

# Automated NREM Sleep Staging Using the Electro-oculogram: A pilot study

Gary Garcia-Molina, Farhad Abtahi and Miguel Lagares-Lemos

**Abstract**—Automatic sleep staging from convenient and unobtrusive sensors has received considerable attention lately because this can enable a large range of potential applications in the clinical and consumer fields. In this paper the focus is on achieving non-REM (NREM) sleep staging from ocular electrodes. From these signals, specific patterns related to sleep such as slow eye movements, K-complexes, eye blinks, and spectral features are estimated. Although such patterns are characteristic of the Electroencephalogram, they can also be visible to a lesser extent on signals from ocular electrodes. Automatic sleep staging was implemented using two approaches: i) based on a state-machine and ii) using a neural network. The first one relied on the recommendations of the American Academy of Sleep Medicine, and the second one used a multilayer perceptron which was trained on manually sleep-staged data. Results were obtained on the data of five volunteers who participated in a nap experiment. Manual sleep staging of this data, performed by an expert, was used as reference. Five stages were considered, namely wake with eyes open, wake with eyes closed, and sleep stages N1, N2, and N3. The results were characterized in terms of confusion matrices from which the Cohen's  $\kappa$  coefficients were estimated. The values of  $\kappa$  for both the state-machine and neural-network based automatic sleep staging approaches were 0.79 and 0.59 respectively. Thus, the state-machine based approach shows a very good agreement with manual staging of sleep-data.

## I. INTRODUCTION

Sleep is understood as a reversible state of unconsciousness, characterized by a decrease of activity and alertness. While some controversy exists about the precise role of sleep, there is no doubt that sleep, by changing so many aspects of physiology and behavior, affects the vast majority of body functions including: immune function, hormonal regulation, metabolism, and thermoregulation [1].

Two distinct types of sleep occur in mammals: rapid eye movement (REM) sleep, and non-REM sleep [2]. Compared to the low voltage, high frequency patterns appearing in the awake electroencephalogram (EEG), non-REM (NREM) sleep is associated with a synchronized EEG pattern in which specific electrographic events take place. These events are sleep spindles, K-complexes, and high-voltage slow wave activity (SWA) within the delta frequency band (between 0.5 and 4.0 Hz) that can be recorded over the entire cortical surface [3].

In humans, NREM is sub-divided into stages 2 and 3-4 (presently named N3) depending on the proportions of each

G. Garcia-Molina is with Philips Research North America, USA, gary.garcia@philips.com

F. Abtahi is with the School of Technology and Health of the KTH Royal Institute of Technology, Sweden, sabbt@kth.se

M. Lagares-Lemos is with Universidad Carlos III de Madrid, Spain, miguel.lagares@uc3m.es

of these electrographic events [4]. During REM sleep, the EEG exhibits an activated pattern similar to that observed during wakefulness. REM sleep is distinguished from wakefulness, primarily by reduced responsiveness and muscle atonia [2].

Sleep is traditionally monitored using a set of simultaneously recorded electro-physiological signals (polysomnography a.k.a. PSG) which includes EEG, electro-oculogram (EOG), electromyogram (EMG), and electrocardiogram (ECG) recordings [5]. Visual sleep scoring of a full night recording is a rather time consuming task. This and the limitations of classical sleep scoring rules (the so called R&K rules [6]) in characterizing certain pathologies has motivated the emergence of computer-based automatic sleep scoring [5].

From the perspective of consumer applications of sleep monitoring, the use of full PSG to characterize sleep is unpractical. In this paper, we explore the possibility of using electrodes located around the eyes (ocular electrodes) to characterize the stages of NREM sleep. The use of ocular electrodes is motivated by a prospective application involving a sleeping mask with embedded electrodes. In view of a nap application, the focus of this paper is on NREM sleep.

From ocular electrodes one can extract information about eye blinks and eye movements. Slow eye movements are particularly relevant to identify the wake to sleep transition. In addition, EEG pattern characteristic of NREM sleep such as K-complexes spindles, and slow wave activity can be observed in the signal captured through ocular electrodes. This paper presents two methods for automatic sleep staging using EOG signals. The first method is based on a state-machine that implements sleep staging rules based on the American Academy of Sleep Medicine (AASM) standard [4]. The second method relies on the use of Neural Networks.

