Wavelet Packet Sub-band Based Classification of Alcoholic and Controlled State EEG Signals

D. Puri[1], R. Ingle[2], P. Kachare[1], M. Patil[1] and R Awale[2]

[1] Department of Electronics and Telecommunication Engineering, RAIT, Navi Mumbai
[2] Department of Electrical Engineering, VJTI, Mumbai
{digspuri@gmail.com; rringle@vjti.org.in; pramod.kachare1991@yahoo.com; mukesh.rait@gmail.com ;rnawale@vjti.org.in}

Abstract: Electro-encephalogram (EEG) is one of the most practiced signals in brain computer interface systems. Several distinct EEG patterns have been analyzed in identifying physiological and psychological states. Work presented here focuses on classification of EEG patterns for alcoholic and controlled states. Third level sub-band energy features are generated for either classes using multi-resolution wavelet packet transformation. A well-known support vector classifier is employed to segregate these features in two well defined classes. Experimental results show significant improvement over wavelet tree feature extraction. Cross-validation tests confirm the greater classification accuracy for proposed technique.

Keywords: *Electroencephalogram, wavelet packet transform, sub-band decomposition, support vector machine.*

1 Introduction

Biological signals such as Electro-Encephalogram (EEG), Electromyogram, Electrocardiogram, etc contain information that is very important in several clinical applications. Despite rapid advances of different neuroimaging techniques, EEG recordings continue to play a vital role in the diagnosis of various psychological states [1]. These signals contain large amount of information about various neural activities [2].

EEG signals are brain waves captured using various electrodes placed on human scalp. EEG signals are highly non-linear, aperiodic and time varying responses having very low frequency and small amplitude [5]. Classification of these signals relies on recognizing the region of brain depending upon the seizure pattern. Different methods have been proposed to classify EEG signals for various applications. Performance of such classification systems mainly depend on the set of features considered for the classification [4]. Most of the existing schemes for extracting spontaneous EEG features are based on Auto-Regressive (AR) models, Fast Fourier Transform (FFT), Short-Time Fourier Transform (STFT) and Wavelet Transform (WT).

AR or FFT models can neither capture transient features nor the time-frequency information in a given signal [5]. STFT alleviates such time-frequency conflict by localizing both time and frequency information over uniformly spaced moving window over entire range of frequencies. WT steps further to adapt window size according to frequency. If basic wavelet function has a finite duration, frequency information obtained from WT seems to be localized in time. Therefore, for non-stationary transient signals such as EEG, WT is superior to FFT as well as STFT [3]. In this paper, we focus on special class of WT, the Wavelet Packet Transform to extract sub-band based features of EEG signals. Energy values for these sub-bands act as features which are further classified using Support Vector Machine.

Rest of the paper is organized as follows: Section II describes acquisition of EEG signals and data description. Section III talks about wavelet packet transform (WPT) used for feature extraction. Section IV includes details of Support vector machine (SVM). Experimental results are discussed in section V. Finally, Section VI concludes the work.

2 Acquisition of EEG Signals

Recording or acquisition of EEG signal is a subject of research by itself. In this paper, all the experiments were conducted on a publicly available database, maintained by UCI machine learning repository [9, 10].

ATLANTIS PRESS

Signals have been taken from 40 healthy (not suffering from any neurological disorders) novices during alcoholic and controlled states, when looking at the pictures of objects, chosen from 1980 Snodgrass and Vander wart picture set [11]. The database is segregated in two sets of Alcoholic and Controlled states abbreviated as 'A' and 'C', respectively. Each set is made of 40 EEG recordings, each with 64 electrodes placed on the human scalp according to International 10-20 system of electrode placement. Fig.1 shows a sample EEG signals from class A. Each sample is one minute long recording with 256Hz sampling for faithful reconstruction. The experiment has been repeated for each subject for 60 trials each of 1 second. Each recording of EEG signal consists of 15360 * 64 samples with voltage in range of microvolts (μV).
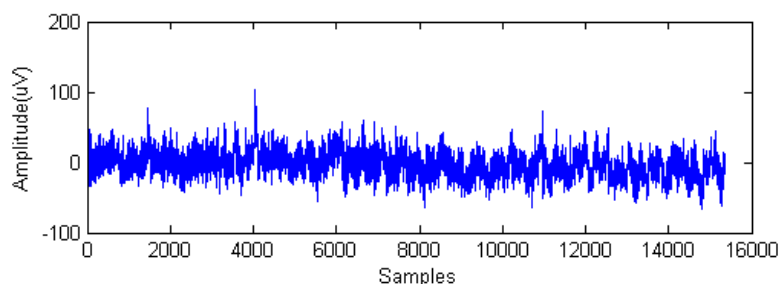

