Computer Engineering and Intelligent Systems
ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online)
Vol 3, No.3, 2012

A Novel Neural Network Classifier for Brain Computer Interface

Aparna Chaparala[1*] Dr. J.V.R.Murthy[2] Dr. B.Raveendra Babu[3] M.V.P.Chandra Sekhara Rao[1]

1. R.V.R.&J.C. College of Engineering, Guntur - 522019, AP, India
2. Dept. of CSE, JNTU College of Engineering, Kakinada, AP, India
3. DELTA Technology & Management Services Pvt. Ltd., Hyderabad, AP, India

* E-mail of the corresponding author: chaparala_aparna@yahoo.com

Abstract

Brain computer interfaces (BCI) provides a non-muscular channel for controlling a device through electroencephalographic signals to perform different tasks. The BCI system records the Electro-encephalography (EEG) and detects specific patterns that initiate control commands of the device. The efficiency of the BCI depends upon the methods used to process the brain signals and classify various patterns of brain signal accurately to perform different tasks. Due to the presence of artifacts in the raw EEG signal, it is required to preprocess the signals for efficient feature extraction. In this paper it is proposed to implement a BCI system which extracts the EEG features using Discrete Cosine transforms. Also, two stages of filtering with the first stage being a butterworth filter and the second stage consisting of an moving average 15 point spencer filter has been used to remove random noise and at the same time maintaining a sharp step response. The classification of the signals is done using the proposed Semi Partial Recurrent Neural Network. The proposed method has very good classification accuracy compared to conventional neural network classifiers.

Keywords: Brain Computer Interface (BCI), Electro Encephalography (EEG), Discrete Cosine transforms(DCT), Butterworth filters, Spencer filters, Semi Partial Recurrent Neural network, laguarre polynomial

1. Introduction

A Brain Computer Interface (BCI) system records the brain signals through Electro-encephalography (EEG), preprocesses the raw signals to remove artifacts and noise, and employs various signal processing algorithms to translate patterns into meaningful control commands. The purpose of BCI is to control devices like computers, speech synthesizers, assistive appliances and neural prostheses by individual with severe motor disabilities, through brain signals. Signal processing plays an important role in BCI system design, as meaningful patterns are to be extracted from the brain signal.

Figure 1 depicts a generic BCI system (Mason S G et al. 2003). The device is controlled through a series of functional components. Electrodes record signals from the users scalp and convert the signals into electrical signals which are amplified. The artifact processor removes the artifacts from the amplified signals. Feature generator transforms the signals into feature values that are the base for the control of device. The feature generator is generally made up of three steps, signal enhancement, feature extraction and dimensionality reduction. Signal enhancement refers to the preprocessing of the signals to increase the signal-to-noise ratio of the signal. Most commonly used preprocessing methods are Surface Laplacian (Mc Farland D et al. 1998 ; Dornhege G et al. 2004), Independent Component Analysis (ICA) (Serby H et al. 2005), and Principal Component Analysis (Guan J et al. 2005). Feature extraction generates the feature vectors and dimensionality reduction, reduces the number of feature. Thus features useful for classification is identified and chosen while artifacts and noise are eliminated in feature generator step. Genetic algorithm (Peterson D A et al. 2005), PCA (Bashashati A et al. 2005), Distinctive sensitive learning vector quantization (DSLVQ)

(Pfurtscheller et al. 2001) are some of the feature selectors used. The feature translator translates the features into control signals. Various classification algorithms based on linear or nonlinear classification methods are available in literature for classifying the features. Bayesian (Curran E et al. 2004), Gaussian (Millan J R 2004), k-nearest neighbor (Blankertz B et al. 2002), SVM (Peterson D A et al. 2005, MLP (Hung C I et al. 2005) are some of the classifiers used. The BCI transducer translates the brain signals into logical control signals. The logical control signals from the feature translator is converted into semantic control signals in control interface. Device controller converts the semantic control signals into physical control signals which control the device.

