SUPPORT VECTOR MACHINE BASED CLASSIFICATION OF EEG SIGNALS FOR ALCOHOLIC AND CONTROLLED STATE USING WAVELET PACKET TRANSFORM

Rahul R Ingle[1], Sonal Oimbe[2], R.N.Awale[3]
[1]VJTI Mumbai, [2]Cummins College of Engineering for women Pune, [3]VJTI Mumbai
Email:rringle@vjti.org.in[1], aboimbe@gmail.com[2], rnawale@vjti.org.in[3]

Abstract
Classification of acquired Electro-Encephalogram (EEG) signals are of recent interest in Brain Computer Interface systems. Literature review distinguished investigation of EEG patterns due to various activities such as, sleeping, reading, meditation, etc. Work presented here focuses on classification of EEG patterns for alcoholic and controlled states. Third level sub-band energy patterns are generated for either classes using multi-resolution wavelet packet transformation. A well known support vector classifier is employed for classification purpose. Experimental results show significant improvement over wavelet tree feature extraction. Cross-validation tests confirm the greater classification accuracy for proposed technique.

Index Terms: Electroencephalogram, wavelet packet transform, sub-band decomposition, support vector machine.

I. INTRODUCTION

Several biological signals such as Electro-Encephalogram (EEG), Electromyogram, Electrocardiogram, etc contain information that is very important in 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]. EEG signals are brain signal waves captured using various electrodes placed on human scalp. These signals contain large amount of information about various neural activities [2].

EEG signals are highly non-linear, aperiodic, time varying responses characterized with small amplitude and very low frequency [5]. In classification of these signals, one of the difficult task is to find the region of brain depending upon the seizure pattern. Different methods have been proposed to such classify EEG signals for various applications. Performance of classification systems mainly depend on feature extraction. 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 cannot capture transient features in a given signal as well as time frequency information cannot be seen [5]. STFT alleviates such time-frequency conflict by localizing both time and frequency information over uniformly spaced moving window for 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 and STFT [3].

In this paper, in order to find highest accuracy for classification of EEG signals for alcoholic and controlled state of human being, Wavelet Packet Transform with Support Vector Machine (WPT-SVM) classifier is proposed. Rest of the paper has organized as follows: Section II describes acquisition of EEG signals and data description. Section III is about wavelet packet transform (WPT) used for feature extraction. Section IV includes details of Support vector machine (SVM). In section V experimental results are discussed. Section VI concludes the work.

II. ACQUISITION OF EEG SIGNALS

All the experiments have been performed on publicly available UCI machine learning database [9,10]. The database consist of two sets of grouped data abbreviated as ’A’ and ’C’ for Alcoholic and Controlled states, respectively. Each data set has 40 signals recorded from 64 electrodes placed on the scalp according to International 10-20 system of electrode placement and sampled at 256 Hz for a minute. 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 experiment has been repeated for each subject for 60 trials of 1 second each. Every EEG signal consists of 15360* 64 samples each with very small voltage, in range of microvolts (μV). Sample EEG signals from class A and class C are shown in the fig. 1.

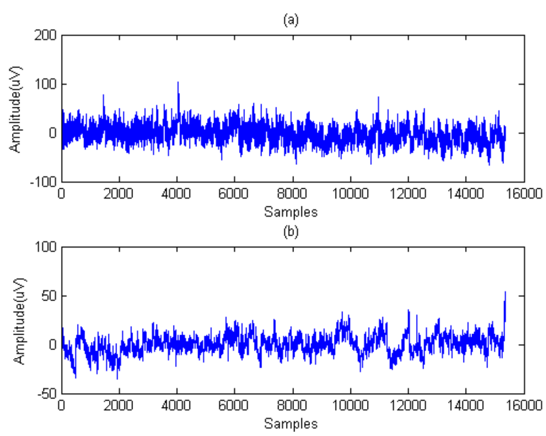

Fig. 1. Time domain EEG signal: (a) Class A (b) Class C.

