- Computer-training

Subject-Training Methods: In subject-training EEG BCI methods, a reliable algorithm is used to transform raw EEG into some form of feedback (e.g., cursor movement on a computer monitor). The subject is then asked to modify their mental state in order to attain a specific goal (e.g., move the cursor to particular position).

Computer-Training Methods: In computer-training methods, various reliable mental states are processed by a computer using pattern recognition techniques in attempts to distinguish among them.

Cross-Over Systems: It should be well-known that a lot of systems have elements of both computer-training and subject-training. For example, a system designed with recorded data by using computer-training could later be tested on-line. For the period of this real-time testing process, the user could learn to improve the performance of the system, thus introducing a subject-training factor.

Review of Bci Feature Extraction: The literature related to feature extraction methods applied in BCI pattern recognition works are discussed as follows.

Raw EEG: Raw EEG has infrequently been used as a feature for BCI EEG pattern recognition. This is liable due to the fact that the amount of data available in Raw EEG would overcome most pattern recognition techniques. For example, a 1 second window recorded at 250 Hz from 32 electrodes would produce 8000 features. Usually, a method should be used to lessen the amount of experimental dimensions or extract related information.

Fast Fourier Transform (FFT): As with Raw EEG, the Fourier Transform (or any Frequency-Domain representation) of EEG as well contains large amounts of data which can overcome several pattern recognition methods. On the other hand, BCI research has been conducted by extracting valuable portions of this frequency-domain data. One such technique is observing Event-Related Desynchronization (ERD) in the mu-rhythm (8 – 12 Hz). Penny, Roberts and Stokes [5] collected EEG data from subjects who were given cues to either imagine or execute right and left hand movements. This data was then transformed using the Laplacian operator, then the FFT was computed. Pineda *et al.* [1] conducted recordings during four conditions: baseline, hand-movement, imagined hand-movement and observed hand-movement. They observed considerably different mu-rhythms in all three of the movement-related trials in evaluation with the baseline.

Auotoregressive(AR): In autoregressive (AR) techniques, a model is formed where a current voltage can be predicted from N past voltages where the model order is N. Thus the model can be represented as:

$$x_{i,e}(t) = -\sum_{i=1}^{N} a_{i,e} x_{i,e}(t-i)$$

where $a_{i,e}$ is the ith order AR coefficient for electrode e. These AR coefficients can be used as features. To obtain these coefficients, EEG data is generally windowed into blocks of data with more than N samples. Then, as the value of t is shifted through the window of data, we obtain numerous model equations which permit us to compute optimum AR coefficients. Thus, these AR coefficients can be used to symbolize the mental state during that window of time.

During preliminary BCI work anderson *et al.* [7] establish that 6th order AR parameters calculated with the Burg method [8] could be used to estimate spectral density which was used to calculate bilateral spectral asymmetry ratios from EEG recorded during two mental tasks (baseline/relaxed and mental arithmetic) for use as a feature in the training/testing of small (1 to 5 hidden units) feed forward artificial neural networks (ANNs) trained with back propagation

Wavelet: Wavelets are fundamentally a concession between time-domain and frequency-domain because they allow the user to view change in frequency bands over time (with less resolution than just time-domain or frequency-domain). The Discrete Wavelet Transform (DWT) can be computed as a series of filters. The filtering and down sampling cycle exposed can be repeated until there is only one sample remaining. Therefore, the maximum number of resulting DWT coefficients is equal to the number of samples in the original signal. The computationally expensive Continuous Wavelet Transform (CWT) can create a lot of additional coefficients, on the other hand the resulting closely spaced scales are extremely correlated and it can be revealed that no information is lost (i.e., using an inverse DWT, the original signal can be reconstructed).

A additional discrete method, called Wavelet Packets, allows the DWT frequency bands (coefficient scales) to be further subdivided.

Complexity: Rezek and Roberts [16] used one complexity measure, Embedding-Space Decomposition (ESD), in BCI experiments. They achieved an average 20% error rate (with 4 out of 8 subjects excluded from this figure) using 4 second windows of EEG data collected during imagined finger movements versus control state. The decision boundary was created by considering whether a complication peak occurred within 3 seconds of an imagined movement.