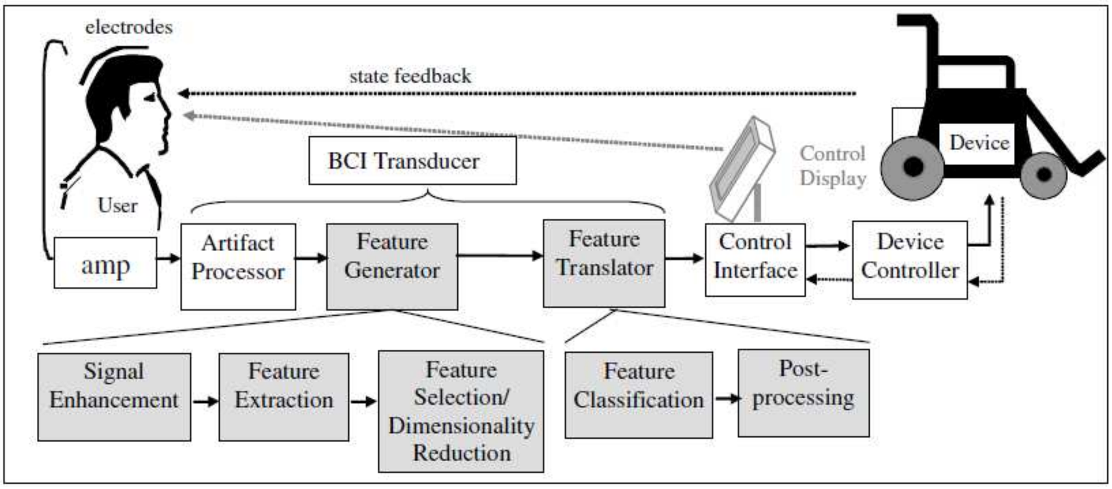

Fig 1: Functional model of a BCI system

In this paper, the proposed BCI system extracts features from the EEG signals using Discrete Cosine transforms. The classification of the signals is done using the Semi Partial Recurrent Neural network with laguarre function in input layer and tanh function in hidden layer with delta learning rule. The paper is organized into four sections, with section I giving introduction to BCI systems, section II concerns with the materials and methods used, section III discusses the result with conclusion in section IV.

2. Materials and Methods

The discrete cosine transform (DCT) is closely related to Karhunen-Loeve-Hotelling (KLH) transform, a transform that produces uncorrelated coefficients (N Ahmed et al. 1983). DCT converts time series signal into basic frequency components. It decomposes the image into set of waveforms. The process of decomposing an image into a set of cosine basis functions is called forward discrete cosine transform (FDCT) and process of reconstructing is called inverse discrete cosine transform (IDCT). Some simple functions to compute the DCT and to preprocess the provided EEG data for BCI system are as follows:

The FDCT (N Ahmed et al. 1983) of a list of n real numbers s(x), x = 0, ..., n-1, is the list of length n is given by:

$$S(u) = \sqrt{2/n}C(u)\sum_{x=0}^{n-1} s(x)\cos\frac{(2x+1)u\pi}{2n} \quad u = 0\dots n \qquad (1)$$

Where C(u) is equal to 1/ square root of 2 for u=0 or is equal to 1 for all other values.

The constant factors are chosen so that the basis vectors are orthogonal and normalized. The inverse cosine transform (IDCT) is given by:

$$s(x) = \sqrt{2/n}\sum_{u=0}^{n-1} C(u)S(u)\cos\frac{(2x+1)u\pi}{2n} \quad \text{x=0… n} \tag{2}$$

Where C(x) is equal to 1/ square root of 2 for x=0 or is equal to 1 for all other values.

A 15 point Spencers filter is used to compute the moving averages of EEG signals and reduce the noise spikes. The obtained data in the frequency domain is filtered using Butterworth filter to remove noise and artifacts in the frequency range of 5-30Hz. The Butterworth filter is a signal processing filter which gives a as flat a frequency response for the pass-band (Giovanni Bianchi et al. 2007). It is one of the most commonly used digital filters and is also called maximally flat magnitude filter. In Butterworth filter, no ripples are formed in the pass-band and are zero on reaching stop-band. It has slower roll-off and more linear phase response when compared to other filters like Chebyshev and elliptic filter. Butterworth filters are advantageously used to filter EEG signals as the pass-band and stop-band are maximally flat, which results in quality output signal for different frequency band.

In a low-pass filter, all low frequency components in the signal are passed through and the high frequency components are stopped. The cutoff frequency divides the pass-band and the stop-band. Thus artifacts in the EEG signal are easily filtered out using a low-pass filter. The low-pass filter can be modified into high-pass filter; when placed in series with others to form band-pass and band-stop filters. The gain G(ω) of an n-order Butterworth low pass filter (S. Butterworth 1930) in terms of transfer function H(s) is given as

$$G^2(\omega) = |H(j\omega)|^2 = \frac{G_0^2}{1+\left(\frac{\omega}{\omega_c}\right)^{2n}} \tag{3}$$

where n is order of filter, ω_c is cutoff frequency and G_0 is the DC gain i.e gain at zero frequency.

The Butterworth filter is used to preprocess the EEG signal to remove high frequency noise or artifacts with cutoff frequencies in a range of 5 - 30 Hz.

