A Study on Sleep EEG Using Sample Entropy and Power Spectrum Analysis

Aaron Raymond See and Chih-Kuo Liang
Department of Electrical Engineering
Southern Taiwan University
Tainan, Taiwan

Abstract — **Research on the automation of sleep stage classification, particularly single channel EEG, has been a challenge for many years. The research aims to look into the analysis and evaluation of feature extraction techniques and classification methods that are important to properly classify sleep stages with limited channels. Sample entropy, and the power spectrum of the harmonic parameters using infinite impulse response filters and wavelet transform were used to extract features from the data taken from Physionet database. A total of 13 features were initially extracted and used for the training and testing of the sleep stage classification system. Analysis of the training data showed a distinct combination patterns between the sample entropy and harmonic parameters with a change in the sleep stage. In addition, a prototype for the sleep stage classification system was implemented. Support Vector Machine (SVM) was utilized for the classification system. While the training data were extracted from several database. Further refinement of the data and the program could be useful for a test the sleep stage classification on other database or data.**

Keywords- sample entropy, harmonic parameters, power spectrum, wavelet transform, sleep stage classification, support vector machine

I. INTRODUCTION

Due to the increase in the number of sleep disorder patients, sleep examination is becoming prevalent. The current standard sleep examination used is the Polysomnograph (PSG) that requires several sensors on the examinee. But today more and more research on the development of single channel Electroencephalograph (EEG) sleep recording and stage classification is becoming more popular for simplicity and comfort [1].

Traditional EEG channels for sleep recording are placed either at C3-A2 or C4-A1 according to the 10-20 system for placement of EEG electrodes on the scalp [2-3]. While classifying different sleep stages is done using Rechtschaffen and Kales (R&K) scoring manual to determine the sleep characteristic features of non-rapid eye movement (NREM) and REM sleep. NREM sleep was divided into four stages namely: transitional sleep (stage 1), light sleep (stage 2), slow wave sleep (stages 3 and 4) and paradoxical sleep also known as the REM sleep [4-5]. In recent sleep stage classification, it was noted that stage 1 and stage 2 can be combined to be light sleep (LS). Then stage 3 and 4 are then combined and is well known as deep sleep or slow wave sleep (SWS). There has been several researches that have used SWS as a substitute to the later 2 stages since there are not much significant difference between the two while other research preferred to use 4 stages to make an analysis [2,7,14]. Thus, to simplify sleep staging the research used 4 stages wake, LS, SWS, and REM.

However, manual scoring is time consuming and error prone therefore many researchers have developed algorithms and programs to automatically classify sleep stages.

One study is on the use of sample entropy. It was found to be able to reflect complexity of the mixture system and showed differences in different sleep stages. It can be noted that wake has the highest value of sample entropy and decreases as a person goes deeper into their sleep but rises again upon reaching REM [6].

Another one is an automated sleep staging program using a single-channel sleep EEG was developed and tested by a group of doctors and researchers. Tests were done on an epoch by epoch basis while classifying sleep from 2 states to 5 states using fast Fourier transform and power spectrum density and fuzzy algorithm. It showed promising results of more than 80% sensitivity and could be a good candidate for diagnostic aid and automatic sleep stage scoring [7].

Furthermore, since the EEG data is considered a non-linear data, Wilson et al. used the wavelet based continuous Morlet transform to analyze EEG data in both the time and frequency domain. Result from using continuous wavelet transform was that delta rhythm in sleep stages 2, 3, and 4 showed significant variations. Moreover, the K-complex waves can be highlighted in the averaged power spectrum density spectrum. It also noted a gradual increase in power from sleep stages 2 to 4 due to the delta wave increase [8]. While another research performed sleep stage classification used wavelet transform and neural network on 2 EEG signals at 200 Hz sampling rate. The EEG signals are subdivided into wavelet packets before going into the classifier. For the wavelet Daubechies (db20) was used [9].

II. MATERIALS AND METHODS

A. Data Set

The raw EEG data were taken from the Physionet (MIT-BIH) Polysomnographic database. The subjects are male with an average age of 42.24 years. The data is taken with a sampling rate of 250 Hz and ADC resolution of 12 bits. While

We would like to thank National Science Council, for providing financial support for this project through NSC 99-2221-E-218 -012.

the EEG recording were taken from channels C3-O1, C4-A1 and O2-A1. The data are first converted from the .mat file to ASCII file before extracting 30-second epochs from the EEG data [10].

