Wavelet Transform Enhancement for Drowsiness Classification in EEG Records Using Energy Coefficient Distribution and Neural Network

Naiyana Boonnak, Suwatchai Kamonsantiroj, and Luepol Pipanmaekaporn

***Abstract*—Reliable classification of drowsy stage in EEG signals have attracted attentions from researchers for many years because of large amounts of brain signal noise. Recent studies have demonstrated that the analysis of EEG signals can get benefits from wavelet transform (WT). Despite of this, experiments do not support the effective use of wavelet features for the discrimination of EEG signals because there is much redundant and irrelevant information contained in wavelet coefficients. Furthermore, extraction of useful features from EEG signals for classification is still an open research question. The novel method present in this paper is to extract useful features for classification of EEG signals based on wavelet transform. This method basically consists of two major steps. The first step is extracting energy coefficients from wavelet transform based on Parseval's theorem to represent the distribution of brain signals. The second step focuses on revising weights of energy coefficients to facilitate a classification method. We show that the energy-based features not only capture meaningful information of wavelet transform, but also are useful for classification. We evaluate the proposed method by using the energy-based features to train a neural network for classification of drowsy and alert signals in EEG records. The experimental results conducted on the MIT-BIH Polysomnographic database have shown that the proposed method achieves 90.27% of accuracy compared to wavelet-based methods.**

***Index Terms*—EEG, drowsiness, alertness, wavelet transform, energy distribution, neural network, classification.**

I. Introduction

Drowsy driving has become a serious problem that leads to thousands of automobile crashes each year [1]. For many years, there are several approaches done to detect the drowsy driver [2]-[5]. In [2], [3], autonomous sensors are used to examine the driver behaviors and the driving performance to infer drowsy driver. However, this method has been unsuccessful because difference of vehicle types and driving conditions. Automatic detection of drowsy driver from relevant body's responses, such as blink rate, eye closure and head movement, using image processing was proposed in [4]. Although this method is fully automatic, it has several restricts. In addition, most drivers are unwilling to be monitored with a camera directly.

Biomedical signals are useful to infer the body's response. Among these signals, Electroencephalogram (EEG) is one of reliable methods to detect several problems of the brain [5] and is a main method in drowsy detection [10]. Fig. 1 illustrates EEG signals of alert and drowsy stages.

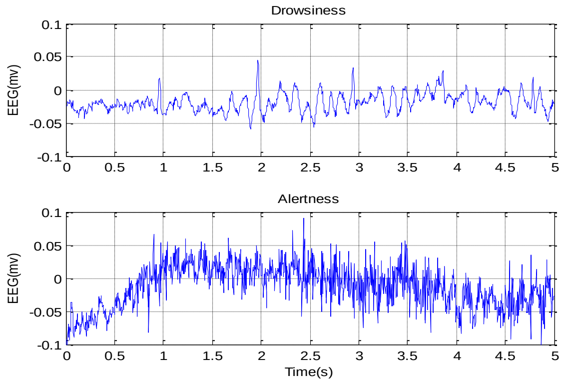

Fig. 1. EEG signals of alertness and drowsiness stages.

Several approaches to the analysis of EEG signals were proposed in the last decade [6]-[9]. Basically these approaches can fall into one of the following schemes: 1) time analysis, 2) frequency analysis and 3) time-frequency analysis. In time domain, the analysis of EEG signals is normally done by extracting statistical values, such as maximum, minimum, mean and standard deviation. Then, these values are used to distinguish between the alertness and drowsiness stages [6]. Despite its efficient computation, these statistical features are not effective enough to capture underlying properties of EEG signals. More common approaches are based on frequency analysis [7]-[9]. , Fourier analysis is a dominant technique that extracts a frequency spectrum of EEG signals. However, it is known that the conventional method of frequency analysis is not highly successful for the analysis of non-stationary signals (i.e., EEG signals) [6], [11].

According to studies [11], [13], [17], the appropriate way to the analysis of the non-stationary signals is using wavelet transform (WT) which is a mathematical tool for signal processing based on time-frequency analysis [21]. The main idea of WT is to decompose a signal into the frequency sub-bands using wavelet transform and then a set of statistical features are extracted from the sub-bands to represent the distribution of wavelet coefficients. Despite this, experiments do not support the effective use of wavelet coefficients [6], [16].