The trend of a time series is estimated using a linear filtering operation as follows:

$$\gamma_t = \sum_{r=0}^{q} a_r X_t(n+r) \tag{4}$$

Where a_r is a set of weights and $\sum a_r = 1$ is a moving average or finite impulse response filter.

The 15 point Spencer filters for moving averages is symmetric in nature. It is given as:

$$\frac{1}{360}(3,-6,-5,3,21,46,67,74,67,46,21,3,-5,-6,-3)$$

The maximum and average energy from each channel are computed and used as attributes. Support vector machine is used to reduce the feature vector.

2.1 Partial Recurrent Neural Network

The neural network where input is fed through successive layers of the network to the output is called feedforward networks. The neural network which has a feedback loop is known as Recurrent Neural Network (RNN). If the feedback is in only one of the layers then it is referred to as Semi Partial Recurrent Neural network (SPRNN). The recurrent networks are dynamic in nature as the feedback loops use unit delay elements. PRNN has feedback in any one of the layers only. PRNNs are easier to use than the RNNs. Time is implicitly represented in PRNN. Simple PRNN consists of two-layer network with feedback in the hidden layer as shown in figure 2. The output of the hidden layer at time t is fed back as additional inputs at time t+1, thus the PRNN works in discrete time steps. The proposed PRNN has laguarre function in the input layer and a tanh function in the hidden layer. The tanh function being asymmetric helps to train faster.

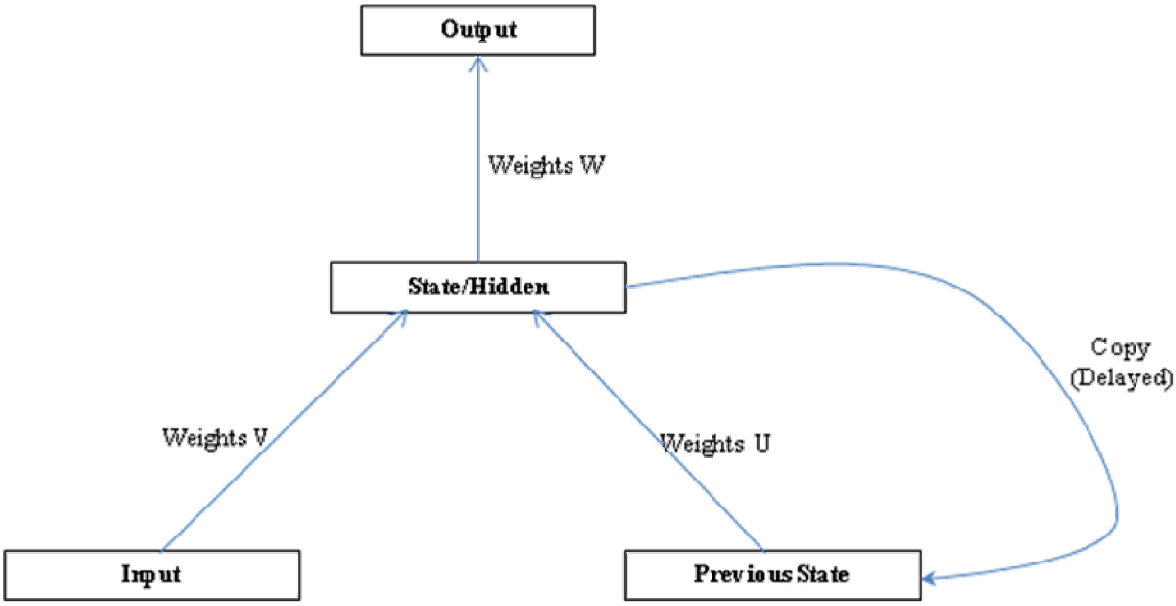

Fig 2: A simple Partial Recurrent Neural network

The output of PRNN when a input vector x is propagated through a weight layer V, and the previous state activation due to recurrent weight layer U,

$$y_j(t) = f(net_j(t)) \tag{5}$$

$$net_j(t) = \left(\sum_i^n X_i(t)v_{ji} + \sum_h^m y_h(t-1)u_{jh} \right) + \theta_j \tag{6}$$

where n is the number of inputs, θ_j is bias, f is output function, m is number of state nodes, and

i, j / h, k denotes the input, hidden and output nodes respectively.

The output of the network with output weights W is,

$$net_k(t) = \sum_j^m y_j(t)w_{kj} + \theta_k \tag{7}$$

The learning of the PRNN at each time step starts with the input vectors fed into the network and it generates an error, the error is backpropagated to find error gradients for each weights and bias. The weights are updated with learning function using the error gradient.

