Automatic Sleep Stage Detection and Classification: Distinguishing Between Patients with Periodic Limb Movements, Sleep Apnea Hypopnea Syndrome, and Healthy Controls Using Electrooculography (EOG) Signals

Emad Malaekah*, Sobhan Salari Shahrbabaki and Dean Cvetkovic
School of Electrical and Computer Engineering, RMIT University, Australia

Abstract

Background: To improve the diagnostic and clinical treatment of sleep disorders, the first important step is to identify or detect the sleep stages. Utilizing the conventional method-known as visual sleep stage scoring-is tedious and time-consuming. Therefore, there is a significant need to create or develop a new automatic sleep stage detection system to assist the sleep physician in evaluating the sleep stages of patients or non-patient subjects. The first aim of this study is to develop an algorithm for automatic sleep stage detection based on Electrooculography (EOG) signals. The second aim is to utilize sleep quality parameters to classify and screen Periodic Limb Movements of Sleep (PLMS) patients and Sleep Apnea Hypopnea Syndrome (SAHS) patients, as distinct from healthy control subjects.

Methods: 10 patients with Periodic Limb Movements of Sleep (PLMS), 10 patients with Sleep Apnoea Hypopnea Syndrome (SAHS), and 10 healthy control subjects were utilised in this study. Several features were extracted from EOG signals such as cross-correlation, energy entropy, Shannon entropy and maximal amplitude value. K-Nearest Neighbour was used for the classification of sleep stages. Several polysomnographical (PSG) features were measured for screening and classification of the sleep disorders, such as the percentage of the sleep stages over the total time of sleep, the duration of the sleep stages, Sleep Latency (SL), and sleep efficacy. A decision tree analysis was utilised for identifying the three groups of subjects.

Results: The overall accuracy, sensitivity and specificity of automatic sleep stage detection were 80.5%, 81.3% and 88.8%, respectively. The Cohen’s Kappa was 0.73. The performance of the classified sleep disorders showed an overall accuracy of 90%. The sensitivity and specificity were 90% and 95%. The Cohen’s Kappa was 0.85.

Conclusion: One advantage of the automatic sleep stage detection method based on Electrooculography (EOG) signals is that it can be utilized with portable sleep stage recording instead of using a multichannel signal. Classification of sleep disorders based on the automatic system is an improvement, in that it can make the screening or diagnostic processes much faster and easier than with other methods.

Keywords: Sleep stage; Sleep scoring; Polysomnography; Electrooculography; Sleep quality

Introduction

The sleep phenomenon has gained reasonable scientific interest for an extended time. Sleep refers to a behavioral state that varies from wakefulness by a loss of reactivity, readily and reversibly, in relation to events within one’s environment [1]. Sleep can be categorized into two primary and distinct behaviors: NREM (Non-Rapid Eye Movement) sleep and REM (Rapid Eye Movement) sleep [2]. NREM is categorized as light sleep, termed N1 and N2 (the latter is further broken down into S1, S2), and deep sleep, which is termed N3 (and further broken down into S3 and S4) [3]. Deep sleep is also known as Slow-Wave Sleep (SWS). The abbreviations W, N1, N2, N3, and R are derived from the new standard of Iber and colleagues [4].

Performing Polysomnography (PSG) entails a comprehensive sleep study assessing numerous physiological signals such as an Electroencephalogram (EEG), an Electrooculogram (EOG), an Electromyogram (EMG), respiratory effort, an Electrocardiogram (ECG), and others. It is the gold standard for measuring sleep states [5], sleep quality, and sleep quantity. The manual scoring of sleep stages based on EEG, EOG and EMG is a subjective and time-consuming process; hence the need for comprehensive and more accurate automatic techniques that are easy to apply and can be used in experimental and clinical ambulatory research.

Several attempts have been made to utilise the EEG or EOG signals only for detecting sleep stages, or to detect only one particular sleep stage, such as Slow-Wave Sleep (SWS) [6,7].