This paper is organized as follows. Section II presents the experimental design, data acquisition protocols, and the algorithms used in this paper for signal analysis (Section II-B) as well as automatic sleep staging (Section II-C). Section III presents the results and discussion. The conclusions are presented in Section IV.

## II. METHODS

### A. Data acquisition

Five volunteers (all males, Age:  $27.6 \pm 3.8$ ) participated in a nap pilot experiment in which they were asked to sleep for at least 30 minutes but no longer than 90 minutes. The participants entered a quiet and dim-lighted room at 14:00 and could remain there (asleep or not) until 17:00. The five

participants in this experiment were selected out of a pool of volunteers on the basis of their willingness and ability to take a nap during daytime.

The signals at EEG standard locations C4 and A1, four ocular signals (1 cm below and above the left outer canthus, and 1 cm below and above the right outer canthus), and chin EMG signals (according to the standard in [4]) were recorded using a BIOSEMI Active-Two system [7] at a sampling frequency of 2048 Hz.

For the purposes of eye-blink detection (see Section II-B.1) we refer to as *vertical signal*, the signal resulting from subtracting the lower from the upper ocular signal of the left eye (the left eye signals were arbitrarily chosen for eye blink detection).

We refer to as left ocular channel (LOC) the signal recorded at the lower left outer canthus referenced to A1. In addition, we refer to as right ocular channel (ROC) the signal recorded at the upper right outer canthus referenced to A1.

Visual scoring of the sleep stages was done, by an experienced sleep scorer, on the basis of 30-second long *epochs* using the EEG channel C4-A1, the EOG signals LOC and ROC, and the chin EMG signal. Although these recordings did not include all signals recommended for PSG recordings, they allow for a sufficiently accurate manual scoring of NREM and REM stages.

### B. Data Processing

Prior to the processing and feature extraction steps, the recorded signals were preprocessed by applying a notch filter at 50 Hz to remove the power-line noise, followed by a sub-sampling step at 128 Hz.

Because of their relevance for sleep, eye blinks, slow eye movements, and K-complexes are detected in the signals. In addition, spectral features are also extracted on a per-epoch based. The role of these features in sleep staging is summarized in Table I.

1) *Eye Blink Detection*: The presence of eye blinks is indicative of wakefulness. The detection of eye blinks in this paper follows the implementation in [8] where the processing of the vertical signal essentially consists of two steps. The first step involves applying a non-linear 10-sample long median filter to reject outliers followed by a FIR band-pass filter in the frequency range from 1.5 to 8 Hz. The second step consists in squaring each sample of the signal resulting from the first step, smoothing the result through a running average filter, and applying a threshold to detect the blinks. An illustration of the result of this process is presented in Fig. 1.

2) *Slow Eye Movement Detection*: Slow eye movements (SEMs) are low frequency (mainly 0.2 to 0.6 Hz) rolling, horizontal, bidirectional and conjugate movements of the eyes. They are a phenomenon typical of the wake-sleep transition [9]. SEM activity starts before the onset of stage 1 sleep, continues through stage 1 then declines progressively during the first minutes of stage 2, completely disappearing when spindles and K-complexes begin [10].

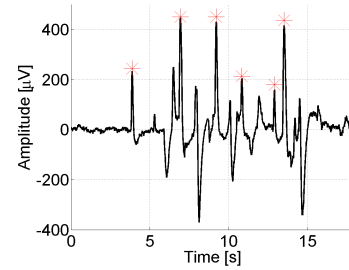


Fig. 1. Detected eye blinks (marked by the stars) using the procedure in [8].

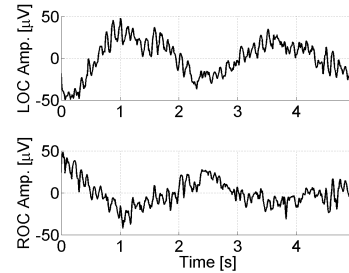


Fig. 2. LOC and ROC signals during SEMs.

Examples of SEMs are depicted in Fig. 2, where the corresponding ROC and LOC signals are conjugate, i.e. have opposite phase. In this particular example, alpha oscillations (8-12 Hz) superpose to the SEMs. Alpha is typical of the wake state and the SEMs indicate the wake-sleep transition.

An estimation of SEM activity for each epoch is extracted based on the difference between the cross-correlations between the signals LOC and ROC for the band from 1 to 6 Hz and the band from 0.5 to 6 Hz.