For many years, there are several methods proposed to enhance WT for EEG classification [6], [11], [15], [16]. For example, Adaptive Neuro-Fuzzy Inference System (ANFIS) is applied to wavelet coefficients instead of standard classification methods for EEG classification [16]. Some

Manuscript received November 29, 2014; revised February 26, 2015.
The authors are with the Department of Computer and Information Science, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand 10800 (e-mail: getaaun9@gmail.com, suwatchaik@kmutnb.ac.th, luepolp@kmutnb.ac.th).

DOI: 10.7763/IJMLC.2015.V5.522

recent studies [11], [15] have experimented that Power Spectral Density (PSD) obtained from discrete wavelet transform of full spectrum EEG outperform other wavelet features. In study [6], the combination of multiple sets of features extracted from time analysis, PSD and wavelet coefficients is employed to differentiate alert and drowsy signals in EEG records where 7 of 19 features are manually selected. According to their experimental results, the multimodal-based features perform better than single features for classification of drowsiness in EEG records. However, automatic extraction of useful features describing alter and drowsy signals in EEG records is still an open research question.

Motivated by the above challenge, we propose a novel method to enhance wavelet coefficients for EEG classification. This method basically consists of two major steps: 1) extraction of energy coefficients and 2) feature revision. The objective of the first step is to find a compact and meaningful representation of wavelet coefficients. This is done by extracting energy coefficients from wavelet transform of brain signals based on Parseval's theorem. The aim of the second step is to facilitate a classification technique (i.e., neural network) by revising the weights of each energy coefficients before the classification is made. In this work, we apply Mapping-Constrained Agglomerative (MCA) [21], an efficient algorithm which automatically finds appropriate locations of clusters by considering both inputs and outputs, as knowledge base to obtain accurate weights of the energy distribution. We evaluate the proposed method by using the energy-based features to train a neural network for the discrimination of drowsiness and alertness stages in EEG signals. The experimental results show impressive performance.

The rest of the paper is organized as follows: Section II gives background of wavelet transform and the energy signal based on the Parseval's theorem. In Section III, we describe our proposed method including extraction of energy coefficients from wavelet decomposition and input-output clustering algorithm. Experimental results and discussion are present in Section IV, following by conclusions and future work in Section V.

II. Background

A. Wavelet Tranform

Wavelet Transform is a mathematical technique that can convert a signal $f(t)$ into shifted and scaled versions of the mother wavelet. The wavelet transform [18] decomposes a given signal $f(t)$ into increasingly details that described by:

$$f(t)=\sum_{j\in z} 2^{j/2} c_j(k)\varphi(2^j t-k)+\sum_{j=0}^{j-1}\sum_{k=0}^{\infty} 2^{j/2} d_j(k)\omega(2^j t-k) \quad (1)$$

where $\varphi(t)$ is a scaling function, $\omega(t)$ is a mother wavelet function. The first component is an approximation of $f(t)$ on the scale index j while the second one is an detail using upper scale index $j+1$, $c_j(k)$ is scaling coefficient and $d_j(k)$ is wavelet coefficient. These coefficients can be calculated as follow:

$$c_j(k)=\int_{-\infty}^{\infty} f(k)\varphi(2^j t-k)\,dt \quad (2)$$

$$d_j(k)=\int_{-\infty}^{\infty} f(k)\omega(2^j t-k)\,dt \quad (3)$$

From (1), the scaling function can be associated with the low-pass filters (LPF) and the wavelet function can be associated the high-pass filters (HPF). The decomposition procedure will start when $f(t)$ was entered into these filters. Then the approximations and the details are the low and high frequency component of signal, respectively. The outputs from them are the detail and approximation coefficients at 1-level shown as A1 and D1. Then the approximation signal of 1-level are sent into 2-level to repeat the decomposition procedure. Next the two outputs are the detail and approximation coefficients shown as A2 and D2. Finally, the signal is decomposed at the expected level. The example of wavelet decomposition for three levels shows in Fig. 2.

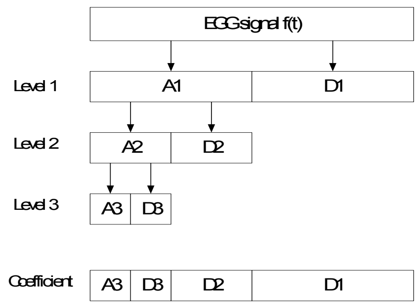

Fig. 2. Wavelet decomposition at the expected level is 3.

