Effect of EEG Time Domain Features on the Classification of Sleep Stages

Sule Yucelbas¹, Seral Ozsen², Cuneyt Yucelbas², Gulay Tezel¹, Serkan Kucukturk³ and Sebnem Yosunkaya³

¹Department of Computer Engineering, Selcuk University, Konya, Turkey; syucelbas@selcuk.edu.tr, gtezel@selcuk.edu.tr
²Department of Electrical and Electronics Engineering, Selcuk University, Konya, Turkey; seral@selcuk.edu.tr, cyucelbas@selcuk.edu.tr
³Sleep Laboratory, Faculty of Medicine, Necmettin Erbakan University, Konya, Turkey; biyolog_serkan@hotmail.com, syosunkaya@gmail.com

Abstract

Background/Objectives: Studies on the field of automatic sleep stage classification have been taking more attention of researchers day by day. Noise in the recordings, nonlinear dynamic feature of EEG signals and some other reasons affect the performance of proposed systems in negative manner. Methods/Statistical Analysis: Sleep can be divided five main stages as Wake, Non-REM1, Non-REM2, Non-REM3 and REM. Almost every proposed method can successfully classify some evident stages like Non-REM2 and REM. But when it comes to the transitions between stages, the systems are not very good in their performances. Thus a different classification strategy was proposed in this study. Five different classifiers were designed especially for transitions between stages using time domain features of EEG, EOG and EMG signals and evaluated these features for each classifier. Sequential backward feature selection process was applied in each classifier to find out which features are dominant in each classification procedure. Artificial Neural Networks was used in designed classifiers. Findings: The highest classification accuracy was obtained as 91.03% for Classifier-3 which predicts stages coming after Non-REM II. The lowest accuracy was recorded as 75.42% for Classifier-2 in which stages are determined after the Non-REM I epochs. Comparatively good results were reached especially if it is taken into account that only used time-domain features of signals. Application/Improvements: The obtained results show that the designed classifiers can be used in automatic sleep staging system, confidently.

Keywords: ANN, Automatic Sleep Stage Classification, EEG, EMG, EOG, Feature Selection

1. Introduction

Automatic sleep staging is not a new field of research but it is among the hard problems of classification. So, studies regarding to this field seem to be heavily continue in next years. Sleep staging is such a process that two sleep experts can conclude different staging patterns for the same records. Rules are standard and staging should be done

R&K Rules of ASSM¹. Sleep consists of 5 stages named as Wake, Non-REM I, Non-REM II, Non-REM II (sometimes called as Slow Wave Sleep- SWS) and REM. The classification of sleep stages is done on this R&K Rules. Sleep is divided small time part named as epochs and each epoch is labelled with these stages according to some special signal characteristics. The hardness of the process lies in seeing special characters of signals. For example, while

*Author for correspondence
one sleep expert can label a signal pattern as sleep spindle in EEG signal, other expert may not see this pattern. Thus, experience and interpretation difference between experts can result different sleep patterns. With the use of automatic sleep staging systems this bottleneck can be eliminated. Because one standardized computer program, which is prepared with the advice of very experienced sleep experts, will SEE and label stages without needing experience and interpretation. Researchers have been trying to do this for years but it can be said that there isn’t any well performed automatic staging system which do its job properly with high classification ratios.