EEG automatic detection has been employed for detecting sleep stages. The method was comprised of four steps: segmentation, extraction of parameters, analysis of cluster, and classification. The parameters compared included the harmonic parameters, Hjorth, and relative band energy [8]. An automatic algorithm used by Liang [9] for detection of SWS utilized one or two EOG/EEG channels. The result of this study obtained 80% sensitivity, and a Cohen’s kappa value of 0.755.

“Sleep disorder” refers to a medical condition in the patterns of sleep of an animal or human being, also known as somniphagy [10]. The classification of sleep disorders is essential in order to differentiate

*Corresponding author: Emad Malaekah, School of Electrical and Computer Engineering, RMIT University, 376-392 Swanston Street, GPO Box 2476V, Melbourne, VIC 3001, Australia, Tel: +613-992-596-41; Fax: +613-992-520-07; E-mail: s3321008@student.rmit.edu.au

Received February 26, 2015; Accepted March 19, 2015; Published March 23, 2015


Copyright: © 2015 Malaekah E, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
between disorders, and to enhance understanding of etiology, pathophysiology and symptoms, thus enabling appropriate treatment [11]. The Pittsburgh Sleep Quality Index (PSQI), which was developed by Buysse [12], has been used as a standardised subjective measure to evaluate sleep quality. PSQI is based on several questions relating to the evaluation of psychometric properties of sleep quality for the duration of one month. The periodic limb movement disorder is defined as a nearly irresistible urge to move legs while asleep [13]. Studies indicate that PLMS occurs in stages 1 or 2 of the sleep period before REM sleep. On the other hand, Obstructive Sleep Apnea Hypopnea Syndrome (OSAHS) has been on the increase in the last fifty years, with significant morbidity rates in both developing and developed countries. OSAHS also causes daylight sleepiness [14]. Sleep apnea hypopnea syndrome leads to fragmentation of sleep and limits the quantity of time spent in the deeper sleep stages 3 and 4.

In this study we focused on two main objectives. The first aim was to develop an automatic sleep stage detection method based on two EOG signals, as compared with the manual sleep stage detection system based on EOG, EEG, and EMG signals. The second important aim was to develop an automatic system for classifying different sleep disorders based on the sleep quality extracted from the sleep stages.

Materials and Methods

Participants and data collection

The PSG data was downloaded from the online database [11]. The PSG signals included three EEG signals (C3-AI, FP1-A1 and O1-A1), two EOG signals, and one submental EMG channel. The Right EOG (REOG) and Left EOG (LEOG) signals were the only signals utilised from this PSG dataset. This PSG data was then recorded from 10 healthy controls: 7 females aged 20-65 years (average age: 40 years), and 3 males aged 20-27 years (average age: 23.5 years). Next, these signals were taken in 10 patients with Periodic Limb Movements of Sleep (PLMS): 8 adult males aged 31-71 years (average age: 51 years), and 2 adult females aged 27-69 years (average age: 48 years). Finally, signals were recorded from 10 patients with Sleep Apnea Hypopnea Syndrome (SAHS): 6 adult males aged 38-73 years (average age: 55 years), and 4 adult females aged 52-74 years (average age: 60 years). The collected data was acquired in a Belgian sleep hospital using a digital 32-channel polygraph (Brainnet System of MEDATEC, Brussels, Belgium). The sample frequency was 200 Hz. The visual sleep stage was scored by an expert according to the AASM criteria.

EOG signal Processing

Pre-processing: The EOG data was segmented into 5-second epochs. The entire EOG vector was processed utilising a zero-phase band pass filter with a Cascaded Integrator-Comb (CIC) filter of order six, for the following different frequency bands: delta (0.5–2 Hz), delta1 (2–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), sigma (12–16 Hz), beta1 (16–20 Hz ), and beta2 (20–30 Hz) (Figure 1). Since an EOG signal might be affected by EEG, EMG, and ECG artifacts, launching a suitable algorithm in order to remove the artifacts and noises is necessary. A cascade of three adaptive filters based on a least mean square algorithm was employed, and by this means EEG, EMG, and ECG artifacts were duly eliminated (more information in [15,16]).

