ISSN:0975-9646

V.Baby Deepa, et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (6) , 2011, 2830-2834

Investigation of Classification Accuracy using Fast Hartley Transform for Feature Extraction

V.Baby Deepa, Dr. P. Thangaraj

***Abstract-* In the first step, the EEG signal from each electrode is converted to the frequency domain using the Fast Hartley Transform. Artifacts in the transformed signal using the frequency domain were removed using a band pass Chebyshev filter such that only frequencies in the range of 5-15 Hz is retained. The minimum energy, maximum energy and the average energy is computed. The computed features are trained and classified using AD Tree, BayesNet and Instance based learners.**

***Keywords:* EEG, FHT, Chebyshev Filter, AD Tree, BayesNet, Instance based learners**

1. INTRODUCTION

A BCI system works by recording the brain signals and applying machine learning algorithms to classify the brain signals and performing a computer controlled action. The most widely used method for signal acquisition is the electro-encephalography (EEG). The reason for the popularity of EEG is due to the non-invasive way of acquiring the brain signals and also it is safe, easy and cheap when compared to other methods .

Most of the existing application prototypes of BCI use EEG signals. Prototypes like “Thought Translation Device” [1] which allows paralyzed patient to write sentences, spelling system [2], “virtual keyboard” based on motor activity [3] are all EEG based BCI systems. Thus the role of EEG processing is crucial in the development of BCI.

The EEG [4] [5]signals is made up of cluster of features. It is imperative to extract the functional features from the EEG data. Identifying and extracting good features from the signals is a crucial step in the design of BCI [6]. Studies [7] show that if the features extracted from EEG are not relevant and the neurophysiological signals employed are not well defined, then the accuracy of the classification algorithm identifying the class of these features, i.e., the mental state of the user is greatly reduced.

As a result, the correct recognition rates of mental states will be very low, which lowers the usability of the BCI or it may even be impossible to use by the user.

2. EEG DATA

The human nervous system communicates through electrical impulses; the functional activity of the brain is reflected on the scalp as variation of the surface potential distribution. Due to the electrical nature of the surface potential variation, it is possible to measure the variation by fixing an array of electrodes to the scalp.

The electrodes measure the voltage between the fixed points, which are then filtered, amplified and recorded as EEG data. The international 10-20 system of electrode placement is the most widely used method of placing the electrodes at specific intervals along the scalp.

Figure 2.1 shows the placement of electrodes according to the 10-20 system. The letter identifies the lobe and the number the hemispheric location.

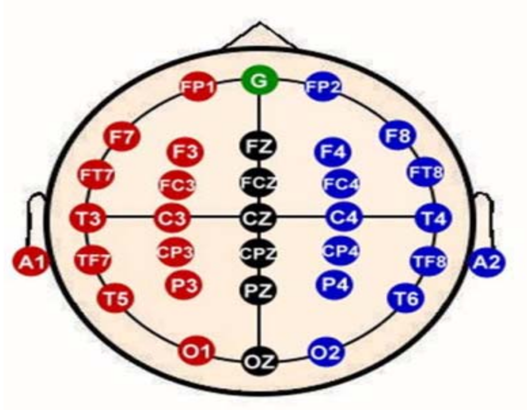

Figure 2.1: Electrodes placement of 10–20 system.

The letters used are: F: Frontal lobe. T: Temporal lobe. C: Central lobe. P: Parietal lobe. O: Occipital lobe. "Z" refers to an electrode placed on the mid-line.

The voltage potentials produced by the brain are at microvolt level, the electrodes conduct this voltage to amplifiers that magnify the signals thousand times. This EEG collected as electrical patterns from the scalp are digitalized and stored as raw records.

The analysis of the EEG signals is complex as large amount of data is received from each electrode. Brain waves are not emitted alone, but the state of brain makes one frequency range more pronounced than the others.

The main problem in automated EEG analysis, as in BCI, is the detection of the different kinds of interference waveforms. These interference waveforms are termed artifacts; artifacts are included in the EEG signal during the recording. The main sources of artifacts are:

- EEG Equipment
- Electrical interference external to recording system
- The leads and the electrodes
- Normal electrical activity from heart, eye blinking, eye movement.

Artifacts can be easily detected on visual inspection but in automated analysis these cause serious misclassification. Recognition and elimination of the artifacts is crucial for the development of practical BCI systems. The eyeblink and eyeball movement are the most severe of the artifacts.