Fig. 1. Time domain sample of alcoholic EEG signal

3 Wavelet Packet Transform

Wavelet is one of the most promising and widely implemented transforms to solve time-frequency conflict. Signal of our interest is digital in nature and hence, we will restrict our discussion to Discrete Wavelet Transform. Wavelet provides multi-resolution analysis for variety of non-linear EEG signals. It explores multiple frequency levels at different resolutions with the help of scaling and dilation parameters, respectively. In this paper, a special class of WT known as Wavelet Packet Transform (WPT), recursive decomposition of signal into low and high frequency sub-bands. Fig. 2 shows diagrammatic representation third level WPT decomposition of a signal. Each decomposition level consist of 2 digital filters with response $g[k]$ and $h[k]$ which are a pair of conjugate mirrors [1]. Dilation and scaling functions decompose the signal $x[n]$ into low-frequency band i.e. coarse approximation and high-frequency bands i.e. detail information [7]. A wavelet packet transform can be represented as;

$$\psi_{j,k}^{i}(t) = 2^{-j/2}\psi^{i}(2^{-j}t - k) \tag{1}$$

where, j is dilation factor and k is scaling factor. $i = 1,2,\ldots,n$, and n is the level of decomposition in wavelet packet tree [12]. Here ψ is called as a mother wavelet and ψ^{i} is obtained by the following recursive relationships,

$$\psi^{2i} = \frac{1}{\sqrt{2}}\sum_{-\infty}^{\infty} h[k]\psi^{i}\left(\frac{t}{2} - k\right) \tag{2}$$

$$\psi^{2i+1} = \frac{1}{\sqrt{2}}\sum_{-\infty}^{\infty} g[k]\psi^{i}\left(\frac{t}{2} - k\right) \tag{3}$$

The wavelet packet coefficients $c_{j,k}^{i}$ corresponding to the signal $f(t)$, can be obtained as,

$$c_{j,k}^{i} = \int_{-\infty}^{\infty} f(t)\,\psi_{j,k}^{i}(t)\,dt \tag{4}$$

The wavelet packet component of signal at a particular sub-band can be obtained as,

$$f_{j}^{i}(t) = \sum_{k=-\infty}^{\infty} c_{j,k}^{i}\,\psi_{j,k}^{i}(t) \tag{5}$$

The extracted wavelet coefficients provide a compact representation that shows the energy distribution of the signal in time and frequency [12].
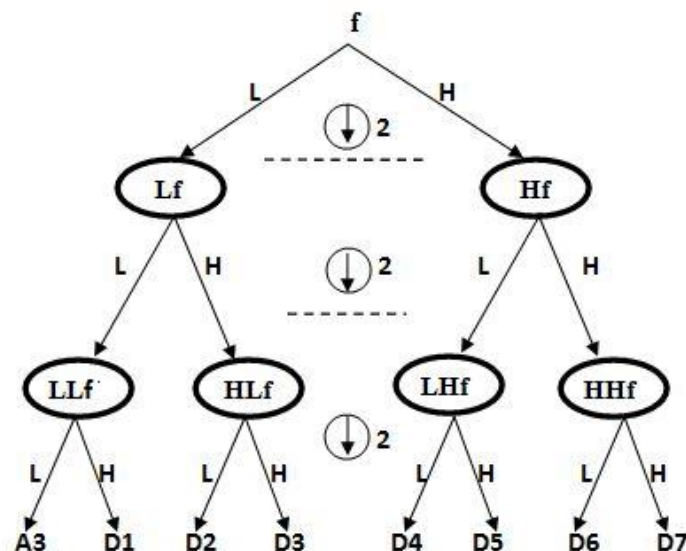
Fig. 2. Proposed Wavelet Packet decomposition scheme

EEG signals are the superposition of different physical activities occurring at different times. Hence, WPT is more suitable for feature extraction of EEG signals. Selection of appropriate wavelet and its decomposition level are two important criterions in modelling of EEG signals. The number of levels depends on the dominant frequency in the signal [1]. Daubechies wavelet of order 1 made more suitable for EEG signal variation. It has smoothening feature which can detect small variations in the EEG signal. For experimentation purpose, we have selected third level decomposition, which transforms EE signal into 8 WPT sub-bands, one approximate sub-band ($A3$) and seven detail sub-bands ($D1 - D7$) shown in Fig. 2. For each sub-band average energy is calculate and normalized using average energy of original signal. Hence, we obtain normalized sub-band energy coefficients sub-band characterizing EEG signal in WPT feature domain. Extracted features are further applied to Support Vector Machine, hard classifier, for training and testing of alcoholic state discrimination system.

