3-30-2015

Detecting Slow Wave Sleep Using a Single EEG Signal Channel

Bo-Lin Su
National Sun Yat-Sen University

Yuxi Luo
National Sun Yat-Sen University

Chih-Yuan Hong
National Sun Yat-Sen University

Mark L. Nagurka
Marquette University, mark.nagurka@marquette.edu

Chen-Wen Yen
National Sun Yat-Sen University

Accepted version. *Journal of Neuroscience Methods*, Vol 243 (March 30, 2015): 47-52. DOI.

Detecting Slow Wave Sleep Using a Single EEG Signal Channel

Bo-Lin Su
Department of Mechanical and Electro-Mechanical Engineering, National Sun Yat-Sen University, Kaohsiung, Taiwan
Yuxi Luo
School of Engineering, Sun Yat-Sen University, Guangzhou 510006, China
Chih-Yuan Hong
Department of Mechanical and Electro-Mechanical Engineering, National Sun Yat-Sen University, Kaohsiung, Taiwan
Mark L. Nagurka
Department of Mechanical Engineering, Marquette University, Milwaukee, WI
Chen-Wen Yen
Department of Mechanical and Electro-Mechanical Engineering, National Sun Yat-Sen University, Department of Physical Therapy, Kaohsiung Medical University, Kaohsiung, Taiwan

Journal of Neuroscience Methods, Vol 243 (March 30, 2015): pg. 47-52. DOI.

Abstract

Background: In addition to the cost and complexity of processing multiple signal channels, manual sleep staging is also tedious, time consuming, and error-prone. The aim of this paper is to propose an automatic slow wave sleep (SWS) detection method that uses only one channel of the electroencephalography (EEG) signal.
New Method: The proposed approach distinguishes itself from previous automatic sleep staging methods by using three specially designed feature groups. The first feature group characterizes the waveform pattern of the EEG signal. The remaining two feature groups are developed to resolve the difficulties caused by interpersonal EEG signal differences.
Results and comparison with existing methods: The proposed approach was tested with 1,003 subjects, and the SWS detection results show kappa coefficient at 0.66, an accuracy level of 0.973, a sensitivity score of 0.644 and a positive predictive value of 0.709. By excluding sleep apnea patients and persons whose age is older than 55, the SWS detection results improved to kappa coefficient, 0.76; accuracy, 0.963; sensitivity, 0.758; and positive predictive value, 0.812.
Conclusions: With newly developed signal features, this study proposed and tested a single-channel EEG-based SWS detection method. The effectiveness of the proposed approach was demonstrated by applying it to detect the SWS of 1003 subjects. Our test results show that a low SWS ratio and sleep apnea can degrade the performance of SWS detection. The results also show that a large and accurately staged sleep dataset is of great importance when developing automatic sleep staging methods.
Keywords: Slow wave sleep, Electroencephalography, Automatic sleep staging, Sleep apnea

1. Introduction

Human sleep can be divided into rapid-eye movement (REM) sleep and nonrapid eye movement (NREM) sleep. NREM sleep can be further classified into Stage 1 (S1), Stage 2 (S2) and Slow Wave Sleep (SWS). Among these sleep stages, SWS has been considered to be the most restorative sleep stage (Van Cauter et al., 2008). However, as sleep quality declines with aging, the total amount of SWS decreases drastically (Van Cauter et al., 2000). In addition, abnormal SWS has also been found to be correlated with a variety of clinical problems

including acute-phase immune system response (Majde and Krueger, 2005), diabetes risk (Tasali et al., 2008), memory consolidation (Diekelmann and Born, 2010), psychiatric disorders (Kyung Lee and Douglass, 2010), and hypertension (Fung et al., 2011).

In-depth understanding of these SWS-related problems can be gained by performing long-term sleep architecture monitoring on a large population. Conventionally, sleep stages are scored using the electroencephalogram (EEG), electromyogram (EMG) and electrooculogram (EOG) signals obtained by a polysomnography (PSG) study. Thus, a large scale sleep architecture monitoring study would require an intensive amount of work. In addition to the cost and complexity of measuring and processing multiple signal channels, manual sleep staging is also tedious, time consuming, and error-prone. In response to these challenges, many automatic sleep staging methods using statistical learning or artificial intelligence techniques have been developed (Park et al., 2000, Agarwal and Gotman, 2001, Caffarel et al., 2006, Porée et al., 2006, Tagluk et al., 2010, Liang et al., 2012 and Pan et al., 2012). The complexity of the sleep staging problems can be furthered alleviated by extracting features from fewer signal channels (Agarwal et al., 2005, Berthomier et al., 2007, Virkkala et al., 2007a, Virkkala et al., 2007b, Malinowska et al., 2009, Güneşa et al., 2010, Levendowski et al., 2012 and Stepnowsky et al., 2013).