Feature extraction: Several features were extracted from the EOG signal in the time and frequency domain, such as variance, Maximal Peak Amplitude Value (MPAV), Minimum Peak Amplitude Value (MPAV), total power, energy entropy, Shannon entropy, and cross-correlation. In order to select the best feature that classified variations in sleep stages and wakefulness, the Sequential Feature Selection method (SFS) was used.

Classification

In this study, the K-Nearest Neighbor was used for classification of sleep and wakefulness stages. The KNN is based on a nonparametric method for different pattern classification approach, which represents as one robust classifier. The KNN classifier works based on a comparison between a new sample (testing data) and baseline (training data). It attempts to find out the K-Nearest Neighbor within the baseline, and indicates a class which seems more normally in the nearest neighbor of K. The value of K might need to be diverse in order to detect the corresponding class between the training and testing data. In this paper, the value of K varies from 1 to 5. The Euclidean distance metric is utilized for calculating the distance between two points. The training and testing data was evaluated based on 10-fold cross-validation.

Smoothing rule

The smoothing rule is one of the common methods for increasing the accuracy of detecting the sleep stages. This rule is used as in the following example: three consecutive readings of N1, N2, and N2 were replaced as sequence N1, N1 and N1.

Classification of sleep disorders

In order to classify the PLMS, SAHS, and healthy control subjects, different PSG sleep stage parameters were first measured, such as Sleep Latency (SL), Sleep Efficiency (SE), number of times subject woke up (NW), Total Sleep Time (TST), Waking After Sleep Onset (WASO), Slow-Wave Sleep (SWS), and Rapid Eye Movement sleep (REM). During the first REM period, characteristics of the sleep stages (N1, N2, N3 and R) were based on the automatic detection system detailed in the previous section.

A decision tree analysis was used to classify the three groups of subjects based on the following rules:

- Rule (1) used the percentage of the sleep stage N1 parameter to separate SAHS patients from PLMS patients and healthy control subjects. If N1(%) was more than 7% and less than 9%, then a subject would be classified as an SHAS patient; if N1(%) was less than 4%, then a subject would be classified as a healthy control; if N1(%) was more than 10%, then a subject would be classified as a PLMS patient.

- Rule (2) used the percentage of the sleep stage N2 parameter to separate the SHAS from PLMS patients. If N2(%) was more than 80%, a patient would be classified as an SHAS patient; If N2(%) was less than 80% and more than 60%, a patient would be classified as a PLMS patient.

- Rule (3) utilised Slow-Wave Sleep Duration (SWS) in minutes and the percentage of sleep stage N3 to distinguish between PLMS and healthy control subjects. If SWS was more than 70 min and N3(%) more than 20%, a subject would be classified as a healthy control. All these rules were based on some percentage of sleep stage

---


---

Figure 1: Block diagram of the automatic sleep stage detection method using EOG signals.
and the time duration of sleep parameters, such as N1%, N2%, and SWS. The automatic classification algorithm is described in Figure 2.

A statistical analysis was conducted using post-hoc tests to ascertain whether there were significant differences between the three groups. Sensitivity and specificity tests, as well as Cohen’s Kappa, were conducted to evaluate the automatic classification algorithm for the three groups.

Results

Automatic sleep stage detection

In this pilot study, we utilised the EOG signal only for detection of the sleep stages of 30 subjects, comprising 10 healthy controls, 10 PLMS patients, and 10 SAHS patients. Several features were extracted from the EOG signal based on different frequency bands as mentioned in the previous section. The overall agreement, sensitivity, and specificity of the detection of sleep stages for healthy control subjects were 83.5%, 85%, and 88% respectively. The Cohen’s Kappa was 0.79. Table 1 shows the confusion matrix with sensitivity and specificity after applying the smoothing rule to one healthy subject as an example. The results show that the best detection was in wakefulness, and in sleep stage N3 (by 91%). The detection of sleep stage N1 after utilizing the smoothing rule was significantly improved. The overall agreement, sensitivity, and specificity for detection of the sleep stages of PLMS patients were 80%, 82%, and 86%, respectively. The Cohen’s Kappa was 0.71, which was lower than the Cohen’s Kappa of the healthy controls. The reason for this is that the normal distribution of sleep stages with healthy controls was much more consistent than that of the PLMS patients. Table 2 shows the confusion matrix, sensitivity, and specificity of the sleep stages of a PLMS patient. It is obvious that the total number of sleep stage N2s was higher than other sleep stages, which increased the detection of stages N1 and R. On the other hand the overall agreement, sensitivity, and specificity for detection of sleep stages with SAHS patients were 78%, 77%, and 80%, respectively, while the Cohen’s Kappa represented was lower than in the other two groups by 0.67. Table 3 shows the confusion matrix, sensitivity, and specificity of sleep stages of an SAHS patient. It is clear that the lower sensitivity was in the wakefulness stages, with an improvement in the detection of sleep stage N1.