B. Data preprocessing

The EEG is first filtered with a level 5 Butterworth high-pass filter with cut-off frequencies between 0.5-30 Hz and then by a band stop filter on the 50-60Hz frequencies. The signal is further filtered using the LabVIEW Wavelet Denoise. It performs noise reduction for 1D signals by using the discrete wavelet transform (DWT) or un-decimated wavelet transform (UWT). The transform used for the system is DWT using db04 wavelet with 7 levels and soft threshold.

C. Feature extraction

Three different methods were used to extract features from the EEG data. These are sample entropy, infinite impulse response (IIR) filter and wavelet multi-resolution analysis and its corresponding power spectrum density.

First, sample entropy is calculated with length (N) = 1500, tolerance for accepting matches (r) = 0.25 and m to be 2.

Second, signals are filtered using a 5th order Butterworth band pass filters to separate the signal into different harmonic frequencies. Signals are separated into Beta 2 (20-45Hz), Beta 1 (12-20Hz), Alpha (8-12Hz), Theta (4-8Hz), Delta 2 (2.5-4Hz) and Delta 1 (0.5-2.5Hz). The single sided scaled amplitude spectrum of a real-valued time domain signal is computed for each harmonic parameter.

Third, the pre-processed data is then decomposed using wavelet multi-resolution analysis with 5 levels on a db04 wavelet to produce several harmonic parameters. Then the Auto Power Spectrum Virtual Instrument (VI) of LabVIEW was used to compute for its single-sided scaled, auto-power spectrum of the time domain signal. It is further scaled due to a large impact of the delta power spectrum. Then it is further narrowed down by computing the power around the peak frequency in the power spectrum. And in the end the percentage of Beta 1, Beta 2, Alpha, Theta and Delta waves are calculated and used as the feature for the classification system. The outputs are then saved into a text file.

D. Support Vector Classification System

For the sleep stage classification Support Vector Machine (SVM) was utilized as the classification tool as it is one of the supervised learning methods used for classification or regression. In this paper the author made use of this method due to its capability to use hyperplane to separate data into several dimensions using the kernel function. It also makes use of soft margin to specify trade-off between hyperplane variations and the size of the margin [11]. Basically there are two types of data for SVM, the separable and the non-separable case. For this research data are considered as non-separable therefore there is the need for the kernel function. The kernel function used is the Radial Basis Function (RBF).

Below is an equation for quadratic programming to solve for SVM instances:

$$L_p = \sum_i \alpha_i - \frac{1}{2}\sum_{i,j} \alpha_i \alpha_j y_i y_j K(x_i \cdot x_j) \qquad (1)$$

Then to substitute the kernel function into the formula:

$$0 \le \alpha_i \le C \qquad (2)$$

$$\sum_i \alpha_i y_i = 0 \qquad (3)$$

Producing,

$$b = \frac{1}{2}[\min(\sum_{\{i|y_i=+1\}} \alpha_i y_i K(x_i, x_j)) + \max(\sum_{\{i|y_i=+1\}} \alpha_i y_i K(x_i, x_j))] \qquad (4)$$

In constructing SVMs, a kernel function $K(x_i,x)$ must be selected. Moreover, for the RBF kernel has simple inputs requiring the data and the C and gamma parameters.

Currently there is no analytical or empirical study that has made a conclusion on which a kernel function is the better than other kernels, therefore SVM performance varies with the choice of kernel and the task on hand [12].

E. Sleep Stage Classification

The sleep stages were divided into four groups namely wake, light sleep (LS), slow wave sleep (SWS) and rapid eye movement (REM) sleep. The 1000 epochs of training instances were extracted from 6 Physionet Polysomnographic database processed using the methods mentioned above. For the C and gamma parameter of the RBF kernel the prototype sleep stage classification system made a trial and error by using the exponential growth of the number 2.

Then a prototype of the sleep stage classification system was created together with a user friendly graphic user interface. LabVIEW was used as the base of the system, from the pre-processing, feature extraction and the sleep stage classification. The core of the classification system utilized the SVM program created by Prof. Lin of National Taiwan University. The current system is an innovation for the application of the SVM program to sleep staging [13].