In addition to cost and complexity, the other challenge for sleep staging is accounting for differences between individual subjects. In general, human physiological signals are qualitatively similar but not quantitatively identical. For example, SWS is characterized by its high-voltage slow-oscillation EEG waveform pattern. However, the actual magnitude and frequency band of the SWS EEG signal depend on the individual being measured.

The aim of this study is to introduce an automatic single EEG channel SWS detection method. This study has several features which distinguish it from previous work. First, the proposed method is less sensitive to the uncertainty caused by individual differences. Second, we address the influence of both the SWS ratio and AHI value, whose impacts on the efficacy of sleep staging have rarely been investigated.

Journal of Neuroscience Methods, Vol 243 (March 30, 2015): pg. 47-52. DOI.

2. Materials and methods

2.1. Study subjects and polysomnography

We recruited 1003 patients from China Medical University Hospital Sleep Centers. Table 1 shows the means and standard deviations of the age, BMI, AHI (apnea-hypopnea index), sleep efficiency and SWS percentage of the test subjects. Note that AHI is defined as the number of apneas and hypopneas per hour of sleep and is often used as a metric to characterize the severity of sleep apnea. PSG data were recorded using a computerized polysomnographic system (Alice 4; Healthdyne, Atlanta, GA, USA). The proposed features were extracted from the C3-M2 EEG channel. The sampling rate for the EEG recording was 100 Hz. A high pass filter of 0.3 Hz and a low pass filter of 35 Hz were used to preprocess the EEG signal. The data were reviewed by an experienced doctor and scored by a sleep technician certified by the Taiwan Sleep Medicine Society. Every 30-second epoch was scored according to the criteria of Rechtschaffen and Kales (Rechtschaffen and Kales, 1968).

Table 1. Clinical characteristics of the test subjects.

Clinical characteristics	Learning set
Number of persons	1003
Age, yr	44.2 ± 14.9
BMI, kg/m^2	26.6 ± 4.6
AHI event/hour	31.7 ± 28.3
Sleep efficiency %	80.2 ± 15.2
SWS %	4.4 ± 7.3

2.2. A neural network based SWS detector

This study formulates the SWS detection problem as a binary classification problem whose goal is to distinguish SWS and non-SWS (NSWS) on an epoch-by-epoch basis. In particular, our binary classifier employs a supervised learning strategy that uses the manually scored sleep staging results as the desired outputs for a neural network model. Specifically, with the 1003 subjects as the training set, this study uses the conventional back-propagation method to determine the optimal connection weights of a multilayer

perceptron (MLP) neural network model by minimizing the mean square output error of the MLP.

2.3. Features for sleep staging

The type and number of EEG features are critical to the efficacy of the SWS detection method. This work proposes nine EEG features which can be divided into three groups, each of which consists of three features. The first group is used to characterize the EEG waveform pattern. Developed from the first feature group, the second and third feature groups are used to overcome the problem of signal differences among individuals. For the sake of convenience, *x, y, and z* will represent the features of the first, second, and third feature groups, respectively. The following three subsections introduce these three feature groups in sequence.

2.3.1. Zero-crossing point features

The first feature group is derived from the zero-crossing points. To generate these features, an epoch is first divided into thirty one-second intervals. For each interval, after subtracting the interval mean from the signal, the zero-crossing points (ZCPs) are located by finding the points where the signal changes sign. With these ZCPs, the time history of a signal is divided into a number of segments, hereafter referred to as "ZCP segments." Since the speed of the signal is closely correlated with the lengths of the ZCP segments, the mean and standard deviation of the ZCP segment lengths are chosen as the first and second features of the first feature group.