Figure 2.2 shows a typical artifact in an EEG signal and figure 2.3 shows eyeblink artifact in EEG waveform recorded by a forehead electrode.
Raw data cannot be used as input of classification algorithm. It is necessary to remove artifacts and extract good features so as to maximize the performance of the system.
The choice of a good pre-processing and feature extraction method has more impact on the final performance rather than the selection of a good classification algorithm.

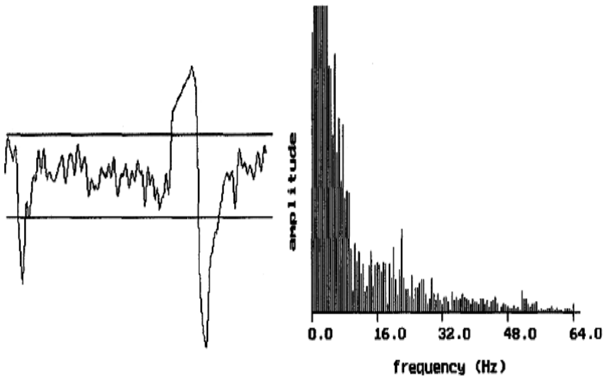

Figure 2.2: Normal EEG from an adult showing muscle and eye blink artifacts

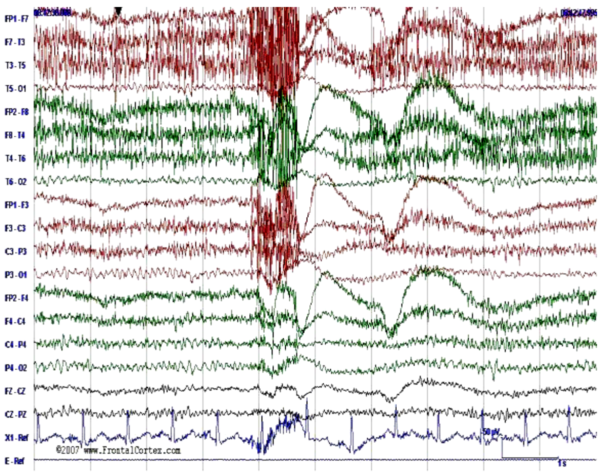

Figure 2.3: Eyeblink artifact in EEG waveform recorded by forehead electrode

Methodology

The IV A dataset used in the brain computer interface. It consists of recordings from five healthy subjects who sat in a chair with arms resting on armrests. Visual cues indicated for 3.5 s which of the following 3 motor imageries the subject should perform: (L) left hand, (R) right hand, (F) right foot. The presentation of target cues was intermitted by periods of random length, 1.75 to 2.25 s, in which the subject could relax. Given are continuous signals of 118 EEG channels and markers that indicate the time points of 280 cues for each of the 5 subjects (aa, al, av, aw, ay). Subject aa was used in our study.

3. FEATURE EXTRACTION USING FAST HARTLEY TRANSFORM

The regular Hartley transform's kernel is based on the cosine-and-sine function, defined (Khayat) as:

$$cas\,(\nu t) = cos(\nu t) + sin(\nu t)$$

Hartley transform compared to Fourier transforms is a real function. The Hartley transform pair can be defined as follows:

$$H(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)\,cas(vt)\,dt \qquad (1)$$

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H(v)cas(vt)dv \qquad (2)$$

A very important property of Hartley Transform is its symmetry:

$$H\{f(t)\} = H(v),\ H\{H(t)\} = f(v) \quad (3)$$

This has the advantage of using the same operation for computing the transform and its inverse. Another important feature is that the transform pairs are both real which provides good computational advantages for Hartley Transform (HT) over the Fourier Transform (FT).