III. RESULTS AND DISCUSSION

A total of 13 features were extracted from the single channel EEG taken from the database. From the training data several patterns were analyzed as sleep stage shifts from one to the other. The sample entropy values wake was found to be the highest and gradually declines as sleep goes deeper. However, it can be noticed that when sleep reaches REM the value rises again to a value between wake and LS as shown in figure 1.

It was also noted that the harmonic parameters of the band pass and wavelet filtered power spectrum provided a certain combination for each sleep stage. By taking the mean of the training data a distinct pattern was presented by each sleep stage as shown in figure 2.

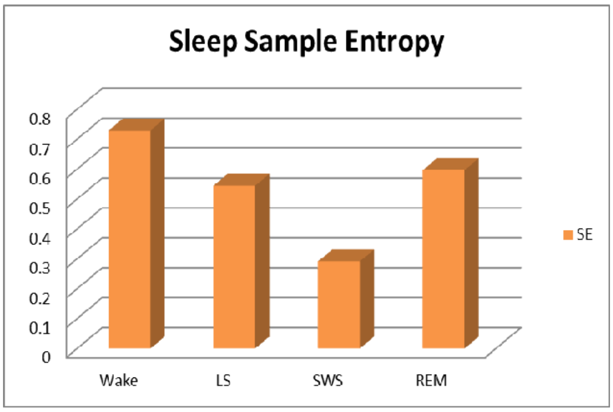

Figure 1. Sleep stage vs. sample entropy analysis

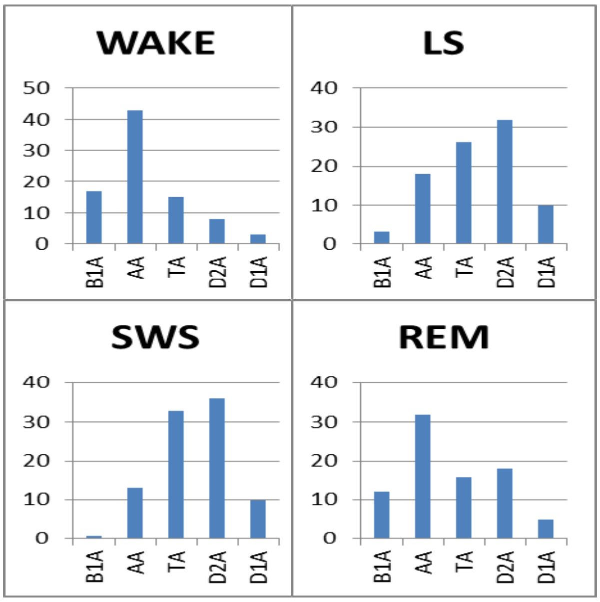

Figure 2. Harmonic parameters in different stages

The next figures will show a comparisson of wake stage and the other sleep stages. The following can be noted from the data.

First, wake showed higher values of alpha and beta with minimal theta and delta waves. Second, LS showed a drop in alpha and beta with increase in the slow waves. Third, when SWS was reached delta and theta waves became the dominant waves. However, during REM it provides a quite mixed frequency making it challenging to determine. Data of the different stages are shown in figures 2.

In future research a more in depth research on the amplitudes and frequency locations can be done to determine some certain conditions during sleep and could be possibly used for medical examinations. It can also be noted that using single channel EEG data without features from other factors is quite challenging. The processed values are still quite inconsistent would require a more advanced algorithm to do the classification. A table of the sample data and the GUI is presented below.