B. Parseval's Theorem

The Parseval's theorem [22] refers to energy of a signal in terms of its Fourier transform. This theorem is useful for evaluating the energy of a signal without prior knowledge about its time domain. Given a signal $f(t)$, the energy coefficients of the signal $f(t)$ are defined by:

$$E=E_i=\sum_{k=1}^{n/2} A_i^2(k)+\sum_{k=1}^{n/2} D_i^2(k) \quad (4)$$

where $E=\sum_{k=1}^{n} f^2(t)=f^2(1)+f^2(2)+\cdots f^2(n)$, $f(t)$ is a signal at time t, $0<t<=n$ and E_i is defined as energy coefficient at its i^{th} level.

Let us give an example. Suppose $f(t)$ = [4, 6, 10, 12, 8, 6, 5, 5] is an input signal. According to the above equation, this signal has the energy E is equal to 446. After we apply Haar wavelet transform, the values of D1 and A1 for its 1-level is [D1|A1] = [$5\sqrt{2}$, $11\sqrt{2}$, $7\sqrt{2}$, $5\sqrt{2}$| $-\sqrt{2}$, $-\sqrt{2}$, $\sqrt{2}$, 0]. It can show that $E_1=\sum_{k=1}^{n/2} A_i^2(K)+\sum_{k=1}^{n/2} D_i^2(K)=446$. Thus the 1st level Haar transform has kept the energy constant. It can be expected that at the 3rd level the energy $E_3=A_3^2+D_3^2+D_2^2+D_1^2$ is also equal to 446.

TABLE I: Frequency Band Level 1-6 of EEG at Sampling Rate 250Hz

Decomposed signals	Frequency bands (Hz)	Level
D1	62.50 – 125	1
D2	31.25 - 62.50	2
D3	15.62 - 31.25	3
D4	7.81 - 15.62	4
D5	3.90 - 7.81	5
D6	1.95 - 3.90	6
A6	0.0- 1.95	6

It is known that the number of decomposition levels related to the sampling rate of the primary signal. In this study, the sampling frequency of the EEG signal is 250 Hz. The frequency band of 1st level is [250/4:250/2], the frequency band of 2nd level is [125/4: 125/2] and the frequency band of the other levels are shown in Table I.

III. Our Metheodology

The main process of our proposed method can be seen in Fig. 3. According to this figure, an EEG signal is segmented and then decomposed with the conventional method of wavelet transform described in the previous section. After that, we extract energy coefficients from the wavelet decomposition to represent the distribution of brain signals. After that, we apply the input-output clustering to adapt weights of each energy coefficient before the classification is made. Finally, a neural network is used to utilize the energy-based features for characterizing alert and drowsy stages in EEG records.

A. Distribution of Signal Energy

The energy coefficients of each resolution level have a wide range of values. For example, the D1 energy coefficients of alert and drowsiness stages are shown in Fig. 4(a) and Fig. 4(b) respectively.

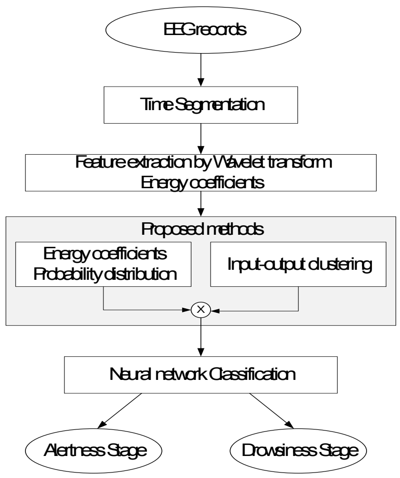

Fig. 3. The process of our proposed method.

Fig. 4(a), the x-axis is the segment number and y-axis is energy coefficient. There are various values of D1 energy coefficient. The minimum value is $0.88 \text{x} 10^{-3}$ and the maximum value is 62.27. The corresponding histograms of the D1 energy coefficient to approximate probability density are shown in Fig. 4(c). Fig. 4(c) shows that the histograms skew to the left. These shapes of the histograms indicate that the small energy coefficient has the higher probability than the large energy coefficient. As shown in Fig. 4(c) and Fig. 4(d), the D1 energy coefficient distribution is not obviously normal distribution. We assume to use the lognormal distribution for describe the energy coefficient distribution that is defined by:

$$y = f(x|\mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{\frac{-(\ln x - \mu)^2}{2\sigma^2}} \qquad (5)$$

where y is the output value of lognormal distribution function, x is input values, μ is mean of x and σ is standard deviation of x. The energy coefficient distributions of the other inputs are not also normal distribution.