Many of the familiar complex relations in the Fourier domain have very similar counter parts in the Hartley domain. Let F (ω) and H (ν) be the FT and HT of a function f(t) the n it is to verify the following:

$$\begin{aligned} &af(t) + bg(t) \Leftrightarrow aF(\nu) + bG(\nu), f(t/a) \\ &H(\nu) = \left[\Re(F(\omega)) - \Im(F(\omega))\right]_{\omega=\nu,} \end{aligned} \qquad (4)$$

$$\begin{aligned} &X(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x(n) cas\left(\frac{2\pi nk}{N}\right), k = 0,1,2,....N-1, \\ &F(\omega) = \left[\varepsilon(H(\nu)) - o(H(\nu))\right]_{\nu=\omega,} \end{aligned} \qquad (5)$$

where, $\Re$, $\Im$, ε, O denote real, imaginary, even and odd parts. Other properties in the Hartley domain are:

$$\begin{aligned} &af(t) + bg(t) \Leftrightarrow aF(\nu) + bG(\nu), \\ &f(t/a) \Leftrightarrow |a| F(\nu),\ f(-n) \Leftrightarrow F(-k), \\ &f * g(t) \Leftrightarrow \frac{1}{2}\begin{bmatrix} F(\nu)G(\nu) + F(-\nu)G(\nu) + \\ F(\nu)G(-\nu) + F(-\nu)G(-\nu) \end{bmatrix}, \\ &\frac{d}{dt} f(t) \Leftrightarrow -\nu F(-\nu), \int f(t)dt \Leftrightarrow -\frac{1}{\nu} F(-\nu), \\ &cas(at) \Leftrightarrow \sqrt{2\pi}\,\delta(\nu - a), \\ &f(t)cos(\nu_0 t) \Leftrightarrow \frac{1}{2}\left[F(\nu - \nu_0) + F(\nu + \nu_0)\right] \end{aligned} \qquad (6)$$

HT's discrete formulation DHT is given by:

$$X(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x(n)\text{cas}\left(\frac{2\pi nk}{N}\right), k = 0,1,..N-1$$

Which is applied to the discrete-time function x (n) with period N. The properties of the DHT are similar to those of the Discrete Fourier Transform (DFT) and Fast Hartley Transform (FHT) [8] which is similar to the familiar Fast Fourier Transform (FFT). Some of the properties of DHT are listed:

$$af(n) + bg(n) \Leftrightarrow aF(k) + bG(k)$$

$$f(-n) \Leftrightarrow F(-k)$$

Obtaining energy values using regular Fast Hartley Transform introduces artifacts associated with EEG signal measurement. To reduce the artifacts we propose a normalization of the obtained energy using Gaussian methods on the Fast Hartley Transform. The normalization provides the benefit to the system performance by desensitizing the system to the signal amplitude variability.

4. CHEBYSHEV FILTER

Chebyshev filters are used to separate one band of frequencies from another. The EEG energy was computed in the 5-15 Hz region to primarily capture the Beta waves in the EEG signal which is closely linked to motor behavior and is generally attenuated during active movements. Chebyshev filter was primarily used for its speed. Chebyshev filters are fast because they are carried out by recursion rather than convolution. The design of these filters is based on the z-transform.

5. MEASURING CLASSIFICATION ACCURACY OF THE PROPOSED FEATURE EXTRACTION SYSTEM USING ALTERNATING DECISION TREE

An Alternating Decision Tree (AD Tree) [9] is a machine learning rule for classification and is a generalization of decision tree that have connections to boosting. It consists of decision nodes and prediction nodes. In each node the decisions are based on the predicate condition. AD trees always have prediction nodes as both root and leaves.

An epoch is classified through AD Tree by following all paths for which all decision nodes are true and summing any prediction nodes that are traversed. This is different from binary classification trees such as Classification and Regression Tree (CART) or C4.5 in which an instance follows only one path through the tree.

The AD Tree algorithm's fundamental element is the rule which consists of a precondition, condition and two scores. A condition is a predicate which is in the form of attribute comparison value. The tree structure can be derived from a set of rules by making note of the precondition that is used in each successive rule. Using 10 fold cross validation the tree is constructed using 21 leaves.

6. MEASURING CLASSIFICATION ACCURACY OF THE PROPOSED FEATURE EXTRACTION SYSTEM USING BAYES NET

Bayesian networks (BNs), also known as belief networks, is a probabilistic graphical models (GMs). The knowledge about uncertain domain is presented in graphical structures. The nodes in the graph represent random variable and edges between the nodes correspond to probabilistic dependencies between the variables. Statistical and computational methods are used to compute the conditional dependencies. Hence, Bayesian Networks combine principles from graph theory, probability theory, computer science, and statistics.

BNs correspond to another graphical model structure known as a directed acyclic graph (DAG). DAGs are popularly used in the statistics, machine learning, and artificial intelligence. The advantages of BNs are that it is both mathematically thorough and easily understandable. Joint Probability Distribution (JPD) of random variables are effectively represented and easily computed through BNs.