In this paper it is proposed to implement a laguarre function in the input layer to provide details of the input's past memory recursively. The laguarre polynomial is given by

$$L_k(u) = \frac{e^u d^k}{k! dx^k}\left(e^{-u}u^k\right) \quad (8)$$

Where k is the order of the polynomial and u is the value for which the polynomial is to be found. It is proposed to use the first order polynomial ie k=1.

The experimental setup consists of 25 neurons in the input layer, 4 neurons in the hidden layer and two neurons in the output layer (one neuron for each class). The hidden layer and the output layer activation functions used are tanh.

3. Results and Discussion

The dataset used for the work is provided by University of Tübingen, Germany, Dept. of Computer Engineering and Institute of Medical Psychology and Behavioral Neurobiology, and Max-Planck- Institute for Biological Cybernetics, Tübingen, Germany, and Universität Bonn, Germany, Dept. of Epileptology(Thomas Lal et al. 2004) was used. 168 instances of a single patient were used to test the proposed algorithm. 80% of the data was used for training and the remaining for testing. The classification accuracy obtained along with the classification accuracy of MLP neural network is shown in figure 3.

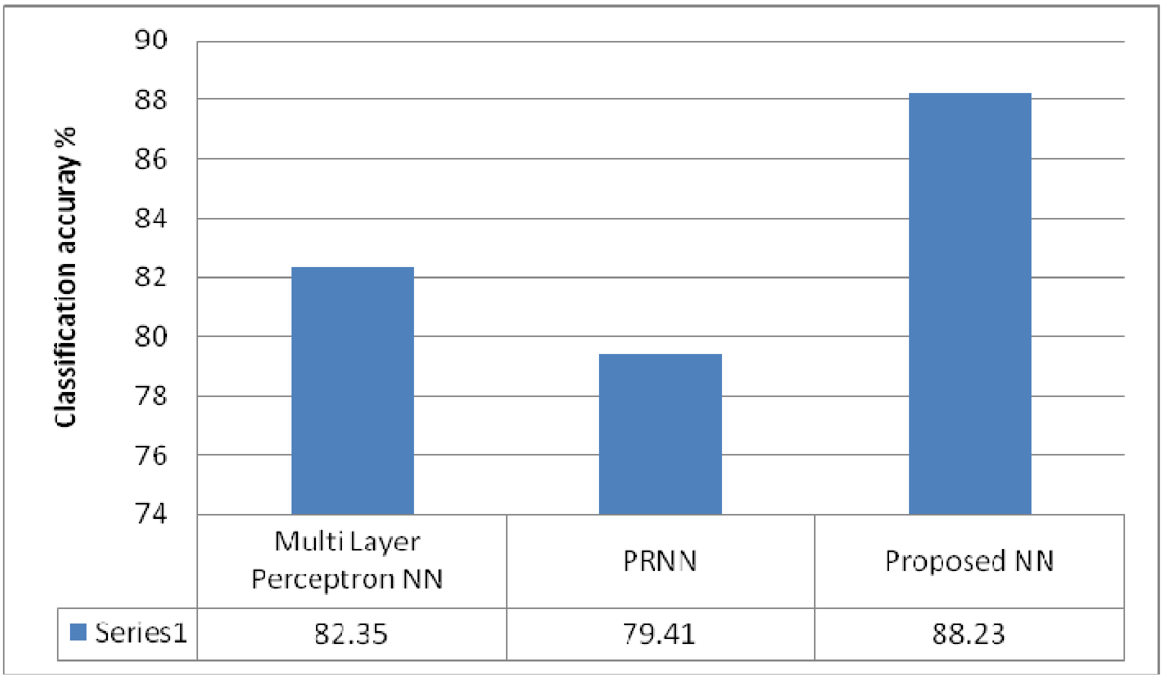

Figure 3 : The classification accuracy of the proposed system

From figure 3, the classification accuracy of the proposed system improves by 10% which is a considerable improvement from regular MLP neural network as well as regular Partial recurrent neural network..

4. Conclusion

In this paper it was proposed to implement a novel neural network based on the partial recurrent neural network with laguarre polynomial in the input layer. Features from the EEG data in time domain was extracted usingdiscrete cosine transform. The frequency of interest was extracted using Butterworth band pass filter. Maximum and average energy for each channel was calculated. The proposed method was implemented using LabVIEW and VC++. The obtained results in the proposed classification method are better than currently available classification algorithms. Further investigation needs to be carried out with other EEG data.

References

Mason S G and Birch G E (2003), “A general framework for brain computer interface design”, IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol.11, 70–85.