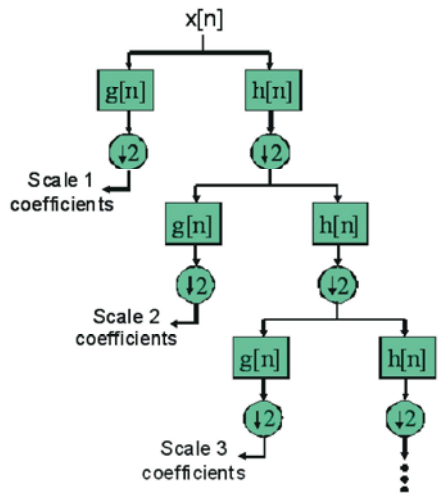

Fig. 2: Block level representation of Discrete Wavelet Transformc

Even as the BCI work conducted using complexity measures is partial, significant research has been conducted in other EEG pattern recognition areas which suggests that complexity could be a useful feature extraction method for BCI systems. The following subsection discusses how these various complexity measures are implemented.

Autoregressive Optimal Model Order (ARMO): One way of quantitatively representing signal complexity is to find the optimal order of an autoregressive (AR) model for that signal. In AR techniques, a model is created where a current voltage can be predicted from p past voltages where the model order is p. Thus the model can be represented as:

$$x_{i,e}(t) = -\sum_{i=1}^{p} a_{i,e} x_{i,e}(t-i)$$

A number of difference methods exist for estimating this optimal order, p. These include Final Prediction Error (FPE), Akaike's Information Criterion (AIC), Schwarz Information Criterion (SIC) [also known as Bayesian Information Criterion (BIC) or the Schwarz Bayesian Criterion (SBC)] and Residual Error.

Spectral Entropy (SE): Spectral Entropy [16] is given by:

$$H = \sum_f p_f \log(1/p_f)$$

where f is frequency and p_f is a normalized Power Spectral Density (PSD) function. One method of obtaining a PSD is using the Fourier Transform.

Approximate Entropy (ApEn): Approximate Entropy is given by [16]:

$$ApEn(m,r,N) = \Phi^{m+1}(r) - \Phi^{m}(r)$$

where:

$$\Phi^{m} = \log[P\|u_{jm} - u_{im}\| \le r]$$

and where r is a fixed tolerance parameter.

Another way of defining this algorithmError! Bookmark not defined. is:

$$ApEn(m,r,L) = \frac{1}{L-m}\sum_{i=1}^{L-m} \log C_i^{m+1}(r) - \frac{1}{L-m+1}\sum_{i=1}^{L-m+1} \log C_i^{m}(r)$$

where m is the length of the pattern, r is the effective filter and $C^m(r)$ is the correlation integral with embedding dimension m and time lag 1.

Embedded Space Decomposition (Taken's Algorithm): One method of representing an embedded space decomposition is using a fractional spectral radius (FSR) [42]:

$$FSR(j) = \frac{\sum_{i=0}^{j} \sigma_i^2}{\sum_{l=0}^{m} \sigma_l^2}$$

where σ_i^2 are the eigenvalues of U^TU and U is an embedding matrix made up of vectors $u^T_1,...,u^T_{N-(i-1)}$ such that $u_i = (x_i, x_{i+j},...,x_{i+(m-1)j})^T$. Thus u_i is the embedding vector whose elements are m samples taken at intervals of J samples along the observed time series. Similar methods were used in other EEG studies as well [16] [32].