4 Support Vector Machine

Support Vector Machine (SVM) is a robust statistical classification technique, primarily employed in case of data scarcity. In our case we have only 80 recordings in database (40 recording per class). Most of the machine learning algorithms are helpless due to such small number of training vectors while, SVM is best known remedy. It is a mathematical analysis to solve n-dimensional optimization using (n-1) dimensional hyper-plane [13]. Selection of appropriate hyper-plane maximizes the average deviation for both the classes. The feature vectors that govern separating hyper-plane are termed as support vectors. For accurate classification only support vectors are required while rest of the dataset becomes redundant.

Mathematically, SVM can be considered as ($f: X \rightarrow Y, x \in X, y \in \{-1, 1\}$) [3]. Hence, the subjective function (f) can be written as,

$$f(x) = \boldsymbol{w}.x + b \tag{6}$$

where, $\boldsymbol{w}$ is the weight matrix. SVM optimizes above equation under the objective function [3],

$$\arg\min \quad \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{m} \zeta_i \tag{7}$$

Above equation generates coefficients of hyper-plane which maximizes separation between $\boldsymbol{w}.x + b = -1$ and $\boldsymbol{w}.x + b = 1$. This places features vectors corresponding positive class on one of the sides of the hyper-plane, while feature vectors corresponding to negative class resides on the opposite side of the hyper-plane. Fig. 3 gives an example to illustrate the concept of the formulation of the SVM. Each training class tolerance is provided in terms of slack variable ($\zeta_i, i = 1, \dots, m$). User controls the slack using positive cost (C). Trade off between margin maximization and slack minimization is controlled using slack penalty (C). The approximate cost value is obtained using various cross validation techniques.
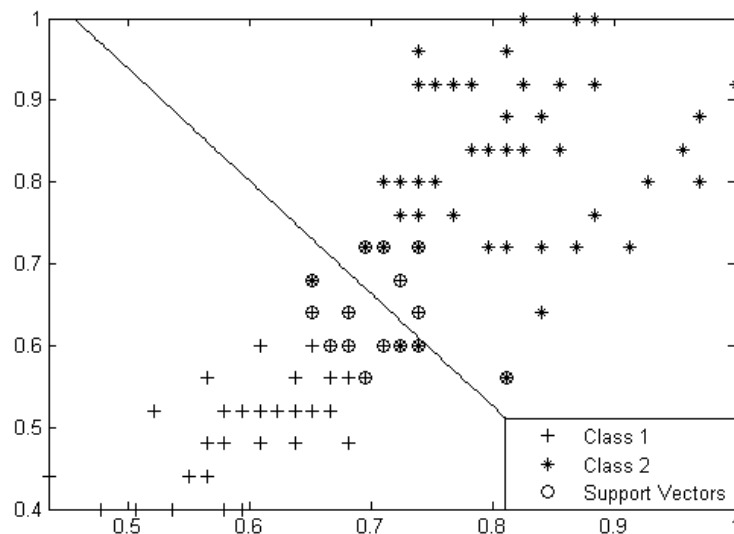
ATLANTIS PRESS

Fig. 3. SVM hyperplane generation to maximize class margin

Table 1. Statistical Parameters for alcoholic state EEG signals

	A1	D1	D2	D3	D4	D5	D6	D7
σ	17.454	6.922	7.148	6.803	7.128	5.245	6.020	5.364
μ	0.386	-0.322	-0.598	0.271	0.284	0.334	0.214	-0.556
β_1	0.296	-0.129	-0.508	0.368	-0.472	-0.108	-0.201	-0.281
β_2	15.081	16.127	4.8110	13.798	5.145	5.462	5.721	6.058
% Energy	87.004	6.787	1.470	1.979	0.700	0.933	0.552	0.571
Max	73.037	48.240	48.723	42.515	57.925	40.313	37.885	32.005
Min	-77.584	-40.655	-44.707	-44.4565	-38.921	-39.461	-41.002	-41.230

Table 2. Statistical Parameters for controlled state EEG signals

	A1	D1	D2	D3	D4	D5	D6	D7
σ	17.658	5.590	6.519	5.401	7.138	5.106	6.020	4.967
μ	-2.898	-0.469	-0.747	-0.538	-0.104	0.146	0.146	-0.554
β_1	-0.077	-0.197	-0.238	-0.143	0.196	0.151	-0.178	-0.019
β_2	5.694	12.972	3.755	5.177	5.125	7.321	6.122	4.919
% Energy	94.551	3.101	0.478	1.292	0.107	0.230	0.029	0.209
Max	69.369	35.843	46.913	32.037	35.778	30.103	37.885	32.135
Min	-111.09	-31.1382	-37.781	-32.181	-38.630	-33.279	-41.007	-29.024

