Energy Computation using DCT for Brain Computer Interface Motor Imagery Classification

[1]Ch.Aparna
Assoc.Professor,
Dept of CSE,
RVR&JC College of Engineering,
Chowdavaram – 522019

[2]Dr J.V.R.Murthy
Professor,
Dept of CSE,
JNTU,
Kakinada – 533004

[3]Dr.B.Raveendra Babu
Director – Operations,
DELTA Technology & Management Services Pvt Ltd,
Hyderabad – 500081

[4]Sri M.V.P.Chandra Sekhara Rao
Assoc.Professor,
Dept of CSE,
RVR&JC College of Engineering,
Chowdavaram – 522019

***Abstract* - Brain computer interface systems are capable to detect and interpret the mental activity into computer interpretable signals giving opportunity for performing computer controlled activities without muscular movement. An challenging area in Brain Computer Interface research is the classification of EEG signals using the raw signals captured which has to undergo some preprocessing, so that the right attributes for classification are obtained. In this paper, we propose to extract the energy component of the EEG signal by processing the data in the frequency domain using Discrete Cosine Transform and application of classification techniques. The proposed method has very good classification accuracy compared to research already carried out in this area.**

Keywords – BCI; Discrete cosine transform(DCT); Butterworth filter; EEG; Naïve Bayes; IB1

I. INTRODUCTION

A brain–computer interface (BCI) is a system that allows communication between the brain and a computer or a robot, without the use of nerves or muscles. By imagining movements of different parts of the body, trained subjects can voluntarily regulate their μ or β rhythms over sensorimotor cortices [3]. Motor imagery (MI) is the state during which the representation of a specific motor action is internally reactivated within the working memory without any overt motor output.

The inputs of the system are usually EEG signals recorded on the scalp's surface. The output is a decision of action among a set of possible ones (for example, a command to a prosthesis). A training session is usually required to build the classifier that allows the decoding of the user's intention [2]. This type of BCI system has become an active research theme due to its relatively robust communication performance and its potential neurophysiological significance for studying the underlying mechanism of motor imagery. In most motor imagery based BCI systems, the identification of brain activity patterns relies on a classification algorithm.

Data from brain signals can be quite high-dimensional, and potentially full of artifacts. Proper application of preprocessing steps can reduce data dimensionality and emphasize portions of the data with discriminative power, thereby reducing computation time and improving classification rates. For accurate classification, it is essential to identify features that need to be used, their properties and how they are used. Feature extraction has been attempted using amplitude values of EEG signals [3], Band Powers (BP) [4], Power Spectral Density (PSD) values [5] [6]. In this paper we investigated the BCI data in the frequency domain using Discrete Cosine Transform and removed unwanted frequencies and noise using Butterworth filter.

For accurate classification for a given BCI system, it is essential to what features need to be used, what their properties are and how they are used. Feature extraction have been attempted using amplitude values of EEG signals [9], Band Powers (BP) [10], Power Spectral Density (PSD) values [11] [12]. In this paper we investigate the BCI data in the frequency domain using Discrete Cosine Transform and remove unwanted frequencies and noise using Butterworth filter.

This paper is organized into five sections. In Section II, the real datasets used in this work are described and the feature extraction method is summarized. In Section III, classification algorithms used in this work are described. Section IV presents the experimental setup and results. Finally, a conclusion summarizing our main findings and our future work is given in Section V.

II. DATASET AND FEATURE EXTRACTION

In this work, Data set 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[7] was used. The experiment task consisted of performing motor imagery of either the left small finger or the tongue in response to a visual cue. The order of the cues was random. Signals were recorded using 8x8 electrodes placed on the contra lateral motor cortex. All recordings were performed with a sampling rate of 1000Hz and after amplification were stored as microvolt values. Each trail was recorded for 3 second duration. The recordings were started only after 0.5

seconds after the visual cue ended to avoid visually evoked potentials. The total number of trials was 278 for one subject.

As frequency domain best represents the essential characteristics, the recorded EEG signal was converted from time domain to frequency domain using Discrete Cosine Transform(DCT). A discrete cosine transform (DCT) expresses a sequence of finitely many data points in terms of a sum of cosine functions oscillating at different frequencies. The discrete cosine transform[13] of a list of n real numbers s(x), x = 0, ..., n-1, is the list of length n given by:

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

$$\text{where} \quad C(u) = 2^{-1/2} \quad \text{for } u = 0$$
$$\qquad\qquad = 1 \quad \text{otherwise}$$

Where each element of the transformed list S(u) is the inner product of the input list s(x) and a basis vector. The constant factors are chosen so that the basis vectors are orthogonal and normalized. The list s(x) can be recovered from its transform S(u) by applying the inverse cosine transform (IDCT):

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

$$\text{where} \quad C(u) = 2^{-1/2} \quad \text{for } u = 0$$
$$\qquad\qquad = 1 \quad \text{otherwise}$$

This equation expresses s as a linear combination of the basis vectors. The coefficients are the elements of the transform S, which may be regarded as reflecting the amount of each frequency present in the inputs. Butterworth filter has been used to remove unwanted features from the transformed signal. Butterworth filter is a filter with a pass-band with no ripple but usually sacrifices some steepness in attenuation. The magnitude of the transfer function for this filter is given by

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

where n is the order of the filter and w_c is the cutoff frequency.

