International Journal of Social Welfare Promotionand Management
Vol. 1, No. 1 (2014), pp.9-16
http://dx.doi.org/10.21742/ijswpm.2014.1.1.02

A Computer Based Method for Classification and Analysis of the Sleep Stages using Different Statistical Parameters

Anand Prakash and Vandana Roy

DoEC, GGITS, Jabalpur, India
anand19prakash@gmail.com, vandana.roy20@gmail.com

Abstract

There are various methods and data used for the classification of sleep stages. In this proposed algorithm single channel Electroencephalograms are used for the detection of these stages. The attribute extraction for the automatic detection is done with the help of the various statistical parameters of the empirical mode decomposition with the zero crossing rate and Hjorth parameter. The extracted attribute are selected based on the Chi-squared evaluation based ranking of the data. These attributes are applied to AdaBoost M1, an ensemble method of machine learning. The classification accuracy of six stages and five stages classification are 91.47% and 92.83% respectively.

Keywords: *EEG, AdaBoost, sleep stages, EMD*

1. Introduction

Electroencephalograms (EEG) are the important bio-signals used for the analysis and detection of various health related problems. Sleep stage detection is one the application where EEG signals are used. The efficient detection of sleep stages leads to the identification and treatment of various brain related problems. The Rech-tschaffen and Kale's (R&K)[1][2] and American Academy of Sleep Medicine (ASAM) [3]criterion are two methods to define the sleep stages. In this paper R&K criterion is used for the study. Six stages of sleep are awake (A), stage 1 (S1), stage 2 (S2), stage 3 (S3), stage 4 (S4) and rapid eye movement (REM). In five stages sleep detection, S3 and S4 are combined to form a single stage known as slow wave sleep (SWS). In four stages, S1 and S2 of five stages are combined to form a single stage. In three stages, the S1, S2, S3 and S4 are combined to form the non-rapid eye movement (NREM) and two stages consist of awake and all other remaining sleep stage.

The sleep stage scoring can be given by the expert scorer. But these scoring often have human error and variation from scorer to scorer [4]. The less time requirement and the better accuracy can be obtained with the help of computer based automatic scoring of sleep stages. Various data had been used by different authors for the detection of the sleep stages. Electroencephalogram (EEG), Electromyogram (EMG) and Electrooculogram (EMG) are mainly used for this purpose. Charbonnier*et. al. [13]* used the multilayer perceptron for stage detection. Spectral and statistical attributes were extracted from the single channel EEG, EOG and EMG data. An accuracy of 85.5 % obtained for the five stages detection. Agrawal*et. al. [5]* extracted spectral attributes with the help of the two channels EEG, two channels EOG and single channel EMG data. These attributes were used for classification with the help of the K-mean clustering. An accuracy of 80% was obtained after this process.

Article history:
Received (July 15, 2014), Review Result (September 23, 2014), Accepted (October 18, 2014)

Print ISSN: 2205-8435, eISSN: 2207-5321 IJSWPM

Authors also used the single channel data for the detection of the various stages. Most of them used single channel EEG for this purpose. Rozhina*et. al.* [6]used the single channel EEG signal for the two stages, three stages, four stages and six stages classification. Spectral and statistical attributes were extracted and used in the ANN. Zhu *et. al. [7]*used the visibility graph of the single channel EEG data for the attribute extraction. The best attributes were selected from these and passed through support vector machine (SVM) with the RBF kernel. The accuracy for the two stages, three stages, four stages, five stages and six stages are 97.9%, 92.6%, 89.3%, 88.9% and 87.5% respectively were obtained. Hassan *et. al. [8]*also used single channel EEG data for the classification in two stages, three stages, four stages, five stages and six stages classification. The spectral and statistical attributes were extracted and classified with the bagging. Maximum accuracy of 95.5% was obtained in the two stages classification.

In our proposed method the attributes of single channel EEG are extracted with statistical attributes of empirical mode decomposition, Hjorth parameters and zero crossing rates. After this the attributes are selected with the help of Chi-square evaluation. At last, the selected attributes is used in the AdaBoost classifier with the REP tree as base learner. Two stages, three stages, four stages, five stage and six stages classification were performed and the result based on the accuracy shows that our method gave best result from previous works.