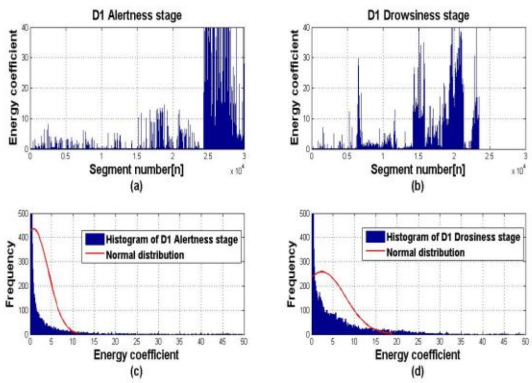

Fig . 4. D1 energy coefficients and their distributions of alert and drowsy signals.

B. Feature Revision Using Input-Output Clustering

In the previous section, we show how to extract a distribution of energy coefficients from wavelet transform. The energy distribution offers a compact and meaningful representation of wavelet coefficients. However, the evaluated weights of each energy coefficient obtained may be not accurate enough to describe alert and drowsy signals because of many noises in EEG records. As a result, this may lead to a negative impact on the classification accuracy obtained.

In this section, we propose to revise weights of each energy coefficient extracted from wavelet transform to facilitate a learning method for classification The idea is to adapt weights of each coefficient of an input vector by combining the likelihood estimates of the nearest cluster.

To estimate the likelihoods, an input-output clustering method is applied in this work. The goal of input-output clustering (IOC) is to find appropriate locations for clusters by considering both inputs and outputs (i.e., drowsiness and alertness). The idea of IOC is to group input data according to the following constraints:

1) The connective data in input region with similar outputs should be group as the same cluster.
2) The non-connective data in input regions with similar outputs should be as different clusters.
3) The non-connective data in input regions with different outputs should be as different clusters.

In this work, we adopt MCA (Mapping-Constrained Agglomerative) algorithm proposed in [21] for weight revision of energy coefficients of an input signal. Given the training data, this algorithm automatically extracts the optimal number of clusters by considering both similar inputs with similar outputs of the training data. It also estimates the means and variances of each cluster in a single pass. We thus use the MCA algorithm to estimate the likelihoods of the input signal.

The following is the outline of MCA algorithm.

1) Given a set of training data $\{(x_k, y_k) | \, k = 1, \ldots, n\}$, initialize $C \leftarrow C_{max}$ and $c_i \leftarrow 1$, where C is the number of cluster; C_{max} is a maximum number of clusters; and c_i is the number of points in the cluster i.
2) Select the first C members of training data and initially assign $m_{Ii} \leftarrow x_i$, $\sigma_{Ii}^2 \leftarrow 0$, $m_{Oi} \leftarrow y_i$, $\sigma_{Oi}^2 \leftarrow 0$ where m_{Ii} and m_{Oi} are the centers of the cluster i in the input space and output space, respectively; σ_{Ii}^2 and σ_{Oi}^2 are the variances of the cluster i in the input space and the

output space, respectively.

3) Compute input and output upper-triangular distance matrices defined as $D_I \in R^{C\times C}$ and $D_O \in R^{C\times C}$, respectively. The elements of D_I and D_O at the a-th row and b-th column defined as d_{Iab} and d_{Oab} where the elements d_{Iab} and d_{Oab} are the distances between the center of the cluster a and the center of the cluster b in the input and output spaces, respectively. They can be computed by:

$$d_{Iab} = \|m_{Ia} - m_{Ib}\| \tag{6}$$

$$d_{Oab} = \|m_{Oa} - m_{Ob}\| \tag{7}$$

where $a,b \in \{1,\ldots,c\}, a \neq b$ and $\| \quad \|$ is the Euclidean norm.