For a given epoch, the third feature can be written as:

equation(1)

$$x_3 = \sum_{i=1}^{I} (e_{i+1} - e_i) \int_{e_i}^{e_{i+1}} |s(t)|\, dt$$

where e_i is the ith ZCP, I is the number of ZCP segments and $s(t)$ denotes the EEG signal. Note that the integral in Eq. (1) is the absolute value of the area under the response curve of EEG signal $s(t)$ and can be numerically determined from the discrete sampled EEG

signal. Apparently, increasing the amplitude of $s(t)$ enlarges this area, and decreasing the speed of $s(t)$ results in a larger $e_{i+1} - e_i$ difference. Together, these two properties characterize the large-voltage low-oscillation EEG signal pattern of SWS.

2.3.2. Features of relative probability density

As addressed in the previous section, one of the fundamental difficulties of physiological signal processing is the variability caused by the differences between individual subjects. The second and third feature groups are proposed to resolve such a difficulty. In particular, for a given first group feature x, the probability density function (PDF) $p(x)$ associated with the training set can be represented as:

equation(2)

$$p(x)=Cp_S(x)+(1-C)p_N(x)$$

where p_S and p_N are the training set PDFs of x for the SWS and NSWS, respectively. In addition, C and $1 - C$ are the ratios of the training set SWS and NSWS, respectively.

According to Bayes' theorem, the optimal decision rule for a classification problem depends on the posterior probabilities of the patterns to be classified (Theodoridis and Koutroumbas, 2008). For the SWS detection problem, the two posterior probabilities of interest are:

equation(3)

$$p(\mathrm{SWS}|x) = \frac{p(x|\mathrm{SWS})p(\mathrm{SWS})}{p(x)} = \frac{p_S(x)C}{p(x)}$$

equation(4)

$$p(\mathrm{NSWS}|x) = \frac{p(x|\mathrm{NSWS})p(\mathrm{NSWS})}{p(x)} = \frac{p_N(x)(1-C)}{p(x)}$$

Apparently, these two posterior probabilities depend on the prior probabilities p(SWS) and p(NSWS) which are simply C and $1 - C$. Considering the fact that these prior probabilities vary from person to person and from night to night, this study introduces the second feature group to account for such a variability.

In this study, MLPs are trained to estimate the posterior probabilities of the training set which consists of multiple PSG records. If C_i represents the ratio of SWS and T_i denotes the length of the total sleep time associated with the *i*th PSG record of the training set, then the MLP is trained to produce outputs f_1 and f_2 to approximate the following two posterior probabilities (Saerens et al., 2002):

equation(5)

$$f_1(x) \approx \frac{p_S(x)C}{p(x)}$$

equation(6)

$$f_2(x) \approx \frac{p_N(x)(1-C)}{p(x)}$$

where the ratio of the SWS of the training set is:

equation(7)

$$C \approx \frac{\sum_{i=1}^{N} C_i T_i}{\sum_{i=1}^{N} T_i}$$

with N as the number of PSG records of the training set. To achieve Bayes optimality for the *i*th PSG record, the outputs the MLP should be:

equation(8)

$$f_1 \approx \frac{p_S(x)C_i}{p_i(x)}$$

equation(9)

$$f_2 \approx \frac{p_N(x)(1-C_i)}{p_i(x)}$$

where p_i is the PDF of the *i*th PSG record and can be written as

equation(10)

$$p_i(x) = C_i p_S(x) + (1 - C_i) p_N(x)$$

Clearly, the Bayes optimality for the entire training set is achieved at a cost: the classification accuracy of individual PSG records is compromised.

To develop features that are adaptive to the variation of SWS ratio, we assume the PDF of a PSG record that awaits staging to be:

equation(11)

$$q(x)=Dp_S(x)+(1-D)p_N(x)$$

Substituting Eqs. (5) and (6) into Eq. (11) to eliminate p_S and p_N yields:

equation(12)

$$\frac{q(x)}{p(x)} \approx \frac{D}{C} f_1 + \frac{1-D}{1-C} f_2$$

Since C, f_1 and f_2 can all be derived from the training set, Eq. (12) can be used to estimate D once $q(x)/p(x)$ is known. Because accurate computation of the probability density is difficult to achieve, we employ an alternative approach. Specifically, if the value of a first group feature x is known to be a, then the corresponding second group feature y is computed from:

equation(13)

$$y = \frac{\int_{a-\Delta x}^{a+\Delta x} q(x)dx}{\int_{a-\Delta x}^{a+\Delta x} p(x)dx}$$

To calculate y, we first determine Δx from the following equality:

equation(14)

$$\int_{a-\Delta x}^{a+\Delta x} p(x)dx = \varepsilon,$$

for a small ε. The upper (lower) limit of the above definite integral should be replaced by the maximum (minimum) value of x if $a + \Delta x(a - \Delta x)$ is larger (smaller) than the maximum (minimum) value of x. In this study, the value of ε is chosen as 0.05. With Δx

Journal of Neuroscience Methods, Vol 243 (March 30, 2015): pg. 47-52. DOI.

found, the numerator of Eq. (13) and, thus, the value of the second group feature *y*, can be readily determined. For convenience, feature *y* will be referred to as the relative probability density of *x* hereafter in this paper.