2. Methodology

2.1. Datasets

The datasets used in this experiment are taken from the Physionet sleep-EDF database[9][10]. Four healthy subjects (ST4001e0, ST4002e0, ST4022e0, and ST4112e0) were chosen randomly for the experiment. These databases consist of two channel EEG (Fpz-Cz and Pz-Oz) and single channel EOG data. The single channel Pz-Oz data is used for the purpose of detection sleep stages due to better accuracy [11][12]. Total numbers of 11055 epochs were obtained from the data. The details of epochs are given below in table I.

Table 1. Epochs in different stages

Stages	A	S1	S2	S3	S4	REM
Number of Epochs	7886	227	1559	360	369	654

2.2. Empirical mode decomposition (EMD)

Empirical mode decomposition (EMD) is a method for decomposing the signal into time frequency estimation of the signal [12]. The decomposed estimation is known as intrinsic mode function (IMF). The minimum criteria for the modes are

- The maximum difference between total number minima and maxima should be one.
- At any point the sum of envelope generated by the local maxima and minima should be zero.

N numbers of IMF p_i generated from this process. In fig.1, a sample signal S1 is taken and then EMD is applied. IMFs are generated after this process which is represented in the rest part of figure. The statistical parameters of each IMF are calculated. These statistical parameters are the different moments of the signal. The statistical parameters these N IMF are given. In equation 1, 2, 3 and 4, p_i is the different IMF of signal and N is number of IMF.

- Mean (m) shows the central tendency of any data. m can be calculated as given in eqn. (1).

$$m = \frac{1}{N}\sum_{i=1}^{N} p_i \tag{1}$$

- Variance (v^2) gives the dispersion of signal data against mean value (m). This is helpful in classification of REM sleep from S1 and S2. $v^2$of each N IMF can be calculated as given in eqn. no. (2):

$$v^2 = \frac{1}{N}\sum_{i=1}^{N}(p_i - m)^2 \tag{2}$$

- Skewness (s) provides asymmetry of the different signal. S of each N IMF with mean (m) and variance (v) can be calculated as given in eqn. no. (3):

$$s = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{p_i - m}{v}\right)^3 \tag{3}$$

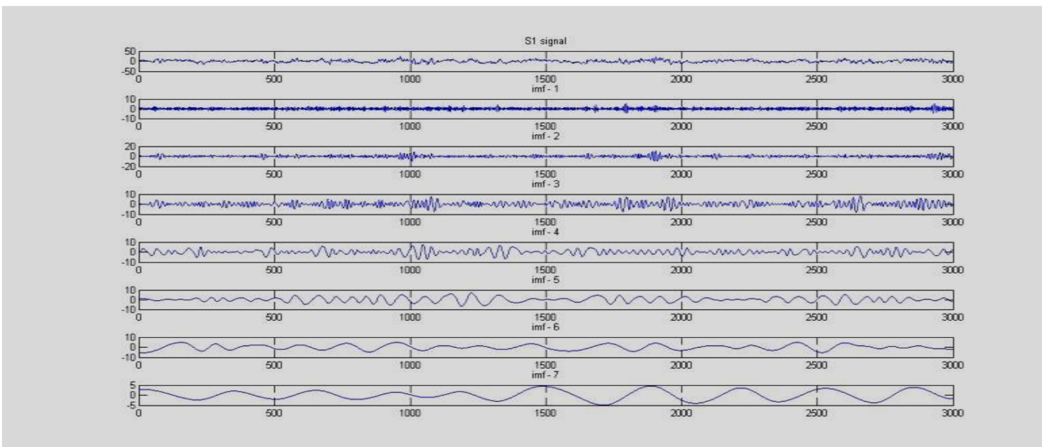

Figure 1. Sample stage 1 signal and its corresponding IMFs

- Kurtosis (k) provides the peakness value of data. Kurtosis of each N IMF with mean (m) and variance (v) can be calculated as given in eqn no (4):

$$k = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{p_i - m}{v}\right)^4 \tag{4}$$

Total 44 attributes were extracted from each IMFs different moment.