III. WAVELET PACKET TRANSFORM

Amongst various methods for extraction of features from EEG signals, WPT is used in this paper. The wavelet packet method is a generalization of wavelet decomposition. WPT provides multiresolution analysis for any type of non-linear EEG signals. Analysis of signals can be done at different frequency levels with different resolutions achieved using scaling and dilation parameters, respectively. WPT has dilation and scaling function, which decomposes the signal x[n] in to low frequency band i.e. coarse approximation and high-frequency bands i.e. detail information [7]. Each decomposition level consist of 2 digital filters with response g[k] and h[k] which are a pair of conjugate mirrors [1]. A wavelet packet transform can be represented as;

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

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

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

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

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

$$c^{i}_{j,k} = \int_{-\infty}^{\infty} f(t)\psi^{i}_{j,k}(t)dt \qquad (4)$$

The wavelet packet component of signal at a particular subband can be obtained as,

$$f^{i}_{j}(t) = \sum_{k=-\infty}^{\infty} c^{i}_{j,k}\psi^{i}_{j,k}(t) \qquad (5)$$

The extracted wavelet coefficients provide a compact representation that shows the energy distribution of the signal in time and frequency [12].

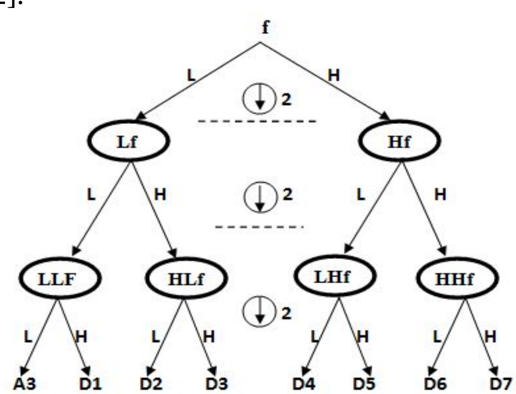

Fig. 2. Proposed Wavelet Packet decomposition scheme

EEG signals are the superposition of different structures occurring at different times. WPT is adopted for feature extraction of EEG signals after decomposition. Selection of appropriate wavelet and level is very important in case of classification of EEG signals. For experimentation purpose, we selected third level decomposition. EEG signal were decomposed to 8 sub-bands, approximate sub-band A3 and 7 detail sub-bands D1 - D7 shown in fig. 2. The number of levels depends on the dominant frequency in the signal, 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. Extracted features are given to SVM classifier for training.

IV. SUPPORT VECTOR MACHINE

SVM is a robust statistical classification technique, primarily employed in case of data scarcity. It is a mathematical analysis to solve n-dimensional optimization using (n-1) dimensional hyper-plane. Hyper-plane is selected in such a way, to maximize the average separation from both the classes. The feature vectors holding separation of hyper-plane are termed as support vectors. For accurate classification only support vectors are required while rest of the dataset becomes redundant. Consider, $f = X \rightarrow Y, x \in X, y \in \{-1,1\}$ [13], and then the subjective function can be written as,

$$f(x) = w.x + b \qquad (6)$$

Where, w is the weight matrix. SVM optimizes above equation under the objective function [3],

$$\arg\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}\varsigma_i \qquad (7)$$

Above equation generates coefficients of hyper-plane which maximizes separation between $w.x + b = -1$ and $w.x + b = 1$. This places positive class on one side of the hyper-plane while negative resides on the other. Each training class tolerance is provided in terms of slack variable ($\zeta_i, i = 1, ...m$). 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 cross validation technique. Fig. 3 gives an example to illustrate the concept of the formulation of the SVM.

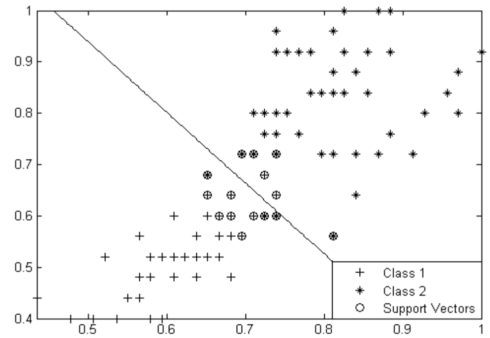

Fig. 3. SVM hyper plane generation to maximize class margin

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
β1	0.296	-0.129	-0.508	0.368	-0.472	-0.108	-0.201	-0.281
β2	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
β1	-0.077	-0.197	-0.238	-0.143	0.196	0.151	-0.178	-0.019
β2	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

V. 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 class. 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 shows variation of feature vectors for alcoholic and controlled state EEG signals across different channels.