2.3.3. Features of percentile rank

In the previous subsection, we assumed that the PDFs of SWS and NSWS were completely known. To further address the issue of individual differences, this subsection assumes that the means of p_S and p_N can vary from person to person. Specifically, for the *i*th PSG record, the PDFs of the SWS and NSWS are assumed to have the following form:

equation(15)

$$p_S^i(x) = p_S(x - m_i)$$

equation(16)

$$p_N^i(x) = p_N(x - m_i)$$

As a result, based on Eq. (10), the PDF for feature *x* of the *i*th PSG record can be represented as:

equation(17)

$$\mathrm{pi}(\mathrm{x}) = \mathrm{p_i}(\mathrm{x}-\mathrm{m_i}) = \mathrm{C_i p_S}(\mathrm{x}-\mathrm{m_i}) + (1-\mathrm{C_i})\mathrm{p_N}(\mathrm{x}-\mathrm{m_i})$$

With Eqs. (10) and (17), the following can be readily shown:

equation(18)

$$\int_{-\infty}^{\alpha} p^i(x)dx = \int_{-\infty}^{\alpha - m_i} p_i(x)dx$$

To account for the unknown m_i, for every *x* of the first feature group, the corresponding third group feature *z* is chosen as the percentile rank of *x* within its PSG record. To demonstrate the rationale behind the third feature group, we first assume that when $x = a$, the corresponding percentile rank is z_a. Based on the definition of cumulative probability, we have:

equation(19)

$$\int_{-\infty}^{\alpha} p^i(x)dx = \frac{z_\alpha}{100}$$

With Eq. (18), Eq. (19) can be rewritten as

equation(20)

$$\int_{-\infty}^{\alpha - m_i} p_i(x)dx = \frac{z_\alpha}{100}$$

Hence, with z_α as a feature, Eq. (20) can be used to estimate m_i.

In summary, the proposed feature set consists of three groups. The first feature group is used to characterize the waveform pattern of the SWS EEG signal. By assuming $p_S(x)$ and $p_N(x)$ to be invariant, the second feature group is used to account for the uncertainty associated with SWS ratio C_i which varies from person to person and from night to night. In contrast, by assuming C_i to be known, we develop the third feature group to help us estimate m_i which characterizes the variation of the means of the first group features, which is dependent on the individual. The intent of using these features simultaneously as the neural network inputs is to effectively improve the SWS detection efficacy of the proposed approach.

3. Results

The first goal of this experimental study is to apply the proposed approach to automatically detect the SWS of the 1003 subject dataset. Derived from a single EEG channel, the first test used only the first feature group $\boldsymbol{x}$, whereas the second test employed feature groups $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$. Note that $\boldsymbol{x} = [x_1 \times x_2\ x_3]^T$ where x_1 and x_2 are the mean and standard deviation of the ZCP segment lengths of the epoch to be scored, and x_3 is defined by equation (3). Similar to $\boldsymbol{x}$, the second feature group $\boldsymbol{y}$ and the third feature group $\boldsymbol{z}$ are also 3 by 1 vectors with y_i as the relative probability density of x_i (Eq. (13)), and z_i is the percentile rank of x_i (Eq. (19)).

In both tests, a single-hidden layer MLP was employed to solve the binary classification problem of SWS detection. The numbers of hidden neurons for these two MLPs were determined based on two

criteria. First, to perform a fair comparison, the number of connection weights of the MLPs employed in the two tests should be approximately the same. Second, the SWS detection performance should be optimized with the minimum number of hidden units. Based on these two rules, via a trial-and-error process, the number of MLP hidden units were chosen as 40 and 18 for the first and second tests, respectively. The resulting kappa coefficients, accuracy sensitivity and positive predictive values (PPV) are shown in Table 2. These results clearly show that the addition of the second and third feature groups improves the efficacy of the proposed automatic SWS detection method.