2.3. Hjorth parameter

Activity, mobility and complexity of any signal can be obtained from the Hjorth parameter [13]. Activity of signal shows its power, variance of the signal. Mobility provides a measure of standard deviation of slope with respect to standard deviation of amplitude. Complexity provides the number of standard slopes actually generated during the average time required for generation of single standard amplitude as given by the mobility. The calculating formulas for these parameters are given in the eqn. 4, 5 and 6.

$$\text{Activity} = \text{var}(X(n)) \tag{5}$$

$$\text{Mobility} = \sqrt{\frac{\text{var}(X(n)\frac{dX}{dn})}{\text{var}(X(n))}} \tag{6}$$

$$\text{Complexity} = \frac{\text{Mobility}(X(n)\frac{dX}{dn})}{\text{Mobility}(X(n))} \tag{7}$$

In above equation, X(n) shows the different sleep stage signals. $\frac{dX}{dn}$ shows the derivative of the signal X(n) in n samples.

2.4. Zero crossing rates

The zero crossing rates provide the number of time domain crossing over the mean line of the signal. This will helpful in determination of the various stages based on different number of crossings in particular window duration.

2.5. Chi-square evaluation

The attributes are passed through the Chi-square distribution. The Chi square distribution is given in below equation 5.

$$\chi^2 = \frac{(O-E)^2}{E} \tag{8}$$

We get different χ^2 value for different attribute. The ranking is provided based on this value in decreasing order. After that the best attributes were selected from the list which gives the best result.

2.6. Ada boost

AdaBoost or Adaptive Boost algorithm is a meta-learning machine learning algorithm developed by the Freund and Schapire in 1996 [14]. In this algorithm the boosting of the learner is done to improve performance by increasing the weighted sum. In the classification process the REP tree has been used for the learner process. REP tree provides the fast-learning algorithm which provides the tree based on the regression/decision based on information gain or variance of the instances. The algorithm for the AdaBoostM1 is given below.

Algorithm 1: AdaBoost M1 work flow
• Input: Generate the sequence of training set $[(X_1, Y_1), [(X_2, Y_2) ... [(X_k, Y_k)]$ Here X_k, k=1, 2, 3… K are the instance And, Y_k, k=1, 2, 3, ….. K are the labels • Take D=1/k • Do it for T iteration
1. Call REP tree learning algorithm with distribution D 2. Get back to hypothesis 3. Calculate the error 4. Update distribution for next iteration
Output of the final hypothesis is the weighted sum of all voted majority hypothesis.

The classification process for Adaboost is done through the ten-cross fold validation. The batch size is 100, number of iterations is 100, and the weighted threshold is 100. For the learner i.e., REP tree, batch size is 100 and number of seed is 2.

3. Proposed methodology

The experimental steps are followed as given in the figure 2. In first step EEG Pz-Oz epochs of 30s are taken. It is done with the help of polyman. In next step each epoch is used for attribute extraction is done with the help of EMD statistical parameters, Hjorth parameters and Zero crossing rate in MATLAB. In next step the extracted attribute was selected with Chi-square evaluation. After that the attributes were used in the machine learning. For the machine learning in the experiment, AdaBoost is used. Tenfold classification process is used for the classification of data. In each fold the $1/10^{th}$ part is used for the testing and rest is used for the classification. In next step next $1/10^{th}$ is used for the testing purpose. All this process is performed in the WEKA.

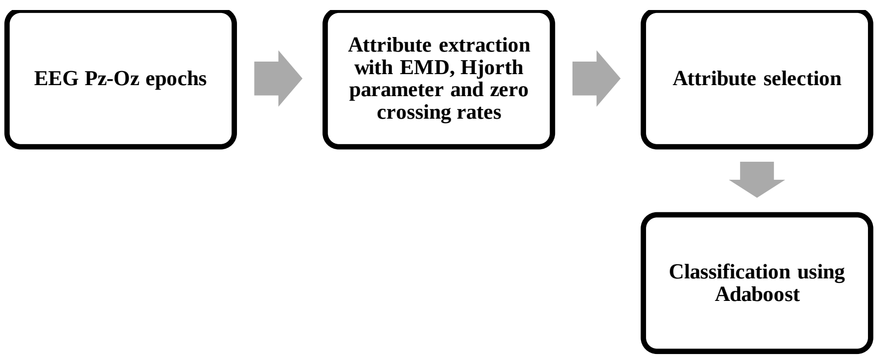

Figure 2. Flow diagram of experiment

4. Results and discussion

This experiment is performed on the computer with Pentium Quad core processor, 2.17 Ghz clock speed and 4 GB RAM on Windows 10 platform.