In order to control a BCI, the user must produce different brain activity patterns that will be identified by the system and translated into commands. In most existing BCI, this identification relies on a classification algorithm [8], i.e., an algorithm that aims at automatically estimating the class of data as represented by a feature vector [9].

III. CLASSIFICATION ALGORITHMS

Classification is one of the most popular data mining tasks with a wide range of applications, and a lot of algorithms have been proposed to build accurate and scalable classifiers. Classification is the process of finding a model (or function) that describes and distinguishes data classes or concepts for the purpose of being able to use the model to predict the class of objects whose class label is unknown. The derived model is based on the analysis of a set of training data. (i.e data objects whose class label is unknown). Classification predicts categorical (Discrete, unorder) labels prediction models continuous-valued functions. That is, it is used to predict missing or unavailable numerical data values rather than class labels. In this paper, Naïve Bayesian and IB1 classification techniques were used and applied to EEG data set to evaluate accuracy.

A. Naïve Bayesian Classifier

Naïve Bayesian [10, 11, 12] is a good classification method. It is simple to train, easy to understand, and performs pretty well for real applications. The Naïve Bayes Classifier technique is particularly suited when the dimensionality of the inputs is high. Despite its simplicity, Naive Bayes can often outperform more sophisticated classification methods.
Steps for Building a Bayesian Classifier

- Collect class exemplars
- Estimate class a priori probabilities
- Estimate class means
- Form covariance matrices, find the inverse and
- determinant for each
- Form the discriminant function for each class

The motivation behind the usage of Bayesian classifier has its roots in the regular study of Bayesian probabilistic theory, which is a branch of mathematical probability and allows us to model uncertainty about the aim and outcome of interest by combining experimental knowledge and observational evidences. The Naive Bayesian classifier is fast and incremental, can deal with discrete and continuous attributes, has excellent performance in real-life problems and can explain its decisions as the sum of information gains. However, its naivety may result in poor performance in domains with strong dependencies among attributes. In this paper, the Naive Bayesian classifier is applied successively enabling it to solve non-linear problems also while retaining all advantages of Bayesian classifier.

B. IB1 Classifier

IB1 classifier uses a simple distance measure to find the training instance closest to the given test instance, and predicts the same class as this training instance. If multiple instances are the same (smallest) distance to the test instance, the first one found is used. The IB1 algorithm, is the simplest instance-based learning classification method[13]. IBL algorithms are derived from the nearest neighbor pattern classifier (Cover & Hart, 1967). They are highly similar to edited nearest neighbor algorithms (Hart, 1968; Gates,1972; Dasarathy, 1980), which also save and use only selected instances to generate classification predictions. While several researchers demonstrated that edited nearest neighbor algorithms can reduce storage requirements with, at most, small losses in classification accuracy, they were unable to predict the expected savings in storage requirements. IBL algorithms are instead incremental and their goals include maximizing classification accuracy on subsequently presented instances.

The similarity and classification functions determine how the set of saved instances in the concept description are used to predict values for the category attribute. Therefore, IBL concept descriptions not only contain a set of instances, but also include these two functions.
In IB1 method, the similarity function used here is:

$$\text{Similarity}(x, y) = -\sqrt{\sum_{i=1}^{n} f(x_i, y_i)} ,$$

where the instances are described by n attributes. We define f(xi yi;) = $(x_i,- y_i)^2$ for numeric-valued attributes and $f(x_i, y_i)$ = $(x_i \neq y_j)$ for Boolean and symbolic-valued attributes. Missing attribute values are assumed to be maximally different from the value present. If they are both missing, then $f(x_i, y_i)$ yields 1. IB1 is identical to the nearest neighbor algorithm except that it normalizes its attributes' ranges, processes instances incrementally, and has a simple policy for tolerating missing values.

IV. EXPERIMENTAL SETUP AND RESULT

A program was developed to handle the time series epoch and the time series graph was plotted. Discrete cosine transform was applied to convert the time domain data to frequency domain. Butterworth filter was used as a band pass filter to eliminate frequencies outside the desired range of 5Hz to 30 Hz. The maximum value, the minimum value and the mean was computed for each channel for all the epochs and recorded. The data so created was used as the input to the decision tree classifier.