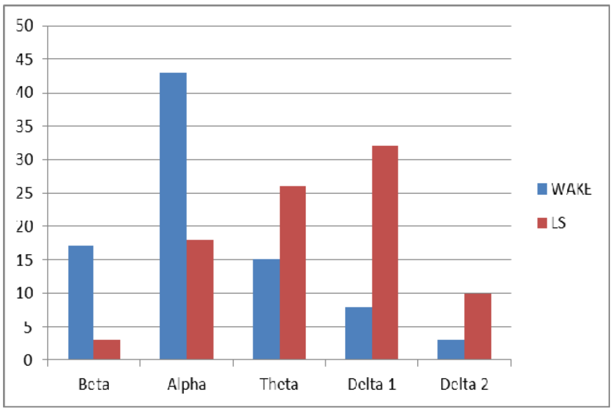

Figure 3. Harmonic frequencies in wake stage and LS stage

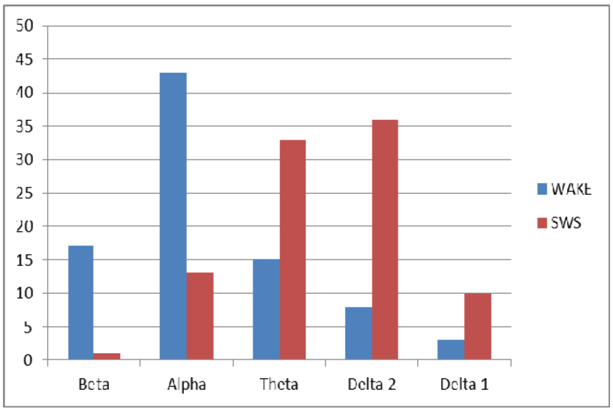

Figure 4. Harmonic frequencies in wake stage and SWS stage

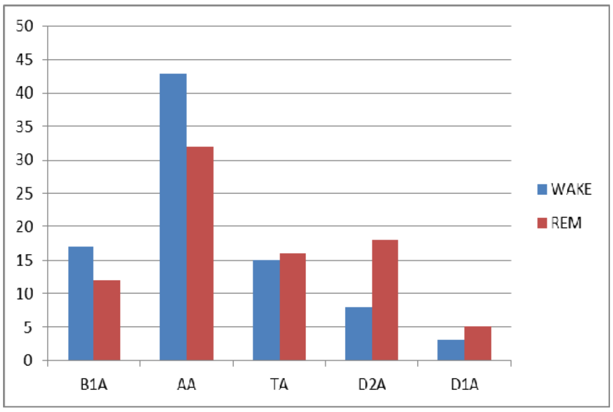

Figure 5. Harmonic frequencies in wake stage and REM stage

IV. CONCLUSIONS

The pattern observed from the two different methods relays a view on how the brain wave changes in different stages of sleep. The active brain activity can be observed through a high sample entropy, alpha and beta values. On the other hand, deep sleep will have higher delta and theta waves and low sample entropy. As for REM mixed signals can be observed but would be confusing to view in terms of sample entropy. Initial results of the training data from the developed sleep stage staging system were able to produce an accuracy of 96.2%. However, further tuning of the data and the program could be useful to test the sleep stage classification on other database or data.

REFERENCES

[1] S. Gunes, K. Polat, & S. Yosunkaya, "Efficient sleep stage recognition system based on EEG signal using k-means clustering based feature weighting," Expert Systems with Applications, vol. 37, no. 12, pp. 7922-7928, December 2010.

[2] M. H. Asyali, R. B. Berry, M. C. K. Khoo, and A. Altinok, "Determining a continuous marker for sleep depth," Computers in Biology and Medicine, vol. 37, pp. 1600-1609, 2007.

[3] T. Penzel and S. Canisius, "Polysomnography," Sleep Apnea, W. J. Randerath, B. M. Sanner and V. K. Somers, Karger, vol. 35, pp.51-60, 2006.

[4] N. Butkov and T. Lee-Chiong, Fundamentals of Sleep Technology, 1st Edition, Lippincott Williams & Wilkins, 2007

[5] A. Rechtschaffen and A. Kales, "A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects," Neurological Information Network, p. 204, 1968.

[6] G. Jiayi, Z. Peng, et al., "Sample entropy analysis of sleep EEG under different stages," IEEE/ICME International Conference on Complex Medical Engineering: pp. 1499-1502, 2007.

[7] C. Berthomier, X. Drouot, et al., "Automatic Analysis of Single-Channel Sleep EEG: Validation in Healthy Individuals," Sleep, vol. 30, no. 11, pp. 1587-1595, November 2007.

[8] D. Wilson, D. D. Rowlands, et al., "Non-Linear signal processing of electroencephalograms for automated sleep monitoring," Biomedical Applications of Micro- and Nanoengineering II , vol. 5651, pp. 95-105, 2005.

[9] E. Oropesa, H. L. Cycon, and M. Jobert., "Sleep Stage Classification using Wavelet Transform and Neural Network", International Computer Science Institute, 1999.

[10] A. L. Goldberger, L. A. N. Amaral, et al. PhysioBank, "PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," Circulation, vol. 101, no. 23, pp. e215-e220, 2000.