The structure of a DAG is made up of set of nodes (vertices) and set of directed edges. The random variables are represented as nodes labeled by the variable names and the dependence between the variables is represented by the edges. Thus, the net is represented in form of circles and arrows as shown in figure 6.1. An edge (or arrow) from node *Xi* (circle representing variable) to node *Xj* represents a statistical dependence between the variables. For instance, the arrow connecting X_1 to X_2 represents the statistical dependence between X_1 and X_2, and the value of X_2 depends upon the value of X_1. Node X_1 is referred to as parent node of X_2 and conversely, X_2 is referred to as child of X_1.

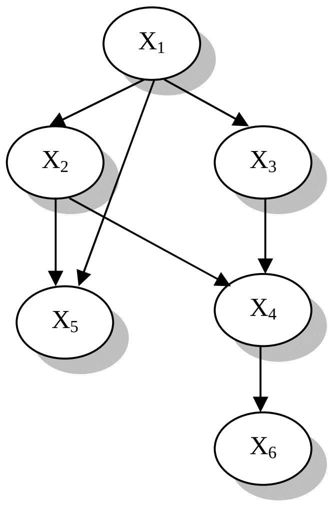

Figure 6.1 Simple Bayesian Network

A BN shows simple conditional independence statement. Each variable is independent of its non descendents in the graph. This characteristic is useful in reducing the number of parameters required to characterize the JPD of the variables, thus the posterior probabilities are efficiently calculated for given conditions. The DAG structure represents the qualitative part of the model, and the quantitative parameters which are conditional probability distribution (CPD) are represented in a table as shown in figure 6.1. The CPD of a node depends only on the values of the parent node. The table lists the local probability that the child node takes for each combination of values of its parents.

7. MEASURING CLASSIFICATION ACCURACY OF THE PROPOSED FEATURE EXTRACTION SYSTEM USING INSTANCE BASED LEARNERS

Instance Based Learner (IBL) classifier uses the class of the nearest k training instances for the class of the test instances. IBL uses a weighted overlap of the feature values of test instance and a memorized example. IBL uses the advantage of global feature weights along with individual feature distance metric.

8. RESULTS AND DISCUSSION

The measured parameters of the three classifiers tested on the proposed feature extraction method are shown in Table 8.1.The confusion matrix is given in Table 8.2.

Table 8.1: Measured Parameters

	Bayes Net	AD Tree	IBL
Classification Accuracy (%)	61.3	58.33	51.78
RMSE	0.497	0.54	0.694
Sensitivity	0.62	0.56	0.49
Specificity	0.61	0.78	0.48

Table 8.2: Confusion Matrix

	Bayes Net		AD Tree		IBL	
Confusion Matrix	Hand	Foot	Hand	Foot	Hand	Foot
Hand	40	40	45	35	45	35
Foot	25	63	35	53	46	32

The classification accuracy line plot is detailed in figure 8.1 with sensitivity and specificity in figure 8.2.

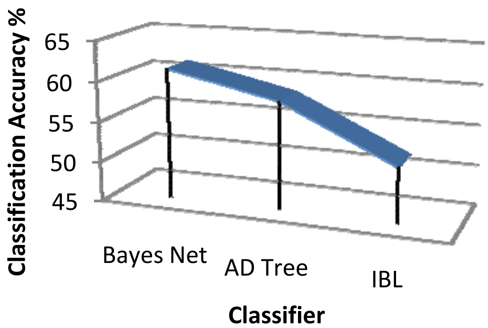

Figure 8.1 : Classification Accuracy of Bayes net, AD tree and IBL

From figure 8.1 it is seen that Bayes net performance is better than AD tree which implies the relationship between the class label and prior probability. Though the computation time is the lowest in IBL, the classification accuracy is the least among the three classifiers. Since the experiment was conducted as a two class problem, specificity and sensitivity helps in identifying the Figure 8.2 shows the sensitivity and specificity plot.