5 Results and Discussion

Following sections 2 and 3, we are going to analyze WPT-SVM system using standard database for alcoholic and normal state EEG signal classification. Before we move to actual training and cross-validation methods take a look at Table 1 and 2. These tables represent various statistical parameters namely: standard deviation (σ), skewness (β_1), kurtosis (β_2), energy, mean (μ), maximum and minimum extracted at third level WPT decomposed sub-bands either classes. These parameters can classify signal by mere careful observation. For example, higher mean, lower kurtosis and skewness for alcoholic state than controlled states. But these could not be generalized and hence, we need a classifier. Similar observation can be seen in Fig. 4 which show variation of feature vectors for alcoholic and controlled state EEG signals across different channels.
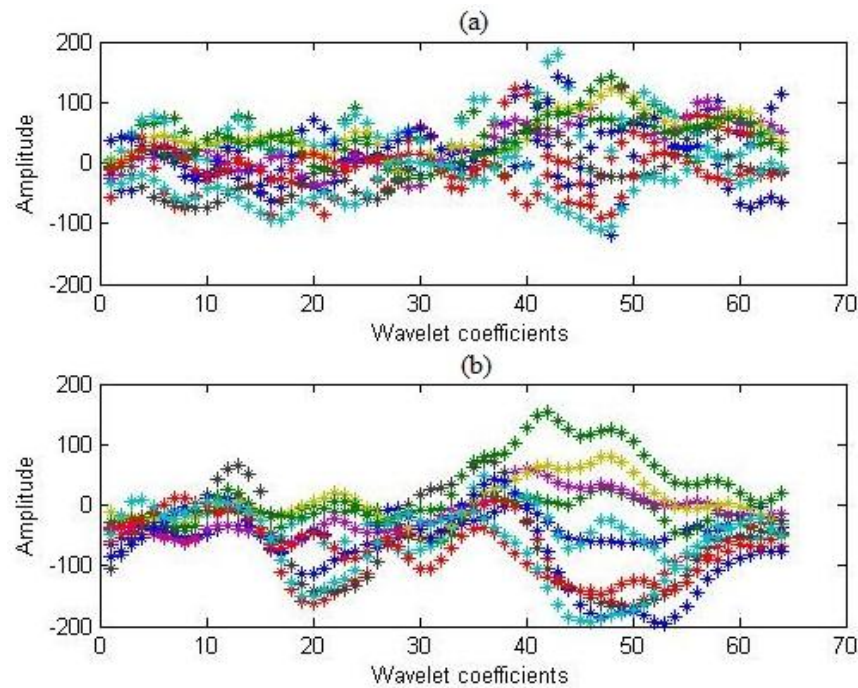
ATLANTIS
PRESS

Fig. 4. Sample of detail wavelet coefficients: (a) Class A (b) Class C

It is predominantly visible that the extracted features are different for both the classes as shown in Table 1 and 2. Therefore these features can be useful for classification purpose. Distribution of various parameters over decomposed wavelet sub-bands can be observed in Fig. 4. The plot consists of eight different lines corresponding to one sub-band each. Class A and Class C do not differ much in Standard deviation on the other hand It can certainly be seen that class A has higher mean corresponding to all sub-bands than class C. Sub-bands 0,1 and 3 have higher skewness and kurtosis in Class A as compared to Class C.

Visual inspection of first level detail sub-bands for either classes, as shown in Fig. 4 indicate significantly distinct patterns. The wavelet coefficients of first level decomposed detail sub-bands from both class are seen in Fig. 5, which shows visible distinction amongst the states. The extracted features from WPT are used to train the SVM. For 60% of vectors provided to train the SVM, WT with SVM has given 86.50% where as WPT with SVM has 88.75% classification accuracy. Consecutively for 80% of vectors provided to train the SVM then WT with SVM has given 94.25% where as WPT with SVM has 95% classification accuracy. Comparison of classification rate has been shown in Table III for both the classes. It is clear that the classification rate for WT with SVM is comparatively low over WPT with SVM. The classification rate of signals shows improvement as length of training vector set increases. The Euclidian distance (ED) shows the similar trend for both WT and WPT based features.
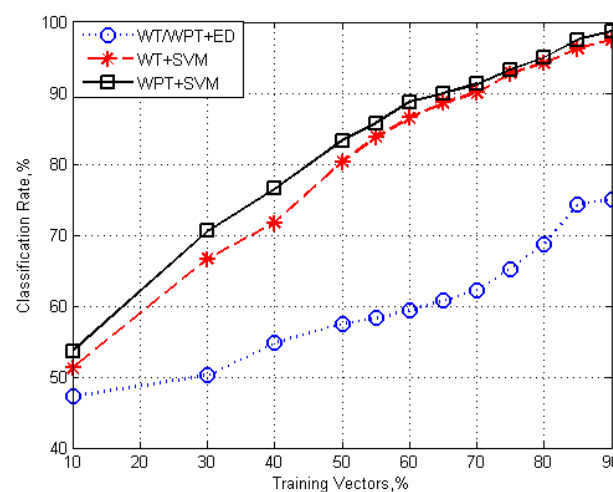