4) Select input a data point (x_k, y_k)
IF $k < n$
$k \leftarrow k+1$
ELSE
$k \leftarrow 1$

5) Find the clusters closest to the current data point is defined as winner input cluster(γ) and winner output cluster (δ).

$$\gamma \leftarrow \|m_{I\gamma} - x_k\| = min\|m_{Ii} - x_k\|$$

$$\delta \leftarrow \|m_{I\delta} - y_k\| = min\|m_{Oi} - y_k\|$$

The cluster centers $m_{I\gamma}$ and $m_{I\delta}$ are the closest center to x_k andy_k.

$$IF\ \gamma = \delta$$

The input cluster and output cluster of (x_k, y_k) satisfies the mapping consistency, we will assign (x_k, y_k) into the cluster γ by performing:

Update the centers and variances of the cluster γ in both input and output space as follows:

$$c_\gamma \leftarrow c_\gamma + 1 \tag{8}$$

$$\text{temp\_}k \leftarrow temp\_k + 1 \tag{9}$$

For the input space:

$$temp\_m_{I\gamma} \leftarrow m_{I\gamma} + \frac{x_i - m_{I\gamma}}{c_\gamma} \tag{10}$$

$$\sigma_{I\gamma}^2 \leftarrow \frac{(c_\gamma - 1)(\sigma_{I\gamma}^2 + m_{I\gamma}^2) + x_i^2}{c_\gamma} - temp\_m_{I\gamma}^2 \tag{11}$$

$$m_{I\gamma} \leftarrow temp\_m_{I\gamma} \tag{12}$$

For the output space:

$$temp\_m_{O\gamma} \leftarrow m_{O\gamma} + \frac{y_i - m_{O\gamma}}{c_\gamma} \tag{13}$$

$$\sigma_{O\gamma}^2 \leftarrow \frac{(c_\gamma - 1)\left(\sigma_{O\gamma}^2 + m_{O\gamma}^2\right) + y_i^2}{c_\gamma} - temp\_m_{O\gamma}^2 \tag{14}$$

$$m_{O\gamma} \leftarrow temp\_m_{O\gamma} \tag{15}$$

Repeat step 3) to update the distance metrics for the input and output spaces.

$IF\ temp\_k \leq n$
Go to step 4)
ELSE
Termination of clustering process.
ELSE
Go to step 4)

Once the clusters were identified, the probability of the nearest cluster is used to adapt weights of energy coefficient vector before a classification is made by a neural network.

C. Neural Network Classifier

To effectively use the energy distribution, a multi-layer feed-forward artificial neural network is trained using standard back-propagation algorithm. As seen in Fig. 5, an input vector that consists of 7 energy coefficients (i.e., D1-D6 and A6) is applied to the input layer. Then, all of the inputs are distributed to each unit in the hidden layers. The output of the final layer is computed by multiplying the output vector from the hidden layers by the weights into the final layer. We use the common asymmetric sigmoid function as activations of these units.

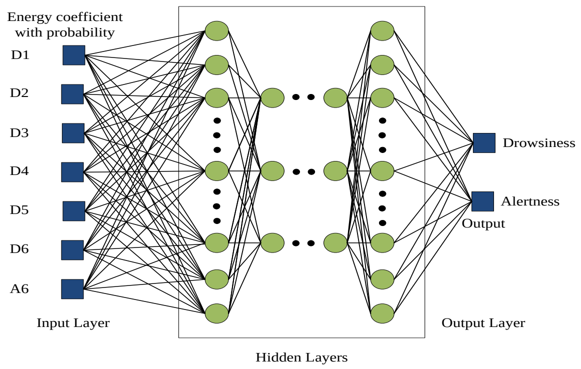

Fig. 5. Multi-layered neural network.

It is known that finding the appropriate number of hidden layers is one of the most critical tasks in neural network design. A network with too few hidden nodes would be incapable of differentiating between complex patterns. In contrast, too many hidden nodes will lead to poor generalization for unseen data and become time-consuming. For this experiment, we found that the best performance is achieved with 5 hidden layers by trial and error and 10 cross-validation procedure.

IV. Experimental Evaluation

A. Data Collection

We use the Massachusetts Institute of Technology (MIT)-Beth Israel Hospital (BIH) polysomnographic database [19]. This dataset contains a collection of eighteen recordings of multiple physiological signals during sleep. The patients were monitored for evaluation of sleep stages. The dataset also contains over 80 hours of four-, six-, and seven-channel polysomnograms, each of them with an EEG signal annotated by experts. The recording time for patients ranged from 2 to 7 hours, and the polysomnograms corresponded to men between 32 and 56 years of age.

B. EEG Signal Processing and Segmentation

Because of long EEG recordings, we segment the EEG signals for the analysis of EEG. In this experiment, we choose the following duration time of signal segmentations: 1, 5 and 10 seconds respectively. Each segment is annotated with a label of "alertness" or "drowsiness" and is used as the training and test sets for classifier evaluation.