Fractal Dimension: Fractal Dimension is a chaotic method of estimating signal complexity. Several fractal dimension algorithms have been tested for EEG, including algorithms by Petrosian[17] and Katz [18], however an algorithm by Higuchi [19, 20, 21, 22, 23, 24] appears to be the most widely used. Under Higuchi's algorithm, one finds the Fractal Dimension, D_F, by a linear least-squares best fit as follows:

$$D_F = \frac{n * \sum(x_k * y_k) - \sum x_k \sum y_k}{n * \sum(x_k^2) - (\sum x_k)^2}$$

where: $y_k = \ln$ IL,(k), $x_k = \ln(1/k)$, $k = k_{min},...,k_{max}$, Q, Q is n minus the total number of different values of k in the interval $[k_{min} k_{max}]$ and:

$$L_m(k) = \frac{\sum_{i=1}^{\left\lfloor \frac{N-m}{k} \right\rfloor} \left| x(m+ik) - x(m+(i-1)k) \right| (n-1)}{\left\lfloor \frac{N-m}{k} \right\rfloor k}$$

Multiple-Electrode Linear Methods: Use of the measures covered in this section towards EEG research is covered in Appendix B.

Global Field Strength Σ: A measure of global field strength, (Σ), is defined as:

$$\sum = \frac{1}{L}\sqrt{\frac{\sum_n \|u_n\|_2^2}{16L}}$$

where u_n is the row vector of a multichannel EEG matrix (where row=time sample and column=electrode), L is the length of the signal (# of rows) and $\| . \|_2$ is the 2-norm.

Global Frequency of Field Changes Φ: A measure of global frequency of field changes, (Φ), is defined as:

$$\Phi = \frac{1}{2\pi}\sqrt{\frac{\sum_n \|u_n\|_2^2}{\sum_n \|(u_n - u_{n-1})/\Delta t\|_2^2}}$$

where Δt is the sampling period, u_n is the row vector of a multichannel EEG matrix (where row=time-sample and column=electrode) and $\| . \|_2$ is the 2-norm.

Spatial Complexity Ω: A measure of spatial complexity, (Ω), is defined as:

$$\log \Omega = -\sum_{i=1}^{16} \xi_i \log \xi_i$$

where ξ represents the normalized ($\xi_i = \lambda_i / \Sigma_i \lambda_i$, where λ_i for i=1:16 represents the eigenvalues of C) eigenvalues of C:

$$C = \frac{1}{L}\sum_n u_n u_n^T$$

where u_n is the row vector of a multichannel EEG matrix (where row=time-sample and column=electrode) and L is the length of the signal (# of rows).

LaPlacian: The Laplacian, or current source density is frequently used in EEG to boost spatial resolution. The majority recent version, the Spline-Laplacian, involves taking the second spatial derivative of the spline function in order to interrupt electric potentials across the scalp and increase spatial resolution [7]. This technique yields estimate of local current passing perpendicular to the skull surface into the scalp. When combined with a high electrode density, it has been revealed that computation of the Spline-Laplacian can increase spatial resolution by a factor of three (compared to raw EEG). This causes the scalp topography to become more exhaustive and allow for a greater number of voltage peaks and troughs which can otherwise be blurred. This in turn helps reduce the spatial resolution confines of EEG. Also, the Spline-Laplacian method eliminates the influence of reference electrode placement which produces a variable bias over electrodes in raw EEG recordings. Though the Laplacian fits into this section as a feature extraction method, it is often considered more of a pre-processing technique.

Principal Components Analysis: Principal Components Analysis (PCA) can also be considered a feature extraction method; but its main idea is dimensionality reduction. It would usually be used as the last step before pattern recognition. For example, one might want to attain the wavelet coefficients from the EEG data and then conduct PCA on that data set to reduce the dimensionality. Decreasing the dimensionality of a signal makes it easier for the pattern recognition method to separate the data into classes. The drawback of PCA is that few useful information may be discarded in them dimensionality reduction procedure.

The most common PCA algorithm involves finding a weight vector which projects the original data into a smaller vector. This weight vector is found by solving for the eigenvectors of the covariance of the original data. These eigenvectors are then sorted by their respective eigen values and a cutoff point is chosen below which all eigenvectors are discarded (thus varying the degree of dimensionality reduction). On the other hand, a neural network can be used to carry out PCA by learning to map the larger data linearly onto a smaller one.