Table 2. Results of the comparative tests for two feature sets.

Features	Kappa	Accuracy	Sensitivity	PPV
x	0.64	0.972	0.610	0.707
x,y,z	0.66	0.973	0.644	0.709

Note that since the number of training examples is much larger than the number of MLP parameters (775,471 epochs versus approximately 200 connection weights), the overfitting problem is unlikely to occur (Amari et al., 1995). This belief has been confirmed by splitting the dataset into training subset (502 subjects) and test subset (501 subjects). The training subset was used to generate optimal connection weights for the MLP whereas the test subset was used to validate the performance of the MLP designed by the training subset. As shown by the results summarized in Table 3, the training and test subsets yield comparable performances which clearly demonstrate the generalization capability of the trained MLP.

Table 3. Results of the training and test subsets.

Dataset	Kappa	Accuracy	Sensitivity	PPV
Training	0.66	0.972	0.636	0.715
Test	0.66	0.973	0.653	0.694

To further evaluate the performance of proposed approach, we stratify the results obtained by the proposed approach in Table 2 to different subgroups. In particular, to evaluate the efficacy of the proposed approach for middle-aged or younger healthy persons, the inclusion criteria for the first subgroup are AHI < 5 and age < 55. Note that an AHI of less than 5 is considered normal. To assess the

performance of the proposed approach for healthy older persons, the second subgroup requires AHI < 5 and age ≥55. As given in Table 4, the SWS ratio of the second (older) subgroup is much smaller than that of the first (younger) subgroup. With a inclusion criterion of AHI ≥5, the third subgroup consists of all the sleep apnea patients of our dataset. As shown by Table 4, the comparative results of the first and second subgroups seem to indicate that the reduction of the SWS ratio can deteriorate the performance of the proposed approach. Similarly, from the results of Table 4, we also find that the proposed approach performs considerably better for the healthy persons (first and second subgroups) than the sleep apnea patients (third subgroup). Hence, the next goal of this study is to investigate the influences of SWS ratio and AHI value on the performance of the proposed approach.

Table 4. Stratify the SWS detection results to different subgroups.

Subgroups	**AHI < 5 Age < 55**	**AHI < 5 Age ≥ 55**	**AHI ≥ 5**
Subjects	132	17	853
AHI	2.2 ± 1.5	1.9 ± 1.4	36.9 ± 27.6
SWS %	9.1 ± 10.0	2.1 ± 6.0	3.7 ± 6.6
Kappa	0.72	0.68	0.64
Accuracy	0.955	0.988	0.975
Sensitivity	0.753	0.582	0.630
PPV	0.745	0.829	0.673

To try to understand the difficulty caused by a small SWS ratio, we selected five different SWS% subsets from our training set. With a 5% SWS ratio interval, these subgroups were arranged such that the differences of their AHI means are statistically insignificant at the 5% level. The results in Table 5 show that, in agreement with the results of Table 4, lower SWS ratios lead to poorer SWS detection results. This phenomenon can be explained by the training set imbalance problem (Sun et al., 2009, He and Garcia, 2009 and Khoshgoftaar et al., 2010). For a binary classification problem, training set imbalance occurs when the number of examples of one class (the majority class) is considerably larger than the number of examples of the other class (the minority class). As the ratios of these two classes of examples become more extreme, the classifier places more emphasis on the majority class in order to minimize the overall error. Unfortunately, minimizing the overall error is often accomplished at the expense of the classification accuracy of the minority class. As a result, the

classifier can easily become one-sided and thus become ineffective in detecting minority class examples. In the present study, SWS is apparently the minority class. Consequently, as the SWS ratio becomes smaller, accurate SWS detection becomes more difficult to achieve.

Table 5. The influence of the SWS ratio.