Table 2. Comparison table of accuracy

Stages	Zhu *et. al.* (2014)	Hassan *et. al.* (2016)	Proposed method
2-stages	97.9%	95.05%	**98.01%**
3-stages	92.6%	89.77%	**94.78%**
4-stages	89.3%	87.49%	**92.66%**
5-stages	88.9%	86.53%	**92.15%**
6-stages	87.5%	85.57%	**90.78%**

Our proposed method with AdaBoost classifier is compared with the other author. These authors also used the sleep-EDF data of Physionet. Zhu *et. al.* [7] used the difference visibility graph and horizontal visibility graph of the single channel EEG. These gave the attributes which were further used in SVM for the classification. The accuracies of two stages, three stages, four stages, five stages and six stages are 97.9%, 92.6%, 89.3%, 88.9%

and 87.5% respectively. Hassan *et. al. [8]* used the statistical and spectral attributes for the classification with bootstrap aggregation. In this experiment, the author got the result for two stages, three stages, four stages, five stages and six stages with 95.05%, 89.77%, 87.49%, 86.53% and 85.57% accuracies respectively. The major reason for the lower accuracy of this author is a smaller number of attributes for the classification. In our proposed methodology of applying the Adaboost classifier with the statistical attributes of EMD, Hjorth and Zero crossing rates gave the better result. The accuracy for the two stages, three stages, four stages, five stages and six stages are 98.01%, 94.78%, 92.66%, 92.15% and 90.78%. In each classification criteria, the proposed work performed better.

Table 3. Confusion matrix of six stages classification

	Proposed method							
Expert scoring		S1	S2	S3	S4	R	A	Sensitivity (%)
	S1	32	58	0	0	83	54	14.1
	S2	6	1312	88	28	93	32	84.2
	S3	0	122	142	87	2	7	39.2
	S4	0	23	59	284	0	3	77
	R	10	119	0	0	479	46	73.2
	A	14	32	5	1	47	7787	98.7
	Specificity (%)	0.3	3.7	1.4	1.1	2.2	4.5	

[Table 3] shows the confusion matrix of the six-stage classification. From the above table, we can observe that the sensitivity for S2, S4, and R were good with 84.4%, 77% and 73.2% respectively, but in case of S1 it gave the poor sensitivity 10%. Most of the S1 data are misclassified in the REM stage. The real scenario has the same problem in detection of S1 from REM is a tedious job. Also, the data of S2 were misclassified in S2 and A stage. The other reason for lower sensitivity of the signal can be the lower number of samples provided for the training set. Average sensitivity is obtained for the S3 stage. Specificity of the S1 is better from all other shows that the lesser amount of data from other states are misclassified in this category. The awake (A) stage shows the least specificity as compared to others.

Table IV shows the confusion matrix of five stage classification. In this stage the data of S3 and S4 are combined to form the SWS. In this table, we can see that when the data of SWS gave the better sensitivity than the S3 and S4 in table III. In the five-stage classification the sensitivity for the S1 is poor. For the S2, R, and A stages, better sensitivities can be observed. Also, specificity is best for the S1 stage and the improvement in specificity when S3 is added with the S4.

Table 4. Confusion matrix of five stage classification

	Proposed method						
		S1	S2	SWS	R	A	Sensitivity (%)
Expert scoring	S1	33	56	0	77	61	14.5%
	S2	7	1311	113	101	27	84.1%
	SWS	0	129	589	2	9	80.8%
	R	13	115	0	470	56	71.9%
	A	13	37	8	43	7785	98.7%
	Specificity (%)	0.3%	3.5%	1.2%	2.1%	4.8%	

We get the better sensitivity when the numbers of stages are less for the classification. From table III and IV we can see the improvement in each stage except for the A.