A number of non-linear methods for dimensionality reduction have also been proposed. These methods include auto associative networks (supervised learning) [25, 26, 27], non-linear extensions to Hebbian learning (unsupervised learning) [28, 29] and Independent Component Analysis (ICA) [30].

Devulapalli [31] conducted BCI research comparing PCA and non-linear PCA by processing raw EEG gathered during two mental tasks (letter composing & mental arithmetic) using a backpropagation-trained neural network for pattern recognition. It was concluded that the non-linear dimensionality reduction methods were superior.

CONCLUSION

The most important confront of an optimized EEG based BCI is dimensionality reduction. The amount of data available in raw EEG overwhelms nearly all pattern recognition techniques (given current computing power). Consequently, feature extraction algorithms, with a goal to extract as much relevant information as possible while reducing data to a small enough set, have been the center of attention of most BCI research. But, it appears that research in that direction is reaching a level.

A promising way of future research could be the use of very large neural networks. This pattern recognition technique can handle thousands of dimensions, still in this case would require tens of thousands of example sets and a significant amount of time and/or computing power for training. It may be likely to create a better BCI by eliminating less of the data through the use of such a large neural network either with current super-computers or with future personal computers.

At the same time it may be possible to further improve current BCI systems. It has been proven that BCI systems are possible using EEG, it is likely that future systems could be implemented using implanted electrodes or innovative non-invasive technologies which could be much faster and have multiple dimensions.

ACKNOWLEDGEMENT

We would like to thank Vellore Institute of Technology University for providing this opportunity to do the research work.

REFERENCES

1. Kennedy, P.R., R.A.E. Bakay, M.M. Moore, K. Adams and J. Goldwaithe, 2000. Direct Control of a Computer from the Human Central Nervous System. IEEE Transactions on Rehabilitation Engineering, 8(2): 198-202.
2. Goettling, G., 1999. Harnessing Brain Power: Georgia Tech research assists direct brain-computer interaction. *Georgia Tech Research Horizons*, Fall. Retrieved March 5, 2003, from: http://gtresearchnews.gatech.edu/reshor/rh-sf99/default.html.
3. Kennedy, P.R., K. Adams, R.A.E. Bakay, J. Goldthwaite, G. Montgomery and M. Moore, 1999. Direct control of a computer from the human central nervous system. *CI Conference New York*, June 16th to 20th.
4. Kennedy, P.R. and R.A.E. Bakay, 1998. Restoration of neural output from a paralyzed patient using a direct brain connection. NeuroReport, 9: 1707-11.
5. Farwell, L.A. and E. Donchin, 1988. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroenceph. and Cl. Neurophys., 70: 510-523.
6. Polikoff, J.B., H.T. Bunnell and W.J. Borkowski Jr., 1999. Toward a P300-based Computer Interface. Technical Report, Applied Science and Engineering Laboratories, Alfred I. Dupont Institute.
7. Pelz, J.B., M.M. Hayhoe, D.H. Ballard, A. Shrivastava, J.D. Bayliss and Von Der Heyde, (unkown date). M. Development of a virtual laboratory for the study of complex human behavior. Technical Report, University of Rochester and Max-Planck-Institute for Biological Cybernetics.
8. Bayliss, J.D. and D.H. Ballard, 1999. Single Trial P3 Recognition in a Virtual Environment. Technical Report, Department of Computer Science, University of Rochester.
9. Bayliss, J.D. and D.H. Ballard, 2000. A Virtual Reality Testbed for Brain-Computer Interface Research., IEEE Trans. on Rehabilitation Engineering, 8(2): 188-190.
10. Bayliss, J.D. and D.H. Ballard, 1998. Single Trial P300 Recognition in a Virtual Environment. Technical Report 98.1, Department of Computer Science, University of Rochester.
11. Penny, W.D., S.J. Roberts and M.J. Stokes, 1998. Imagined Hand Movements Identified From the EEG Mu-Rhythm, Technical Report, Department of Electrical Engineering, Imperial College.
12. Pineda, J.A., B.Z. Allison and A. Vankov, 2000. The Effects of Self-Movement, Observation and Imagination on Mu Rhythm And Readiness Potentials (RPs): Towards A Brain-Computer Interface (BCI). IEEE Trans. Rehabil. Eng., 8(2): 219-22.
13. Anderson, C., S. Devulapalli and E. Stolz, 1995. Determining Mental State from EEG Signals Using Neural Networks. Scientific Programming, Special Issue on Applications Analysis, 4(3): 171-183.
14. Keirn, Z.A. and J.I. Aunon, 1990. A new mode of communication between man and his surroundings. IEEE Transactions on Biomedical Engineering, 37(12): 1209-1214.
15. Proakis, J.G. and D.G. Manolakis, 1992. Digital Signal Processing. MacMillan, New York.