In their study, Hae-Jeong Park et al. compared Rule Based Reasoning (RBR) and Case Based Reasoning (CBR) in sleep stage classification and found that CBR reached better results. Masaaki Hanaoka et al. used decision trees for the same purpose but not obtained good performance -approximately 70%- with their applications. Arthur Flexer et al. applied Hidden Markov Model (HMM) to a single channel sleep EEG signal to score sleep stages and reached performances near 80% but they classified sleep stages as Wake-light sleep and deep sleep. In another study, Luka´s Zoubek et al. extracted time and frequency domain features from the EEG, EOG and EMG signals and evaluated sequential feature selection on these features to have highest accuracy. They reached accuracies about %80s: the highest accuracy for Non-REM III with 92.90% and the lowest accuracy as 64.56% for Non-REM I. In 2008, Kristina Susmakova and Anna Krakovska conducted a comprehensive study to search the discrimination ability of 73 measures computed for various PSG channels (818 measures in total). They used Fisher Discriminant Analysis (FDA) to see the discrimination capability of their measures. Farideh Ebrahimi et al. used Wavelet Packet Coefficients in sleep staging. They achieved a classification accuracy of 93% but they used signals from Physionet database and they combined Non-REM I and REM stages (they have very similar characteristics). In their study, Sheng-Fu Liang et al. utilized Fuzzy Inference System to classify five sleep stages and obtained 87% accuracy. In their another study, Anna Krakovska and Kristina Susmakova tried to find optimal combination of features in sleep stage scoring by using quadratic discriminant analysis. They obtained an agreement ratio of 74% with two sleep experts. Besides, Sheng-Fu Liang et al. used Automatic Stage Scoring of Single-Channel Sleep EEG by Using Multiscale Entropy and Autoregressive Models. In another study, Arnaud Brignol et al. utilized Phase space and power spectral approaches in automatic sleep stage scoring. Also they compared short and standard epoch lengths in their system.

There are many other studies like above mentioned ones. Some of them have reached good results as 90% but when the details of these studies are read, they either used very clean signals from the databases available in Internet or classified stages as Wake-light sleep-deep sleep. A huge part of all night sleep consists of Non-REM II stage and it is not so hard to determine this stage for automatic classifiers. Thus, even a poor classifier can reach accuracies near 75%-80%. The hardness of sleep staging lies in discrimination of stage transitions. In literature, there is no study about this discrimination of transitions. By knowing this, we proposed 5 classifiers which were prepared for transitions from stages Wake, Non-REM I, Non-REM II, Non-REM III and REM. We prepared input data by extraction time domain features from the EEG, EOG and EMG signals for each classifier. By applying Sequential Feature Selection (SFS) to these features, we tried to determine which features are important in which stage transitions. Artificial Neural Networks (ANN) was utilized in designed classifiers. Details and results of the study are given in the succeeding chapters.

2. Material and Methods

2.1 Used Dataset

We used EEG, EOG and EMG signals from the Polysomnographic (PSG) recordings of fifteen subjects which were recorded at Meram Medicine Faculty of Necmettin Erbakan University in Konya/Turkey. The distribution of epochs in each subject to each stage is shown in Table 1.

For the purpose of pre-processing, EEG and EOG signals were filtered with 5th order butterworth bandpass filter with the cut-off frequencies of 0.3Hz-35Hz. The cut-off frequencies of the used filter for the EMG signals on the other hand were 1Hz-100Hz. After filtering process, ECG artefact removal and outlier elimination processes were conducted. Figure 1 shows the original and cleaned EEG signals in an epoch after pre-processing processes.
2.2 Feature Extraction and Proposed Classifiers

As stated in introduction, most automatic sleep stage classifiers fail in transitions between stages. Thus, we proposed five different classifiers as Classifier-1, Classifier-2, Classifier-3, Classifier-4 and Classifier-5. We used the following features extracted from EEG, EOG and EMG signals for all classifiers:

1. Energy of EEG signal
2. Mean absolute value of EEG signal
3. Zero crossing rate of EEG signal
4. "Standard deviation/mean absolute value" ratio of EEG signal
5. Skewness of EEG signal:

\[ x_{skw} = \frac{\sum_{n=1}^{N} (x(n) - x_{m})^3}{(N - 1)x_{std}^3} \]  

(1)

Here, \( N \) is the number of data points in x signal, \( x_{m} \) is the mean value and \( x_{std} \) is the standard deviation of the x signal.