According to our experimental results, we obtained the best performance when using a collection of 1s segments of EEG records (i.e., 71.84% of mean square errors). Thus, we use this signal collection for this experiment.

The Daubechies 2 mother wavelet is used to decompose the EEG segments of 1s duration in six levels. The D1-D6 and A6 energy coefficient are the input features. The output categories are "0 1" and "1 0" to represent alert and drowsiness stages, respectively. There are 107,460 EEG segments (i.e., 84,000 alertness and 23,460 drowsiness records).

C. Lognormal Distribution Parameters

To verify the distribution density of energy coefficient using lognormal distribution, we examine the distribution by using plots of probability density. Three probability distribution functions are compared: Lognormal, Gamma and Exponential distributions. The probability density plots of D1 energy coefficient are shown in Fig. 6.

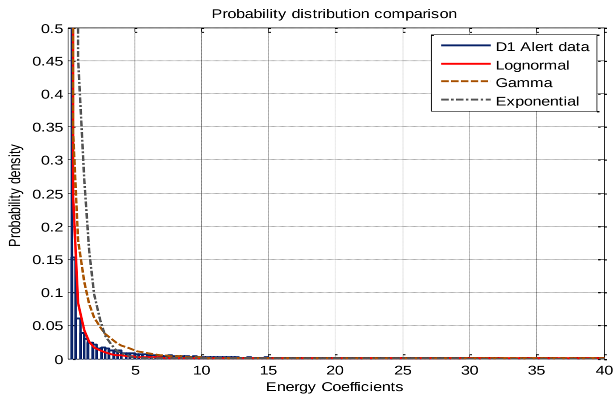

Fig. 6. Three types of probability density plots of D1 energy coefficient.

Fig. 6 illustrates the superior fitting with lognormal distribution compared the other distributions. To verify the energy coefficient distribution of other levels, we will also use similarly procedure. The 2-parameters of lognormal distribution of all levels are shown in Table II. All of X^2 are less than X^2 with alpha 0.05 and degree of freedom (*df*) 60-1-2 show that all level of energy coefficient are lognormal distribution. Table II shows the average (μ) and standard deviation (σ) of each feature.

TABLE II: THE 2-PARAMETERS μ AND σ OF LOGNORMAL DISTRIBUTION

Levels	Alert			Drowsiness		
	X^2	μ	σ	X^2	μ	σ
D1	28.79	-3.3081	2.2145	34.66	-2.3055	2.9983
D2	40.53	-2.5248	1.3406	4.74	-4.0171	2.2933
D3	8.32	-2.0058	1.4186	8.45	-3.9381	2.4419
D4	48.04	-1.4719	2.1933	4.01	-4.5752	2.8558
D5	7.24	-2.9783	1.7575	7.21	-5.1284	2.5997
D6	2.31	-3.8484	1.6395	5.44	-5.3381	2.5598
A6	6.1	-1.3302	1.3407	11.22	-2.3425	2.1860

Since the data are divided into 60 bins and we have estimated two parameters, the calculated value may be tested against the chi-square distribution with 60 -1 -2 = 57 degrees of freedom. For this distribution, the critical value for the 0.05 significance level is 75.61. Since all $X^2 < 75.61$, we do not reject the null hypothesis that the data are log-normal distributed.

D. Results and Discussions

Table III compares the accuracy performance of three different feature extraction methods based on wavelet transform (WT). Our proposed method that uses the energy coefficient with probability distribution (ECPD) to represent the distribution of EEG signals is compared to the two baseline methods, including Power Spectrum Density (PSD) [11] and the method that uses multi-modal features from time analysis, PSD, and WT, named Multi-Modal Analysis (MA) [5].

TABLE III: THE PERFORMANCE COMPARISON OF SIGNAL FEATURE EXTRACTION METHODS

Method	Number of extracted features	Accuracy (%)
Energy coefficients with probability distribution (ECPD)	**7**	**90.27**
Multimodal Analysis (MA) [6]	7	85.66
Power Spectrum Density (PSD) [11]	12	66.67

TABLE IV: THE CONFUSION MATRIX OF CLASSIFCATION RESULTS

Method		Alertness	Drowsiness
ECPD	**Alertness**	**98.09**	17.55
	Drowsiness	1.91	82.45
MA [6]	**Alertness**	87.46	16.14
	Drowsiness	12.54	**83.86**
PSD [11]	**Alertness**	84.12	50.78
	Drowsiness	15.88	49.22

As seen in Table III, ECPD achieves the best performing method for feature extraction in EEG signals with 90.72% of the classification accuracy. MA that uses a set of multiple features from time analysis, PSD, and WT, also performs the second best one with 85.66% of accuracy on this collection of EEG data. We also compare ECPD with PSD obtained from DTW. As seen in Table III, ECPD largely outperforms PSD for the measure of accuracy. Furthermore, the number of extracted features from ECPD is less than that of PSD.