In this paper, Naïve Bayesian and IB1 classification techniques were applied on BCI data set. WEKA, a Data Mining tool was used to calculate classification accuracy and the results are shown in the Figures 1 and 2 and in the Table 1 .

	Predicted		
Actual		a	b
	a	108	32
	b	37	101

Fig1. Confusion Matrix of Naïve Bayesian Classification

	Predicted		
Actual		a	b
	a	113	27
	b	43	95

Fig2. Confusion Matrix of IB1 Classification

TABLE1. Classification Accuracy

Classification Algorithm	Correctly classified	Incorrectly Classified
Naïve Bayesian	75.18	24.82
IB1	74.82	25.18

The classification accuracy of Naïve Bayesian and IB1 classification techniques are as shown in Fig3.

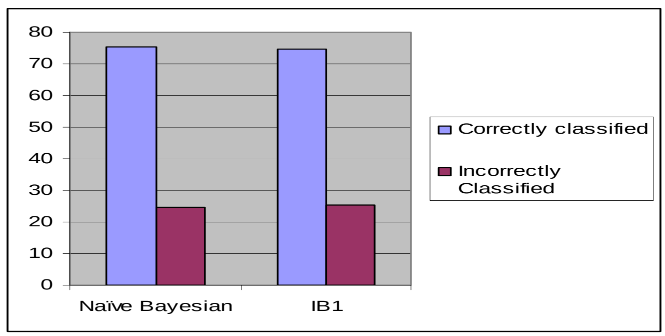

Fig3. Bar chart for Classification accuracy

V. CONCLUSION

In this paper we proposed a frequency domain method for preprocessing the EEG data using Discrete cosine transform and eliminated unwanted frequencies to remove noise and other brain related neural activity using Butterworth filter. The proposed method in preprocessing the data could be used successfully for classification. Different Classification techniques are applied on EEG data set and the results are compared.

REFERENCES

[1] G. Pfurtscheller, D. Flotzinger, and J. Kalcher, “Brain–computer interface—A new communication device for handicapped persons,” J. Microcomput. Appl., vol. 16, pp. 293–299, 1993.

[2] N. Birbaumer, A. R. Murguialday, and L. Cohen, “Brain-computer interface in paralysis,” *Curr. Opin. Neurol.*, vol. 21, no. 6, pp. 634–638, Dec. 2008.

[3] M. Kaper, P. Meinicke, U. Grossekathoefer, T. Lingner, and H. Ritter. Bci competition 2003- data set iib: support vector machines for the p300 speller paradigm. IEEE Transactions on Biomedical Engeneering, 51(6):1073{1076, 2004.

[4] G. Pfurtscheller, C. Neuper, D. Flotzinger, and M. Pregenzer. Eeg-based discrimination between imagination of right and left hand movement. Electroencephalography and Clinical Neurophysiology, 103:642{651, 1997.

[5] S. Chiappa and S. Bengio. Hmm and iohmm modeling of eeg rhythms for asynchronous bci systems. In European Symposium on Artificial Neural Networks ESANN, 2004.

[6] J. R. Millan and J. Mourino. Asynchronous BCI and local neural classi_ers: An overview of the Adaptive Brain Interface project. IEEE Transactions on Neural Systems and Rehabilitation Engineering, Special Issue on Brain-Computer Interface Technology, vol. 11. no . 2, 2003.

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

[8] D. J. McFarland, C. W. Anderson, K.-R. Muller, A. Schlogl, and D. J. Krusienski. Bci meeting 2005-workshop on bci signal processing: feature extraction and translation. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 14(2):135 { 138, 2006.

[9] R. O. Duda, P. E. Hart, and D. G. Stork. Pattern Recognition, second edition. WILEYINTERSCIENCE, 2001.

[10] Duda R. and Hart P. Pattern Classification and Scene Analysis. New York: John Wiley & Sons, 1973.

[11] Join G. H. Enhancements to the Data Mining Process. Ph.D. Thesis, computer Science Dept, Stanford University, 1997.

[12] Domingos P. and Pazzani M. Beyond Independence: Conditions for the Optmality of the Simple Bayesian Classifier. In Proc. 13th Intl. Conf. Machine Learning, p105-112, 1996.

[13] D.W. Aha, D. Kibler, and M.K. Albert, ªInstance-Based Learning Algorithms, Machine Learning, vol. 6, pp. 37-66, 1991.

[14] Jonathan R. Wolpaw, Nils Birbaumer, Dennis J. McFarland, Gert Pfurtscheller, and Theresa M. Vaughan. Brain-computer interfaces for communication and control. Clinical Neurophysiology, pages 767–791, MAR 2002.