| Subjects | Stages | Wake | Non-REM I | Non-REM II | Non-REM III | REM | TOTAL |
|----------|--------|------|-----------|------------|-------------|-----|-------|
| Subject-1 |       | 19   | 84        | 716        | 34          | 199 | 1052  |
| Subject-2 |       | 153  | 126       | 560        | 93          | 265 | 1197  |
| Subject-3 |       | 142  | 264       | 665        | 57          | 92  | 1220  |
| Subject-4 |       | 63   | 62        | 436        | 109         | 121 | 791   |
| Subject-5 |       | 76   | 106       | 557        | 75          | 93  | 97    |
| Subject-6 |       | 73   | 112       | 423        | 100         | 143 | 851   |
| Subject-7 |       | 140  | 142       | 412        | 86          | 117 | 897   |
| Subject-8 |       | 178  | 88        | 545        | 28          | 171 | 1010  |
| Subject-9 |       | 164  | 94        | 495        | 91          | 125 | 969   |
| Subject-10 |      | 224  | 46        | 403        | 0           | 66  | 739   |
| Subject-11 |      | 131  | 90        | 383        | 55          | 49  | 708   |
| Subject-12 |      | 57   | 136       | 656        | 23          | 162 | 1034  |
| Subject-13 |      | 79   | 54        | 456        | 77          | 162 | 828   |
| Subject-14 |      | 67   | 62        | 552        | 91          | 157 | 929   |
| Subject-15 |      | 335  | 86        | 337        | 62          | 84  | 904   |
| TOTAL     |       | 1091 | 1552      | 7596       | 981         | 2006| 13226 |
6. Kurtosis of EEG signal:

\[ x_{kurt} = \frac{\sum_{n=1}^{N} (x(n) - x_m)^4}{(N-1)x_{std}^4} \]  

(2)

7. Hjorth mobility of EEG signal:

\[ \text{Mobility} = \sqrt{\frac{\text{var}(y(t))}{\text{var}(y(t))}} \]  

(3)

Here, \( \text{var}(\cdot) \) stands for variance.

8. Hjorth complexity of EEG signal:

\[ \text{Complexity} = \frac{\text{Mobility}(y(t))}{\text{Mobility}(y(t))} \]  

(4)

9. Entropy of EEG signal\(^1\)

10. Spectral entropy of EEG signal\(^1\)

11-20: Features 1-10 were calculated for left eye EOG signal

21-30: Features 1-10 were calculated for right eye EOG signal

31-40: Features 1-10 were calculated for (right eye EOG – left eye EOG) signal.

41-50: Features 1-10 were calculated for EMG signal

**Classifier-1:** Classifier-1 was designed to predict which stage will come after the current epoch which is stage Wake. According to the sleep experts and R&K rules\(^1\), Wake stages can be followed by either Wake stages or Non-REM I stages. Thus, a two-class system shown in Figure 2(a) was formed.

There were 1562 epochs of Wake stages and 384 epochs of Non-REM I stages in 15 patients’ data.

**Classifier-2:** Classifier-2 was designed to predict which stage will come after the current epoch which is stage Non-REM I. According to the sleep experts and R&K rules\(^1\), Non-REM I stages can be followed by Wake, Non-REM I, Non-REM II and REM stages. Thus, a four-class system shown in Figure 2(b) was formed.

---

**Figure 1.** EEG signals before and after the pre-processing steps.
There were detected 775 epochs whose previous epochs were belonging to Non-REM I stage. Among them, 82 epochs were belonging to Wake stage, 276 epochs were of Non-REM I stage and 404 of them were in Non-REM II stage. Only 15 epochs were REM stage of these 775 epochs.

**Classifier-3:** Classifier-3 was designed to predict which stage will come after the current epoch which is stage Non-REM II. A five-class system shown in Figure 2(c) was formed.

7624 epochs were found whose previous epochs were belonging to Non-REM II stage. The distribution of them was as follows:

- Wake: 194
- Non-REM I: 153
- Non-REM II: 7182
- Non-REM III: 36
- REM: 59

**Classifier-4:** Classifier-4 was for predicting the stages after Non-REM III stages. Again all of the stages can follow Non-REM III stage but we didn't observe any REM stage following Non-REM III. Thus, a four-class system shown in Figure 2(d) was formed.