16. Roberts, S.J., W. Penny and I. Rezek, 1998. “Temporal and Spatial Complexity Measures for EEG-Based Brain-Computer Interfacing”, Medical & Biological Engineering and Computing, 37(1): 93-99.
17. Petrosian, A., 1995. “Kolmogorov Complexity of Finite Sequences and Recognition of Different Preictal EEG Patterns,” Proc. IEEE Symposium on Computer-Based Medical Systems, pp: 212-217.
18. Katz, M., 1998. “Fractals and the analysis of waveforms,” Comput. Biol Med., 18(3): 145-156.
19. Higuchi, T., 1988. ”Approach to an irregular time series on the basis of the fractal theory”, Physica D, 31: 277-283.
20. Ciszewski, J., W. Klonowski, R. Stepien, W. Jernajczyk, A. Karlinski and K. Niedzielska, 1999. "Application of Chaos Theory for EEG-signal Analysis in Patients with Seasonal Affective Disorder", Med. Biol. Eng. Comput., 37(Supplement 1): 359-360.
21. Klonowski, W., E. Olejarczyk and R. Stepien, 2000. Nonlinear Quantifiers of EEG-signal Complexity, Proceedings of 2000 International Symposium on Nonlinear Theory and its Applications (NOLTA'2000), Dresden, Germany, pp: 261-264.
22. Klonowski, W., E. Olejarczyk and R. Stepien, 2001. Chaocity and Dimensional Complexity of EEG-signal, Proceedings of 2000 International Symposium on Nonlinear Theory and its Applications (NOLTA'2001), Tokyo, Japan.
23. Esteller, R., G. Vachtsevanos, J. Echauz and B. Litt, 1999. "A Comparison of Fractal Dimension Algorithms Using Synthetic and Experimental Data," Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS 1999), Orlando, FL, May 30 - Jun. 2.
24. Accardo, A., M. Affinito, M. Carrozzi and F. Bouquet, 1997. “Use of the fractal dimension for the analysis of electroencephalo-graphic time series”, Biol. Cybern., 77: 339-350.
25. Kramer, M.A., 1992. Autoassociative neural networks. Computers and Chemical Engineering, 16(4): 313-328.
26. Kramer, M.A., 1991. Nonlinear principal component analysis using autoassociative neural networks. American Institute of Chemical Engineering Journal, 37(2): 233-243.
27. Kirby, M.J. and R. Miranda, 1996. Circular nodes in neural networks. Neural Computation, 8(2): 390-402.
28. Sanger, T.D., 1991. Optimal hidden units for two–layer nonlinear feedforward neural networks. International Journal of Pattern Recognition and Artificial Intelligence, 5(4): 545-461.
29. Karhunen, J. and J. Joutsensalo, 1994. Representation and separation of signals using nonlinear PCA type learning. Neural Networks, 7(1): 113-127.
30. Jutten, C. and J. Herault, 1991. Blind seperation of sources, part i: An adaptive algorithm based on neuromimetic architecture. Signal Processing, 24: 1-10.
31. Devulapalli, S., 1996. Non-Linear Principal Component Analysis and Classification of EEG During Mental Tasks (MS Thesis). Fort Collins, Colorado: Colorado State University.
32. Bhattacharya, J., 2000. “Complexity Analysis of Spontaneous EEG”, Acta. Neurobiol. Exp., 60: 495-501.