One interesting point reveled in this table is that PSD, a commonly used method for the classification analysis of EEG signals, is not effective enough to achieve the encouraging performance since there are redundant and noisy information extracted from wavelet transform. Furthermore, the results support the superiority of the distribution of energy coefficients for enhancement of wavelet transform. The improvements are also consistent in the number of correct and incorrect classification on the test data made by the proposed method. As seen in Table IV, ECPD achieves the best correct classification of drowsy (98.09%) and alert (82.45%) signals with the lowest incorrect classification rates compared to MA and PSD respectively.

We also examine the effect of feature revision. Fig. 7 illustrates the average values of energy coefficient distribution of alert (blue line) and drowsy (red line) class before/after the revision process. As seen in this figure, the average values of energy coefficients at each level are not significantly different between alert and drowsy signals (the above graph). This may lead to a negative impact on classification methods used for making a decision. However, after the weight revision is performed, the obvious differences of alert and drowsy signals can be easily identified. Consequently, this can facilitates the classification method for the discrimination of brain signals. This result also highlights the use of input-output clustering for feature revision.

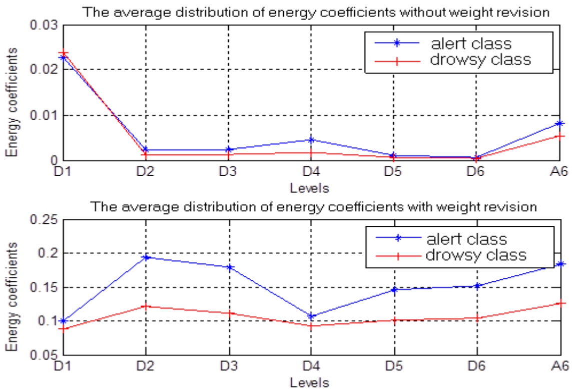

Fig. 7. The average distributions of energy coefficients without weight revision (top) and with weight revision (bottom) on the training data.

V. Conclusions and Future Work

In this work, we tackle the deficient use of wavelet transform for reliable classification of EEG signals. We have presented a novel method to enhance wavelet transform for classification of EEG signals. This method basically extracts energy coefficients which captures meaningful information of wavelet transform based on the Parseval's property. Then, the process of feature revision is applied to each energy coefficient to facilitate the classification method. We evaluate the proposed method by using the energy-based features to train a neural network for characterization of alert and drowsy signals. The experimental results conducted on MIT-BIH Polysomnographic database have shown that the proposed method achieves encouraging performance in comparing with standard wavelet-based features.

Although the proposed method is promising for classification of EEG signals, a fair computation is usually required. Consequently, it might be not suitable to apply this method for real-time signal processing. We thus plan to improve the efficiency issue of our method for future work.

References

[1] National Highway Traffic Accidents 2011 report, Department of Highways, Ministry of Transportation and Communications, Thailand.

[2] T. H. Chang, C. S Hsu, C. Wang, and L. K. Yang, "Onboard measurement and warning module for irregular vehicle behavior," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 3, pp. 501-513, 2008.

[3] A. Sahayadhas, K. Sundaraj, and M. Murugappan, "Detecting driver drowsiness based on sensors: A review," *Sensors*, vol. 12, no. 12, pp. 16937-16953, 2012.

[4] P. Smith, M. Shah, and N. D. V. Lobo, "Determining driver visual attention with one Camera," *IEEE Transactions on Intelligent transportation Systems*, vol. 4, no. 4, pp. 205-218, 2003.

[5] A. Crepel, P. Gelisse, M. Burean, and P. Genton, *Atlas of Electroencephalography*, Paris: Eurotext, 2005.

[6] G. Corre, L. Orosco, and E. Laciar, *Automatic Detection of Drowsiness in EEG Records Based on Multimodal Analysis*, Elsevier, Med Eng Phys, 2013, pp. 1-6.