SWS ratio%	0-5	5–10	10–15	15–20	>20
Subjects	32	34	29	28	22
Mean of AHI	5.090	5.041	5.083	5.021	5.131
Kappa	0.39	0.56	0.69	0.73	0.84
Accuracy	0.965	0.943	0.937	0.929	0.935
Sensitivity	0.494	0.501	0.605	0.619	0.828
PPV	0.316	0.622	0.810	0.876	0.919

Based on the PSG recordings of three healthy subjects and three OSA patients, it has been shown that automatic sleep staging methods can be less accurate for the sleep apnea patients than for healthy subjects (Park et al., 2000). However, to the best of our knowledge, this potential problem has never been studied in-depth. Therefore, this study examines the influence of the AHI on the performance of the proposed approach. Toward this goal, based on the degree of severity of sleep apnea, Table 6 presents the training results of four selected subsets of patients (healthy AHI < 5, mild $5 \leq$ AHI < 15, moderate $15 \leq$ AHI < 30, severe AHI ≤ 30). To eliminate the influence of the SWS ratio, these subsets were chosen such that the differences between their SWS ratio means are statistically insignificant at the level of 0.05.

Table 6. The influence of the AHI value.

AHI	<5	5–15	15–30	>30
Subjects	61	96	72	82
SWS%	5.02%	5.05%	5.03	5.02%
Kappa	0.56	0.50	0.53	0.52
Accuracy	0.955	0.950	0.960	0.960
Sensitivity	0.647	0.549	0.484	0.464
PPV	0.534	0.506	0.626	0.633

As demonstrated by Table 6, the proposed approach is less effective in dealing with sleep apnea patients. This performance degradation problem also appears in human scoring. Specifically,

Journal of Neuroscience Methods, Vol 243 (March 30, 2015): pg. 47-52. DOI.

several inter-scorer reliability studies have found that human sleep staging results for obstructive sleep apnea (OSA) patients have a comparatively lower degree of agreement than do those of healthy subjects (Norman et al., 2000, Collop, 2002, Danker-Hopfe et al., 2004 and Basner et al., 2008). It is possible that the signal infidelity caused by the increased movement artifacts, arousals, and other apnea symptoms result in the performance degradation of the proposed approach for apnea patients. It is also likely that such a performance drop comes from the imperfect training set that contains incorrect human staging decisions. Unfortunately, the extent of their influence and interactions is unclear. Therefore, a possible future research project is to compare the degree of performance degradation of the proposed approach to the degree of increased inter-scorer disagreement in dealing with apnea patients.

With an imperfect training set, building a "perfect" automatic sleep staging method does not seem to be a realistic expectation. Although PSG has been considered the gold standard in the diagnosis of sleep disorders, for the development of automatic sleep staging methods, we also need a sufficiently large and complete PSG dataset that possesses highly accurate, if not perfect, sleep staging results. Therefore, it is believed that the results reported in this study have been hampered by the imperfection of the training set employed in this study. For the very same reason, we also expect better results from the proposed approach if the quality of the training set can be effectively improved.

The results reported in Table 1, Table 2, Table 3, Table 4, Table 5 and Table 6 are all based on the MLP classifier trained by the entire dataset of 1003 subjects. Considering the impacts of the severity of sleep apnea and ratio of SWS, it seems possible to improve the SWS detection results by tailoring the proposed approach to subgroup of persons that share high degree of similarity in terms of SWS ratio and AHI value. To demonstrate this potential, a MLP was exclusively designed by using dataset of the first subgroup studied in Table 4. Compared to the results of Table 4 which were obtained by the MLP trained by the entire dataset of 1003 subjects, this specially designed MLP improves kappa coefficient (from 0.72 to 0.76), accuracy (from 0.955 to 0.963), sensitivity (from 0.753 to 0.758) and positive predictive value (from 0.745 to 0.812).

In this study, the proposed approach used features extracted from the C3-M2 channel. Identical experimental studies have also been conducted by using C4-M1 channel. Although not reported here, the results obtained by these two EEG channels are quite comparable. However, considering spatially dependent EEG potentials from the various regions of the scalp, a possible future work is to try to use features extracted from other EEG channels in order to optimize the performance of the proposed approach. For example, in addition to the conventional EEG electrode locations (such as the 10–20 system of electrode placement), one can also try to use signals measured by forehand leads.

4. Conclusion

Based on newly developed signal features, this study proposed and tested a single-channel EEG-based SWS detection method. In addition to introducing a group of features to characterize the high-amplitude low-frequency SWS EEG waveform patterns, two additional groups of features were also developed to resolve the difficulties caused by interpersonal differences of the EEG signals. The effectiveness of the proposed approach was demonstrated by applying it to detect the SWS of 1003 subjects. In addition, this study investigated the influence of the ratio of slow wave sleep and the severity of sleep apnea on the efficacy of the proposed method. Our test results show that a low SWS ratio and sleep apnea can degrade the performance of SWS detection. The results of this study also show that a sufficiently large and accurately staged PSG dataset is of great importance when developing high performance automatic sleep staging methods.