Figures 3, 4, and 5 show the hypnograms of visual sleep stage scoring vs. automatic scoring for a healthy control subject, a PLMS patient, and an SAHS patient, respectively, as opposed to with automatic sleep stage detection. It can be observed that due to some sleep stages was scored as sleep stage N2 or N3 which made the view of hypnogram included some non-corrected classification. However, all of the visual scores were consistent with 84% for a healthy subject, 86% for a PLMS patient, and 79% for SAHS patients.

Automatic classification of the PLMS, SAHS, and healthy control subjects

Figure 2 shows the automatic classification algorithm used to classify the patients with PLMS, the patients with SAHS, and the healthy control subjects. The significant sleep parameters, such as N1(%), N2(%), and SWSD were used to identify the three groups on the basis of the thresholds as described in the previous section. Table 4 shows the post-hoc test analysis for the three groups of participants. There were significant differences between the PLMS patients and the healthy control participants, particularly in sleep stages N1, N2, and SWS duration. Furthermore, there were significant differences between the SAHS patients and the healthy control participants in the following sleep parameters: SWS duration, REM duration, and sleep stages N2, N3, and R. The participants with PLMS differed from the SAHS patients within some sleep parameters, such as in sleep stage N2. Figure 4 shows the bar plots of the three groups. The SL, WASO, and NW were all significant sleep stage parameters for the PLMS patients. Sleep stage N2 was a significant sleep parameter for the SAHS patients. The sensitivity and specificity of identification in the PLMS patients, SAHS patients, and healthy controls were 90% and 95%, respectively (Table 5). The level of accuracy and Cohen’s kappa were 90% and 0.85, respectively.

Discussion

In this pilot study, we aimed to use EOG signals for automatic sleep


Figure 2: The automatic classification algorithm for PLMS patients, SAHS patients, and healthy participants based on decision tree analysis. Where N1 and N2 are the sleep stages and SWSD is slow-wave sleep duration in min.

Figure 3: The hypnogram of visual sleep stage scoring vs. automatic scoring for a healthy control subject.

Figure 4: The hypnogram of visual sleep stage scoring vs. automatic scoring for a PLMS patient.

Figure 5: The hypnogram of visual sleep stage scoring vs. automatic scoring for an SAHS patient.
Table 1: The confusion matrix of a healthy control subject.

<table>
<thead>
<tr>
<th>Visual detection</th>
<th>Automatic detection</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1 (%)</td>
<td>N2 (%)</td>
</tr>
<tr>
<td>W</td>
<td>575</td>
</tr>
<tr>
<td>N1</td>
<td>22</td>
</tr>
<tr>
<td>N2</td>
<td>15</td>
</tr>
<tr>
<td>N3</td>
<td>3</td>
</tr>
<tr>
<td>R</td>
<td>11</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>90</td>
</tr>
</tbody>
</table>

Table 2: The confusion matrix of a PLMS patient.

<table>
<thead>
<tr>
<th>Visual detection</th>
<th>Automatic detection</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1 (%)</td>
<td>N2 (%)</td>
</tr>
<tr>
<td>W</td>
<td>524</td>
</tr>
<tr>
<td>N1</td>
<td>50</td>
</tr>
<tr>
<td>N2</td>
<td>25</td>
</tr>
<tr>
<td>N3</td>
<td>10</td>
</tr>
<tr>
<td>R</td>
<td>186</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>88</td>
</tr>
</tbody>
</table>

Table 3: The confusion matrix of an SAHS patient.