There were observed 937 epochs whose previous epochs were belonging to Non-REM III stage. Among them, only 12 epochs were belonging to Wake stage, 4 epochs were of Non-REM I stage and 21 of them were in Non-REM II stage. Remaining 910 epochs were Non-REM III stage again.

**Classifier-5:** For estimating stages coming after REM stages, a four-class system shown in Figure 2(e) was formed. Again, all stages can follow REM stages according to but we didn't see and Non-REM III stage coming after REM stage.
The distribution of 1919 epochs-which are the epochs coming after REM stages- is given as: Wake: 38, Non-REM I: 14, Non-REM II: 19, REM: 1848

3. Application Results

Sequential Backward Feature Selection (SBFS) process\(^3\) was applied in each classifier to find out which features are dominant in each classification procedure. In this feature selection procedure, all of the features are used in the beginning and one feature is dropped from the feature set in each step. The selection of that feature is done according to the classification performance. That is, if the classification accuracy will be highest when the related feature is dropped from the set, that feature is selected to be dropped from the set. The procedure is continued by deleting one feature in one step until 2 feature remains in the feature set. Then, the best feature configuration giving the highest classification accuracy will be selected as the best feature set.

While conducting above mentioned feature selection procedure, one best ANN parameter – number of Hidden Node Number (HNN) - was searched to find best classification result. One and/or two hidden layer and gradient descent learning algorithm (traingdx) was used in ANN. Learning Rate (LR), Momentum Constant (MC) and Maximum Iteration Number (MAX ITER) parameters were fixed at values 2, 0.8 and 1000 respectively. We didn’t change these parameters because changing them didn’t resulted important differences in classification results. The number of optimum hidden nodes was found by changing HNN between 1 and 100 with steps 1. Also, 3-fold cross validation\(^4\) technique was utilized in each ANN classification trial.

The obtained classification accuracies for each classifier are shown in Table 2. For the first classifier, the change of classification accuracy with respect to used features and the feature number are given in Figure 3. As shown in the figure, a few number of features is unsuccessful as well as high number of features.

When the results in Table 2 are analyzed, highest classification accuracy among the five classifiers was reached in Classifier-3 with a classification accuracy of 91.03%. The aim in this classifier was to determine the stage of epochs whose previous epochs was in Non-REM II. The largest number of input dataset is in this classifier: 7624. The lowest accuracy on the other hand was reached in Classifier-2. This classifier finds out which stage will come

| Classifier: | Best feature combination | Optimum ANN parameters | Highest classification accuracy (%) |
|-------------|--------------------------|------------------------|-------------------------------------|
| Classifier-1 | 3,4,5,7,12,15,22,27, 31,32,34,41,42 (13 features) | One hidden layer: HNN_1: 32 | 88.23 |
| Classifier-2 | 2,3,4,10, 13,17,25,26,31,34,38,39 (12 features) | Two hidden layer: HNN_1=53, HNN_2: 7 | 75.42 |
| Classifier-3 | 1,2,5,8,13,14,18,22,23,24,27,33,38,39,41,42 (16 features) | Two hidden layer: HNN_1: 27, HNN_2: 3 | 91.03 |
| Classifier-4 | 1,2,3,4,6,8,10, 17,22,34,41,42,49 (13 features) | Two hidden layer: HNN_1: 78, HNN_2: 9 | 78.64 |
| Classifier-5 | 2,4,7,8,12,13,14,17,25,28,32,34,37,38,39,42,49 (17 features) | Two hidden layer: HNN_1: 31, HNN_2: 12 | 80.31 |
after Non-REM I stages. There were only 775 epochs in dataset for this classifier. Also, when the studies in literature were searched, it can be seen that it is difficult to discriminate between Non-REM I, Wake and REM stages.