[11] W.S. Noble, "What is a support vector machine?," Nature Biotechnology, vol. 24, no. 12, pp. 1565-1567, December 2006.

[12] A. H. Khandoker, M. Palaniswami, et al., "Support Vector Machines for Automated Recognition of Obstructive Sleep Apnea Syndrome From ECG Recordings," IEEE Transactions on Information Technology in Biomedicine, vol. 13, no. 1, pp. 37-48, January 2009.

[13] C.-C. Chang and C.-J. Lin, LIBSVM: a library for support vector machine, 2001.

[14] E. N. Bruce, M. C. Bruce, and S. Vennelaganti, "Sample entropy tracks changes in electroencephalogram power spectrum with sleep stage and aging," J. Clin Neurophysiol, vol. 26, no. 4, pp. 257-266, August 2009.

TABLE I. SAMPLE DATA

Stage	Sample Entropy	Beta 2	Beta 1	Alpha	Theta	Delta 2	Delta 1	Beta 2	Beta 1	Alpha	Theta	Delta 2	Delta 1
0	0.788	2	22	53	22	0	0	24	32	22	17	2	3
0	1.095	1	7	73	14	1	4	5	11	71	5	4	4
1	0.584	6	8	30	46	3	8	2	4	23	31	35	5
1	0.481	2	24	19	41	5	9	1	6	17	27	39	11
3	0.252	3	4	5	68	4	15	0	2	19	23	36	21
3	0.229	0	0	3	63	2	31	0	0	9	45	36	10
4	0.247	7	37	17	37	1	1	5	10	43	14	14	13
4	0.348	1	3	63	19	3	10	5	8	43	37	2	5

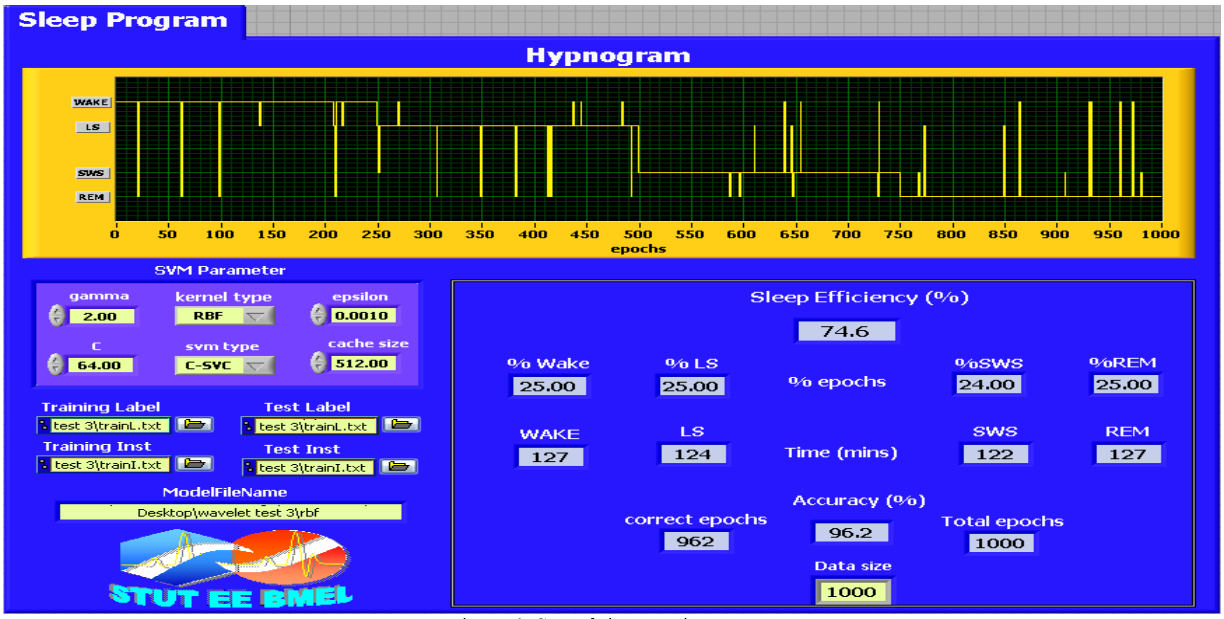

Figure 6. GUI of sleep staging system