<table>
<thead>
<tr>
<th>Visual detection</th>
<th>Automatic detection</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1 (%)</td>
<td>N2 (%)</td>
</tr>
<tr>
<td>W</td>
<td>547</td>
</tr>
<tr>
<td>N1</td>
<td>29</td>
</tr>
<tr>
<td>N2</td>
<td>38</td>
</tr>
<tr>
<td>N3</td>
<td>3</td>
</tr>
<tr>
<td>R</td>
<td>11</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>85</td>
</tr>
</tbody>
</table>

Table 4: The post-hoc t-tests of the differences in the sleep parameters between the three groups.

Table 5: Confusion matrix, sensitivity, and specificity of the three groups.
An automatic method was previously developed for detection of SWS based on two EOG channels [22]. This study employed the amplitude criterion for detecting SWS, and beta power [18-23] was utilised to reduce the artefact. The result shows inter-rater reliability between the visual and the developed automatic method of 73%, with a Cohen’s kappa value of 0.70. The sensitivity and specificity were 79% and 96%, respectively. Another study employed two-channel electrooculography for automatic sleep stage classification particular to the left mastoid (M1) [17]. The synchronous Electroencephalographic (EEG) activity during SWS and S2 were detected by calculating peak-to-peak and cross-correlation amplitude differences in the 0.5 to 6 Hz range, and between the two EOG channels. The result indicated epoch by epoch agreement between the visual and the developed automatic method of 72%, with a Cohen’s kappa value of 0.63.

The second aim of this study was to utilise the sleep stages that were detected based on an automated system for the purpose of classifying PLMS patient, SAHS patients, and healthy control subjects. We provide significant evidence that supports the use of the PSG sleep stage features—such as sleep stage N1(%), N2(%), and SWSD—for an automatic classification of PLMS patients, SAHS patients, and healthy control subjects. The percentage of sleep stage N1 was the most significant feature distinguishing patients with SAHS from healthy control subjects. The study found that sleep stage N(%) was between 7% and 9% for 7 SAHS patients. Conversely, 8 healthy subjects had a percentage of sleep stage N1 for less than 5% of total sleep duration. However, some patients with PLMS show higher percentages in sleep stage N1, which led the mean average of sleep stage N1(%) to be higher than in the other two groups. Figure 6 presents evidence that the mean average of sleep stage N1(%) for less than 5% of total sleep duration.

The percentage of sleep stage N2 was used to distinguish between SAHS and PLMS patients. This study found that most of the SAHS patients had a higher percentage of sleep stage N2 (above 80%) than PLMS patients. In Figure 2, we show the threshold that was used to distinguish between these two groups of patients. The SWSD and the percentage of sleep stage N3 were used to separate the PLMS patients from the healthy control subjects. We found that most of the patients with PLMS had less SWSD compared to healthy subjects. The longest SWSD of the healthy controls was above 70 min. The reason for using the percentage of sleep stage N3 is because some patients with PLMS had similar SWSD to the healthy control group. Figure 2 shows the threshold that was used to separate these two groups. The overall accuracy was 90%, and the Cohen’s kappa was 0.85.

**Conclusion**

In conclusion, this paper aimed to develop an automatic method for detection of the sleep stages based on EOG signals, and then utilised these sleep stages for classification of PLMS patients, SAHS patients, and healthy control subjects. There was a significant advantage which supports utilising automatic sleep stage detection based on only EOG signals on ambulatory sleep recordings. The sensitivity of identifying PLMS, SAHS and healthy control participants was 90%, 90%, and 90%, respectively. This study suggests that using an automatic classification system in screening processes is more effective and efficient compared to some standards such as PSQI.

**Acknowledgements**

The authors gratefully acknowledge the financial support of the Ministry of Higher Education of the Kingdom of Saudi Arabia.

**References**


15. The DREAMS Subjects Database.