Fig. 5. Classification accuracy vs percentage vector

In case of WPT, 3-level decomposition has given 8 sub-bands, whereas the wavelet transform has given only 4, WPT divides both approximated and detailed sub-band for further decomposition, so features extracted from the EEG signals using WPT gives better results as compared to WT.

6 Conclusion

In this work, we concentrated on the classification of EEG signals for alcoholic and controlled states. The third level WPT sub-band energy feature vector with SVM classifier is proposed here. Decomposition of lower as well as higher sub-bands is one of the advantages of WPT over former WT approaches. Class separation is exploited using energy based features along with SVM classifier. Discrimination power of WPT and WT energy features is compared using several training The WPT-SVM gives better results with slightly higher processing time when compared with WT-SVM and other traditional methods.

Effect of wavelet kernel and decomposition level on feature discrimination power needs to be analyzed. In several parameters of SVM kernel will be optimized under given problem framework.

References

[1] Guler, N., Ubeyli, E.D.: Multiclass Support Vector Machines for EEG-Signals Classification, Information Technology in Biomedicine, IEEE Trans., vol.11, no.2, pp.117-126, (2007).

[2] Shen, M., Lin, L., Chen, J., Chang, C.Q.: A Prediction Approach for Multichannel EEG Signals Modeling Using Local Wavelet SVM, Instrumentation and Measurement, IEEE Trans., vol.59, no.5, pp.1485-1492, (2010).

[3] Lin, K., Chen, M.: On the Design and Analysis of the Privacy-Preserving SVM Classifier, Knowledge and Data, IEEE Trans., vol.23, no.11, pp.1704-1717, (2011).

[4] Mammone, N., Foresta, F., Morabito, F.C.: Automatic Artifact Rejection From Multichannel Scalp EEG by Wavelet ICA,IEEE Sensors Journal, vol.3, pp.533-542,(2012).

[5] Lekshmi, S.S., Selvam, V., Rajasekaran, M. P.: EEG signal Classification using Principle Component Analysis and Wavelet Transform with Neural Network, International conference on communication and signal processing, IEEE conference, vol.1, pp.456-459, (2014).

[6] Das, A.B., Bhuiyan M.I.H., Alam S.M.S.: Statistical parameters in the dual tree complex wavelet transform domain for the detection of epilepsy and seizure, International Conference on Electrical Information and Communication Technology (EICT), vol.2, pp.1-6, (2014).

[7] Yan, G., Yang, B., Chen, S.: Automated and Adaptive Feature Extraction for Brain-Computer Interfaces by using Wavelet Packet, International Conference on Machine Learning and Cybernetics, IEEE conference, vol.no.1, pp.4248-4251, (2006).

[8] Ke, L., Li, R.: Classification of EEG signals by multi-scale filtering and PCA, International Conference on Intelligent Computing and Intelligent Systems, ICIS. IEEE conference, vol.1, pp.362,366, (2009).

[9] Ingber, L.: Statistical mechanics of neocortical interactions: Canonical momenta indicators of electroencephalography, Physical Review E., vol.55,pp. 4578-4593, (1997).

[10] Ingber, L.: Statistical mechanics of neocortical interactions: Training and testing canonical momenta indicators of EEG, Mathematical Computer Modelling, vol. 27, pp.33-64, (1998).

[11] Snodgrss, J. G. , Vanderwart, M.: A standardized set of 260 pictures: norms for the naming agreement, familiarity, and visual complexity, Journal of Experimental Psychology: Human Learning and Memory, vol.6, pp.174-215, (1980).

[12] Wang, D., Miao, D., Xie, C.: Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection, Journal Elsevier of Expert Systems with Applications, vol.38, pp.14314-14320, (2011).

[13] Upadhyay, R., Manglick, A., Reddy, D.K., Kankar, P.K.: Channel optimization and nonlinear feature extraction for Electroencephalogram signals classification, Journal Elsevier of Computers and Electrical Engineering, (2015).