References

Agarwal and Gotman, 2001. R. Agarwal, J. Gotman. Computer-assisted sleep staging. *IEEE Trans Biomed,* 48 (12) (2001), pp. 1412–1423

Agarwal et al., 2005. R. Agarwal, T. Takeuchi, S. Laroche, J. Gotman. Detection of rapid-eye movements in sleep studies. *IEEE Trans Biomed,* 52 (8) (2005), pp. 1390–1396

Amari et al., 1995. S. Amari, N. Murata, K.R. Müller, M. Finke, H.H. Yang. Statistical theory of overtraining-Is cross-validation asymptotically effective? *Adv Neural Inf Process Syst,* 8 (1995), pp. 176–182

Basner et al., 2008. M. Basner, B. Griefhahn, T. Penzel. Inter-rater agreement in sleep stage classification between centers with different background. *Somnologie,* 12 (1) (2008), pp. 75–84

Berthomier et al., 2007. C. Berthomier, X. Drouot, M. Herman-Stoïca, P. Berthomier, J. Prado, D. Bokar-Thire, *et al.* Automatic analysis of single-channel sleep EEG: validation in healthy individuals. *Sleep,* 30 (11) (2007), pp. 1587–1595

Caffarel et al., 2006. J. Caffarel, G.J. Gibson, J.P. Harrison, C.J. Griffiths, M.J. Drinnan. Comparison of manual sleep staging with automated neural network-based analysis in clinical practice. *Med Biol Eng Comput,* 44 (1-2) (2006), pp. 105–110

Collop, 2002. N.A. Collop. Scoring variability between polysomnography technologists in different sleep laboratories. *Sleep Med,* 3 (1) (2002), pp. 43–47

Danker-Hopfe et al., 2004. H. Danker-Hopfe, D. Kunz, G. Gruber, G. Klösch, J.L. Lorenzo, S.L. Himanen, *et al.* Interrater reliability between scorers from eight European sleep laboratories in subjects with different sleep disorders. *J Sleep Res,* 13 (1) (2004), pp. 63–69

Diekelmann and Born, 2010. S. Diekelmann, J. Born. The memory function of sleep. *Nat Rev Neurosci,* 11 (2) (2010), pp. 114–126

Fung et al., 2011. M.M. Fung, K. Peters, S. Redline, M.G. Ziegler, S. Ancoli-Israel, E. Barrett-Connor, *et al.* Decreased slow wave sleep increases risk of developing hypertension in elderly men. *Hypertension,* 58 (4) (2011), pp. 596–603

Güneşa et al., 2010. S. Güneşa, K. Polata, S. Yosunkayab. Efficient sleep stage recognition system based on EEG signal using k-means clustering based feature weighting. *Expert Syst,* 37 (12) (2010), pp. 7922–7928

He and Garcia, 2009. H. He, E.A. Garcia. Learning from imbalance data. *IEEE Trans Knowl Data Eng,* 21 (9) (2009), pp. 1263–1284

Khoshgoftaar et al., 2010. T.M. Khoshgoftaar, J. Van Hulse, A. Napolitano. Supervised neural network modeling: an empirical investigation into learning from imbalanced data with labeling errors. *IEEE Trans Neural Netw*, 21 (5) (2010), pp. 813–830

Kyung Lee and Douglass, 2010. E. Kyung Lee, A.B. Douglass. Sleep in psychiatric disorders: where are we now? *Can J Psychiatry,* 55 (7) (2010), pp. 403–412

Levendowski et al., 2012. D.J. Levendowski, D. Popovic, C. Berka, P.R. Westbrook. Retrospective cross-validation of automated sleep staging using electroocular recording in patients with and without sleep disordered breathing. *Int Arch Med,* 5 (1) (2012), p. 21

Liang et al., 2012. S.F. Liang, C.E. Kuo, Y.H. Hu, Y.S. Cheng. A rule-based automatic sleep staging method. *J Neurosci Methods,* 205 (1) (2012), pp. 169–176

Majde and Krueger, 2005. J.A. Majde, J.M. Krueger. Links between the innate immune system and sleep. *J Allergy Clin Immunol,* 116 (6) (2005), pp. 1188–1198