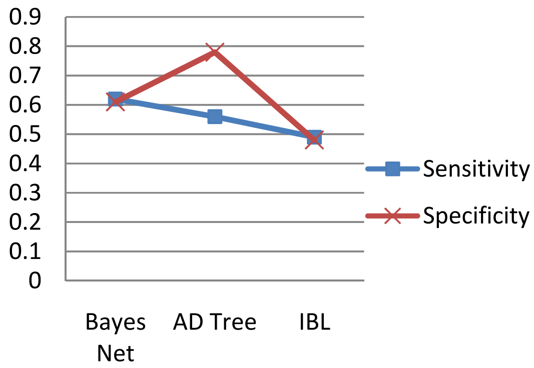

Figure 8.2 Sensitivity and specificity.

9. CONCLUSION

In this paper the proposed feature extraction technique was used to compare the classification accuracy of three classifiers namely Bayes net, AD tree and IBL. The average classification accuracy obtained was 57.13 %. Further investigation is needed to improve the classification accuracy. Support Vector Machines to reduce features and measure the classification accuracy.

REFERENCES

[1]. Birbaumer N., Kubler A., Ghanayim N., Hinterberger T., Perelmouter J., Kaiser J., Iversen I., Kotchoubey B., Neumann N. and Flor H.(2000) 'The Thought Translation Device (TTD) for completely paralyzed patients'. In IEEE Transactions on Rehabilitation Engineering, volume 8, pages 190–193.
[2]. Donchin E. Spencer K.M. and Wijensinghe R. (2000) 'The mental prosthesis: Assessing the speed of a P300-based brain-computer interface'. In IEEE Transactions on Rehabilitation Engineering, volume 8, pages 174–179 .
[3]. Scherer R., Muller G.R., Neuper C., Graimann B. and Pfurtscheller G.(2004) 'An asynchronously controlled EEG-based virtual keyboard improvement of the spelling rate'. In IEEE Transactions on Biomedical Engineering, volume 51(6), pages 979–984.
[4]. Casson , Alexander J. and Rodriguez-Villegas Esther (2008) 'On data reduction in EEG monitoring: Comparison between ambulatory and non-ambulatory recordings' Engineering in Medicine and Biology Society. . 30th Annual International Conference of the IEEE Digital Object Identifier: 10.1109/IEMBS.2008.4650553. Page(s): 5885 – 5888
[5]. Jahankhani P. Revett K. and Kodogiannis V.(2007) ' Data Mining an EEG Dataset With an Emphasis on Dimensionality Reduction' Computational Intelligence and Data Mining,. IEEE Symposiumon on Digital Object Identifier: 10.1109/CIDM.2007.368903 Page(s): 405 - 412 .
[6]. Choi Kyoung ho Sasaki M. (2002) 'EEG data classification with several mental tasks', Systems, Man and Cybernetics, IEEE International Conference on Volume:6.
[7]. Nielsen K.D. Cabrera A.F. Omar Feixdo Nascimento (2006) 'EEG based BCI-towards a better control' Brain-computer interface research at aalborg university, Neural Systems and Rehabilitation Engineering, IEEE

Transactionson Volume: 14,Issue: 2 Digital Object Identifier: 10.1109/TNSRE.2006.875529 , Page(s): 202 – 204.
[8]. Bracewell R.N.(1995) 'Computing with the Hartley Transform' Computers in Physics 9 (4), 373–379.
[9]. Pfahringer B., Holmes G., and Kirkby R.(2001) 'Optimizing the induction of alternating decision trees'. Proceedings of the Fifth Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining, pp: 477-487.

AUTHORS

V.Baby Deepa, received her Bachelor's and Master's in Computer Science from Barathidasan University, Trichy and did her M.Phil. as well in the same university. She has 12 years of teaching experience and is Assistant professor in the faculty of Software Engineering, she is serving as the head for the same faculty in M.Kumarasamy College of Engineering, Karur. She has presented more than 15 papers on various topics including national, international conference and journals. She is a research scholar of Anna University Chennai and her research area is Fuzzy and Data Mining.

Dr.P.Thangaraj, did his graduation and post graduation in Mathematics at Madras University.He completed his M.Phil degree in the year 1993 from Bharathiar University. He completed his research work on Fuzzy Metric Spaces and awarded Ph.D degree by Bharathiar University.He completed the post graduation in Computer Applications at ICNOU in 2005. He completed his Master of Engineering degree in Computer Science in the year 2007 at Vinayaka Missions University. Currently he is a professor and Head of Computer Science and Engineering in Bannariamman College of Engineering and Technology. His current area of research interest is in Fuzzy Metric Spaces and Data Mining.