[7] P. P. Weiss, "Classification of electroencephalogram (EEG) signal based on fourier transform and neural network," PhD Thesis, 2013.

[8] R. Upadhyay, P. K. Kankar, P. K. Padhy, and V. K. Gupta, "Classification of drowsy and controlled EEG signals," in *Proc. the IEEE Int. Conf. on Engineering*, December 2012.

[9] A. Picot, S. Charbonnier, and A. Caplier, "On-line automatic detection of driver drowsiness using a single electroencephalographic channel," in *Proc. the 30th Annual International Conf. IEEE Engineering in Medicine and Biology Society, EMBC'08*, 2008.

[10] L .N. Boyle, J. Tippin, A. Paul, and M. Rizzo, "Driver performance in the moments surrounding a microsleep," *Traffic Psychology Behavior Transportation Research*, vol. 11, no. 2, pp. 126-136, 2008.

[11] A. Correa and E. Leber, "An automatic detector of drowsiness based on spectral analysis and wavelet decomposition of EEG records," in *Proc. the 32nd International Conf. of the IEEE EMBS Buenos Aires*, Argentina, Aug 31-Sep 4, 2010, pp. 1405-1408.

[12] A. Subasi, "Automatic recognition of alertness from EEG by using neural networks and wavelet coefficients," *Expert Systems with Applications*, vol. 28, no. 4, pp. 701-11, 2005.

[13] M. K. Kiymik, M. Akin, and A. Subasi, "Automatic recognition of alertness level by using wavelet transform and artificial neural network," *Journal of Neuroscience Methods*, Elsevier B.V., vol. 139, pp. 231-240, 2004.

[14] A. Subasi, "Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients," *Expert Systems with Applications*, vol. 28, pp. 701-711, 2005.

[15] D. Easwaramoorthy and R. Uthayakumar, "Analysis of biomedical EEG signals using wavelet transforms and multifractal analysis", *IEEE*, pp 544-549, 2010.

[16] I. Guler and E. Ubeyli, "Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients," *Journal of Neuroscience Methods*, Elsevier B.V., pp. 1-9, 2005.

[17] N. Hazarika, J. Chen, A. Tsoi, and A. Sergejew, "Classification of EEG signals using the wavelet transform," *Signal Processing*, Elsevier Science B.V., vol. 59, pp 61-72, 1997.

[18] C. Burrus, R. Gopinath, and H. Guo, *Introduction Wavelets and Wavelet Transforms: A Primer*, Prentice-Hall, Upper Saddle River, NJ, 1998.

[19] Y. Ichimaru and G. B. Moody, "Development of the polysomnographic database on CD-ROM," *Psychiatry and Clinical Neurosciences*, vol. 53, pp. 175-177, 1999.

[20] Goldbeger, L. Amaral, N. Glass, J. Hausdorff, P. Ivanov, R. Mark, J. Mietus, G. Moody, C. Peng, and H. Stanley, "PhysioNet components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp.215-220, 2000.

[21] J.-S. Wang and C. S. G. Lee, "Self-adaptive neuro-fuzzy inference systems for classification applications," *IEEE Transactions on Fuzzy Systems*, vol. 10, no. 6, pp. 790-802, 2002.

[22] Singh, R. Prasad, and S. D. Sapre, *Communication Systems, 2E*, Tata McGraw-Hill Education, pp. 93-96, 2008.

Naiyana Boonnak is currently a master student at the Department of Computer and Information Science, King Mongkut's University of North Bangkok, Thailand. She holds bachelor's degrees in computer science from Prince of Songkla University, Thailand. Her research interests include artificial intelligence and artificial neural network.

Suwatchai Kamolsuntiroj is currently a lecturer at the Department of Computer and Information Science, King Mongkut's University of North Bangkok, Thailand. He holds master's degrees in computer science. He also earned his doctoral degree in computer engineering from Kasetsart University, Thailand, graduating in 2013. His current research interests involve fuzzy logic, neural network and artificial intelligence.

Luepol Pipanmaekaporn is currently a lecturer at the Department of Computer and Information Science, King Mongkut's University of North Bangkok, Thailand. He holds bachelor's and master's degrees in computer science. He also earned his doctoral degree in computer science from Queensland University of Technology, Australia, graduating in 2013. His current research interests involve information retrieval, web mining, and data mining.