4. Discussion and Conclusion

There are many automatic sleep stage classification studies in literature. While using a very large set of features ranging from time-, frequency-based features to dynamic, statistical and nonlinear features, problems regarding to reaching high accuracies still remains. Especially, methods generally fail in correctly classifying transition epochs between stages. Thus, we concentrated our attention at transition epochs. We designed 5 classifiers to predict which stage will come after the stages Wake, Non-REM I, Non-REM II, Non-REM III and REM respectively by using time domain features of EEG, EOG and EMG signals and ANN classifier. With the use of 50 features extracted from the above signals, we applied Sequential Backward Feature Selection in each classifier to determine which features are determinative in selecting stages for each classifier. By doing so, we obtained highest classification accuracy of 91.03% for Classifier-3 which predicts stages coming after Non-REM II. Lowest obtained accuracy was recorded as 75.42% for Classifier-2 in which stages are determined after the Non-REM I epochs. We reached comparatively good results especially if it is taken into account that we only used time-domain features of signals. We studied also on integrating frequency-domain and other statistical analysis techniques into this study to reach higher results especially for Classifier-2 and Classifier-4.

5. Acknowledgement

This study is supported fully by the TUBITAK Research Project under Grant No. 113E591. And also, this study is supported by the Scientific Research Projects of Selcuk University.

6. References

1. Rechtschaffen A, Kales A. A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. US Government Printing Office: Washington; 1969.
2. Park HJ, Oh JS, Jeong DU, Park KS. Automated sleep stage scoring using hybrid rule-and case-based reasoning. Computers and Biomedical Research. 2000; 33:330–49.
3. Hanaoka M, Kobayashi M, Yamazaki H. Automated sleep stage scoring by decision tree learning. Proceedings of 23rd Annual EMBS International Conference, Istanbul: Turkey; 2001. p. 1751–4.
4. Flexera A, Gruber G, Dorffner G. A reliable probabilistic sleep stager based on a single EEG signal. Artificial Intelligence in Medicine. 2005; 33:199–207.
5. Zoubek L, Charbonnier S, Lescq S, Buguet A, Chapotot F. Feature selection for sleep/wake stages classification using...
data driven methods. Biomedical Signal Processing and Control. 2007; 2:171–9.

6. Susmakova K, Krakovska A. Discrimination ability of individual measures used in sleep stages classification. Artificial Intelligence in Medicine. 2008; 44:261–77.

7. Ebrahimi F, Mikaeili M, Estrada E, Nazeran H. Automatic sleep stage classification based on EEG signals by using neural networks and wavelet packet coefficients. Proceedings of 30th Annual International IEEE EMBS Conference, Vancouver: BC; 2008. p. 1151–4.

8. Liang SF, Chen YH, Kuo CE, Chen JY, Hsu SC. A fuzzy inference system for sleep staging. Proceedings of IEEE International Conference on Fuzzy Systems, Taipei: Taiwan; 2011. p. 2104–7.

9. Krakovska A, Mezeiova K. Automatic sleep scoring: A search for an optimal combination of measures. Artificial Intelligence in Medicine. 2011; 53:25–33.

10. Liang SF, Kuo CE, Hu YH, Pan YH, Wang YH. Automatic stage scoring of single-channel sleep EEG by using multiscale entropy and autoregressive models. IEEE Transactions on Instrumentation and Measurement. 2012; 61(6):1649–57.

11. Brignol A, Al-ani T, Drouot X. Phase space and power spectral approaches for EEG-based automatic sleep–wake classification in humans: A comparative study using short and standard epoch lengths. Computer Methods and Programs in Biomedicine. 2013; 109:227–38.

12. Borda M. Fundamentals in Information Theory and Coding. Springer; 2011.

13. Improving feature selection techniques for machine learning [Internet]. [Cited 2015 Apr 01]. Available from: http://scholarworks.gsu.edu/cs_diss/27.

14. Devijver PA, Kittler J. Pattern recognition: A statistical approach. Prentice-Hall; 1982.