Malinowska et al., 2009. U. Malinowska, H. Klekowicz, A. Wakarow, S. Niemcewicz, P.J. Durka. Fully parametric sleep staging compatible with the classical criteria. *Neuroinformatics,* 7 (4) (2009), pp. 245–253

Norman et al., 2000. R.G. Norman, I. Pal, C. Stewart, J.A. Walsleben, D.M. Rapoport. Interobserver agreement among sleep scorers from different centers in a large dataset. *Sleep,* 23 (7) (2000), pp. 901–908

Pan et al., 2012. S.T. Pan, C.E. Kuo, J.H. Zeng, S.F. Liang. A transition-constrained discrete hidden Markov model for automatic sleep staging. *Biomed Eng Online,* 11 (2012), p. 52

Park et al., 2000. H.J. Park, J.S. Oh, D.U. Jeong, K.S. Park. Automated sleep stage scoring using hybrid rule- and case-based reasoning. *Comput Biomed Res,* 33 (5) (2000), pp. 330–349

Porée et al., 2006. F. Porée, A. Kachenoura, H. Gauvrit, C. Morvan, G. Carrault, L. Senhadji. Blind source separation for ambulatory sleep recording. *IEEE Trans Inf Technol Biomed,* 10 (2) (2006), pp. 293–301

Rechtschaffen and Kales, 1968. A. Rechtschaffen, A. Kales. *A Manual of Standardized Terminology. Techniques and Scoring Systems for Sleep Stages of Human Subject.* National Institutes of Health, Washington DC (1968) [Pub. No. 204]

Saerens et al., 2002. M. Saerens, P. Latinne, C. Decaestecker. Any reasonable cost function can be used for a posteriori probability approximation. *IEEE Trans Neural Netw,* 3 (5) (2002), pp. 1204–1210

Stepnowsky et al., 2013. C. Stepnowsky, D. Levendowski, D. Popovic, I. Ayappa, D.M. Rapoport. Scoring accuracy of automated sleep staging from a bipolar electroocular recording compared to manual scoring by multiple raters. *Sleep Med,* 14 (11) (2013), pp. 1199–1207

Sun et al., 2009. Y. Sun, A.K.C. Wong, M.S. Kamel. Classification of imbalanced data: a review. *Int J Pattern Recognit Artif Intell,* 23 (4) (2009), pp. 687–719

Tagluk et al., 2010. M.E. Tagluk, N. Sezgin, M. Akin. Estimation of sleep stages by an artificial neural network employing EEG, EMG and EOG. *J Med Syst,* 34 (4) (2010), pp. 717–725

Tasali et al., 2008. E. Tasali, R. Leproult, D.A. Ehrmann, E. Van Cauter. Slow-wave sleep and the risk of type 2 diabetes in humans. *Proc Natl Acad Sci,* 105 (3) (2008), pp. 1044–1049

Theodoridis and Koutroumbas, 2008. S. Theodoridis, K. Koutroumbas. *Pattern Recognition.* (4th ed.)Academic Press, San Diego (2008)

Van Cauter et al., 2000. E. Van Cauter, R. Leproult, L. Plat. Age-related changes in slow wave sleep and REM sleep and relationship with growth hormone and cortisol levels in healthy men. *J Am Med Assoc,* 84 (7) (2000), pp. 861–868

Van Cauter et al., 2008. E. Van Cauter, K. Spiegel, E. Tasali, R. Leproult. Metabolic consequences of sleep and sleep loss. *Sleep Med,* 9 (Suppl. 1) (2008), pp. S23–S28

Virkkala et al., 2007a. J. Virkkala, J. Hasan, A. Värri, S.L. Himanen, K. Müller. Automatic sleep stage classification using two-channel electro-oculography. *J Neurosci Methods,* 166 (1) (2007), pp. 109–115

Virkkala et al., 2007b. J. Virkkala, J. Hasan, A. Värri, S.L. Himanen, K. Müller. Automatic detection of slow wave sleep using two channel electro-oculography. *J Neurosci Methods,* 160 (1) (2007), pp. 171–177

Corresponding author at: Department of Mechanical and Electro-mechanical Engineering, National Sun Yat-Sen University, No. 70, Lienhai Road, Kaohsiung 80424, Taiwan, Tel.: +1 886 07 5254232; fax: +1 886 07 5254299.