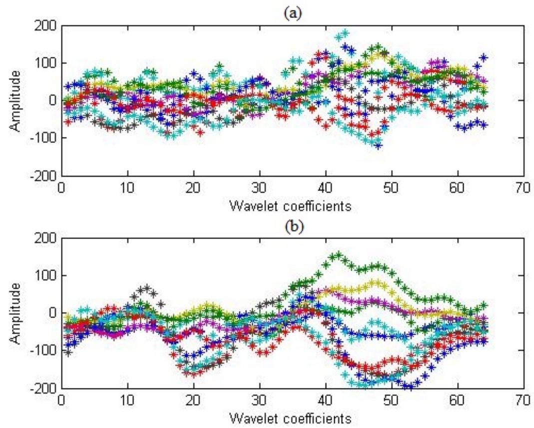

Fig.4. Sample of detail wavelet coefficients: (a) Class A (b) Class C

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.

The wavelet coefficients of first level decomposed detail sub-bands from both class are seen in fig. 4, which shows visible distinction amongst the states. The extracted features from WPT are used to train the support vector machine (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.

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. Fig. 6 shows the effect of increasing percentage of training feature sets and table III clarifies with absolute values.

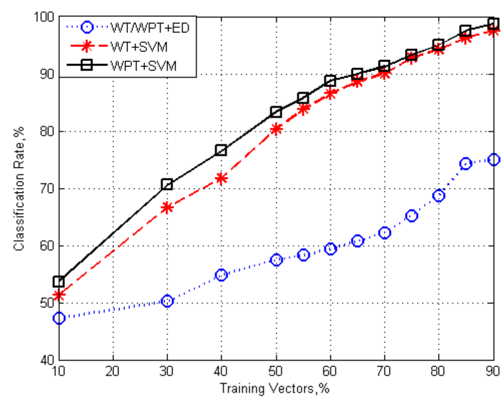

Fig.5 Percentage classification accuracy at various training shares.

TABLE 3. CLASSIFICATION ACCURACY (%)

TRAINING (%)	ED	WT+SVM	WPT+SVM
10	47.22	51.25	53.65
30	50.25	66.65	70.55
40	54.75	71.75	76.50
50	57.50	80.25	83.25
55	58.33	83.75	85.75
60	59.38	86.50	88.75
65	60.71	88.65	90.00
70	62.25	90.05	91.25
75	65.25	92.75	93.25
80	68.75	94.25	95.00
85	74.25	96.25	97.50
90	75.00	97.50	98.75

VI. CONCLUSION

EEG pattern classification has been a recent interest of many researchers. In this work, we concentrated on the classification of EEG signals for alcoholic and controlled states. Third level WPT sub-band energy features with the 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. WPT-SVM gives better results in minimum processing time when compared with WT-SVM and other traditional methods. The classification accuracy can further be improved by using advanced machine learning techniques.

REFERENCES

[1] Fatma 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, March 2007.

[2] Minfen Shen; Lanxin Lin; Jialiang Chen; Chang, C.Q., ”A Prediction approach for Multichannel EEG Signals Modeling Using Local Wavelet SVM,” Instrumentation and Measurement, IEEE Trans., vol.59, no.5,

[3] Keng-Pei Lin; Ming-Syan Chen,”On the Design and Analysis of the Privacy-Preserving SVM Classifier,” Knowledge and Data Engineering, IEEE Trans., vol.23, no.11, pp.1704-1717, Nov. 2011.

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

[5] S.S.Lekshmi; V.Selvam;M.Pallikonda Rajasekaran, ”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, April 3-5, 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, 13-15 Feb., 2014.

[7] Guo-zheng Yan; Bang-Hua Yang; Shuo Chen, ”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, 13-16 Aug., 2006.

[8] Li Ke; Rui Li,”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, 20-22 Nov., 2009.

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

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

[11]J. G. Snodgrss and M. Vanderwart.”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]Deng Wang.; Duoqian Miao; Chen Xie, ”Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection” journal Elsevier Expert Systems with Applications, vol.38, pp.1431414320, 2011.

[13]Wadhwa, N.; Rubinstein, M.; Durand, F.; Freeman, W.T., ”Riesz pyramids for fast phase-based video magnification,” IEEE International Conference on Computational Photography (ICCP), vol no., pp.1,10,2-4 May 2014.