4. Conclusion

In our proposed method, we got the better accuracy from previous author with increase of minimum 0.11% in stage 2. The maximum accuracy increase was observed in the case of five stage classification. The sensitivity of the stage 1 is poor in both the cases of five stage and six stage classification. This can be increased with the help of increasing number of instances and selecting other attribute for better difference of stage 1 form other stages. The overall sensitivity obtained in six states is 90.6% and five stages are 92.2%. The specificity of the proposed method has good value for each of them. The value on the higher side is the awake stage. We can observe that various factor affects the accuracy, sensitivity and specificity. Number of instances in training data gives the better value of these parameters. Other factor is the attribute chosen for the classification of the data. Some time they work better in classifying some part and they leads to the misclassification. Number of classification stages can also affect the classification parameter. This proposed work will work better for the five-stage classification due to the better sensitivity and specificity of the different stages of sleep.

References

[1] A. Rechtschaffen and A. Kales, “Manual of Standardized Terminology, Techniques and Scoring Systems for Sleep Stages of Human Subjects”, U.G.P. Office, Washington DC Public Health Service, **(1968)**.

[2] A.L.C.C. Iber, S. Ancoli-Israel and S.F. Quan, “The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specification,” American Academy of Sleep Medicine, Westchester, USA, **(2005)**.

[3] N.A. Collop, “Scoring Variability betweenPolysomnography Technologists in Different Sleep Laboratories,” Sleep Medicine, Vol.3, pp. 43-47, **(2002)**.

[4] S. Charbonnier, L. Zoubek, S. Lesecq and F. Chapotot, “Self-evaluated Automatic Classifier as a Decision-support Tool for Sleep/Wake staging”,Computer Biology in Medicine, Vol. 41, pp. 380-389, **(2011)**.

[5] R. Agarwal and J. Gotman, “Computer-assisted Sleep Staging,” IEEE Transactions on Biomedical Engineering, Vol. 48, No. 12, pp. 1412-1423, **(2012)**.

[6] M. Ronzhina, O. Janouek, J. Kolrov, M. Novkov, P. Honzk and I. Provaznk, “Sleep Scoring using Artificial Neural Networks”, Sleep Medicine Reviews, Vol. 16, No. 3, pp. 251-263, **(2012)**.

[7] G. Zhu, Y. Li and P. Wen, “Analysis and Classification of Sleep Stages Based on Difference Visibility Graphs from a Single Channel EEG Signal,” IEEE Journal of Biomedical and Health Informatics, Vol. 99, No. 99, pp. 1-8, **(2014)**.

[8] A.R. Hassan, S. K. Bashar and M.I.H. Bhuiyan, "On the Classification of Sleep States By Means of Statistical and Spectral Features from Single Channel Electroencephalogram," Journal of Neuroscience Methods, Vol. 271, pp.2238-2243, **(2015)**.

[9] B. Kemp, "The Sleep-EDF Database", Jun**(2013)**. [Online]. Available: http://www.physionet.org/physiobank/database/sleep-edf/

[10] B. Kemp, A.H. Zwinderman, B. Tuk, H.A.C. Kamphuisen and J.J.L. Oberye, "Analysis of a Sleep-Dependent Neuronal Feedback Loop: the Slow-Wave Micro continuity of the EEG," IEEE Transaction in Biomedical Engineering, Vol. 47, No. 9, pp. 1185-1194, September **(2000)**.

[11] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.K. Peng and H.E. Stanley, "PhysioBank, Physiotoolkit, and Physionet: Components of a New Research Resource for Complex Physiologic Signals," Circulation, Vol. 101, pp. e215-e220, June 13, **(2000)**.

[12] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung and H.H. Liu, "The Empirical Mode Decomposition Method and the Hilbert Spectrum for Non-stationary Time Series Analysis", Procurement Royal Society London, Vol. 454A, pp. 903-995, **(1998)**.

[13] S. Charbonnier, L. Zoubek, S. Lesecq and F. Chapotot, "Self-evaluated Automatic Classifier as a Decision-support Tool for Sleep/Wake staging", Computer in Biology Medicine, Vol. 41, pp. 380-389, **(2011)**.

[14] Y. Freund and R.E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", Journal of Computer System and Science, Vol. 55, No. 1, pp. 119-139, **(1997)**.