McFarland D and Wolpaw J R (1998),”EEG-based communication and control: short-term role of feedback “, IEEE Transactions on. Rehabilitation Engineering, vol. 6, 7–11.

Dornhege G, Blankertz B, Curio G and Muller K R (2004), “Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms”, IEEE Transactions on Biomedical Engineering, vol. 51, 993–1002

Serby H, Yom-Tov E and Inbar G F (2005), “An improved P300-based brain–computer interface”, IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 13, 89–98

Guan J, Chen Y, Lin J, Yun Y and Huang M (2005), “N2 components as features for brain computer interface”, Proceedings of 1st International Conference on Neural Interface and Control (Wuhan, China), 45–9.

Peterson D A, Knight J N, Kirby M J, Anderson C W and Thaut M H (2005), “Feature selection and blind source separation in an EEG based brain computer interface”, EURASIP J. Appl. Signal Process. 19 3128–40

Bashashati A, Ward R K and Birch G E (2005), “A new design of the asynchronous brain computer interface using the knowledge of the path of features”, Proceedings of 2nd IEEE-EMBS Conference on Neural Engineering (Arlington, VA), 101–4

Pfurtscheller G and Neuper C (2001), “Motor imagery and direct brain–computer communication”, Proceedings of IEEE vol.89, 1123–34

Curran E, Sykacek P, Stokes M, Roberts S J, Penny W, Johnsrude I and Owen A M (2004), “Cognitive tasks for driving a brain–computer interfacing system: a pilot study”, IEEE Trans. on Neural Systems and Rehabiitational Engineering Vol. 12, 48–54

Millan J R (2004), “On the need for on-line learning in brain–computer interfaces”, Proceedings of Annual International Joint Conference on Neural Networks (Budapest, Hungary)

Blankertz B, Curio G and Muller K R (2002), “Classifying single trial EEG: Towards brain–computer interfacing”, Advances in Neural Information Processing Systems vol 14, 157–64

Hung C I, Lee P L, Wu Y T, Chen L F, Yeh T C and Hsieh J C (2005), “Recognition of motor imagery electroencephalography using independent component analysis and machine classifiers”, Arificial Neural Networks and Biomedical Engineering, 33, 1053–70.

N. Ahmed, T. Natarajan (1983), “Discrete-Time Signals and Systems”, Reston Publishing Company.

Giovanni Bianchi and Roberto Sorrentino (2007). “Electronic filter simulation & design”, McGraw-Hill Professional. 17–20. ISBN 9780071494670.

S. Butterworth (1930), “Wireless Engineer” , vol. 7, 536–541.

Thomas Lal, Thilo Hinterberger, Guido Widman, Michael Schröder, Jeremy Hill, Wolfgang Rosenstiel, Christian Elger, Bernhard Schölkopf, Niels Birbaumer.(2004), “Methods Towards Invasive Human Brain Computer Interfaces”, Advances in Neural Information Processing Systems (NIPS)

Aparna Chaparala, is working as an Associate Professor in computer science and engineering department of R.V.R. & J.C. College of Engineering, Chowdavaram, Guntur. She has 9 years experience in teaching. She completed her M.Tech in Computer Science & Engineering. She is doing her research in Data Mining area. Presently pursuing Ph.D from J.N.T.U, Hyderabad. She has published 5 papers in international journals.

Dr J.V.R. Murthy is presently working as a professor in the department of CSE at J.N.T.U., Kakinada. He did M.Tech in CSE at IIT. He has over 20 years of teaching experience and 3 years of industrial experience. A Memento of Appreciation was awarded for "good performance and on schedule completion of People Soft HRMS project" by Key Span Energy Corporation, New York. He has more than 15 publications in national and international journals. His interested areas of research include Data Warehousing, data mining and VLDB.

Dr B. Raveendra Babu has obtained Masters degree in Computer Science and Engineering from Anna University, Chennai. He received Ph.D. in Applied Mathematics from S.V University, Tirupati. He is currently leading a Team as Director (Operations), M/s.Delta Technologies (P) Ltd.,Madhapur, Hyderabad. He has 26 years of teaching experience. He has more than 25 international & national publications to his credit. His interested areas of research include VLDB, Image Processing, Pattern analysis and Wavelets.

M.V.P.Chandra Sekhara Rao, is an Associate Professor in the department of computer science and engineering in R.V.R. & J.C. College of Engineering, Chowdavaram, Guntur. He has over 15 years of experience in teaching. He completed his B.E and M.Tech in Computer Science & Engineering. He is doing research in the area of Data Mining. Presently pursuing Ph.D from J.N.T.U, Hyderabad. He has published 5 papers in international journals and presented a paper in international conference.