3) *K-complex detection*: The K-complex (KC) is a major grapho-element of the sleep EEG. The K-complex manifests as a well-delineated negative sharp wave immediately followed by a positive component standing out from the background EEG, lasting for longer than 500 milliseconds. The KC amplitude is usually maximal when recorded using frontal derivations [4].

It is accepted that the KC (also called vertex wave) is a reliable sign for advanced drowsiness, becoming more frequent with the deepening of sleep [11]. The K-Complex detection procedure in this paper is done as follows:

- Band pass filtering in the 0.5 to 3 Hz,
- applying a smoothing moving average filter,
- finding the negative and positive peaks of the signal,
- and checking for the amplitude threshold of a K-complex ( $75\mu\text{V}$  peak-to-peak).

An example of a detected K-Complex in the LOC is shown in Fig. 3.

4) *Spectral feature estimation*: Activity in the following frequency bands is relevant for sleep scoring [4]: i) delta ( $\delta$ ) band (0.5-4 Hz) which is particularly prominent in N3, ii) theta ( $\theta$ ) band (4-7 Hz) which is usually seen during drowsiness, light sleep and REM sleep, iii) alpha ( $\alpha$ ) band

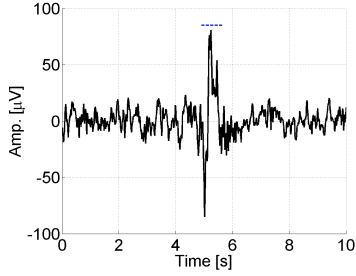


Fig. 3. Detected K-complex in the LOC.

(8-12 Hz) which is typical of the awake (with eyes closed) state, iv) the sigma ( $\sigma$ ) band (11-15 Hz) which corresponds to sleep spindles that are characteristic of N2, and iv) the beta ( $\beta$ ) band (15-30 Hz) characteristic of the awake (with eyes open) state. Although these bands are primarily considered for EEG analysis, they are also visible, though to a lesser extent, in EOG recordings.

In addition to the per-epoch estimation of the absolute power in the above mentioned frequency bands, the power percentage of each band in a given epoch was also estimated. The percentage estimation consisted in applying first a band-pass filter in the 0.5- 30 Hz band, analyzing the 30-second long epoch into one-second long segments (*osl-segment*) with 0.5 seconds overlap, fitting an autoregressive (AR) model [12] with order 15 to each *osl-segment*, estimating the power spectral density from the AR model, and identifying the frequency band in which the maximum power occurs. The power percentage of a given band in an epoch is then equal to the number of times in which the (per *osl-segment*) maximum power occurred in that band divided by the total number of *osl-segments* in the epoch (i.e. 59).

We refer to the  $\delta$ ,  $\theta$ , and  $\alpha$  power percentages as  $\delta\%$ ,  $\theta\%$ , and  $\alpha\%$  respectively. The absolute power in each band is referenced by the corresponding Greek letter.

### C. Automatic Sleep Staging

For the purpose of a consumer application, this paper focusses on five stages, namely wake with eyes open (WO), wake with eyes closed (WC), and sleep stages (N1, N2, and N3) as defined by the standard [4]. Automatic staging is performed on a per-epoch basis. Two types of approaches were tested, the first one implements a state machine while the second one relies on a neural network.

1) *State machine based*: In accordance with the standard rules for sleep scoring [4], a state machine based scoring algorithm was devised in which the next epoch estimated stage depends on the current stage (second row from the top in Table I) in accordance with the conditions specified in Table I.

2) *Neural network based*: To use a neural network (NN) based classifier, features were extracted from an epoch and from both ocular channels (LOC and ROC). The following features were used: 1)  $\delta\%$ , 2)  $\theta\%$ , 3)  $\alpha\%$ , 4)  $\beta\%$ , 5) the ratio  $(\alpha + \beta)/(\theta + \delta)$  (which is associated with sleep depth [13]), 6) the scoring result of the previous epoch, 7) to 10) ratios

TABLE I  
SLEEP-STAGE TRANSITION TABLE

	Current stage				
	WO	WC	N1	N2	N3
$\alpha\% \geq 50\%$	WC	WC	WC	WC	WC
$\alpha\% < 50\%$		N1	N1		
SEMs are present		N1	N1	N1	N1
Eye blinks are present	WO	WO	WO	WO	WO
Sleep spindles are present			N2	N2	N2
K-complexes are present			N2	N2	N3
Slow wave activity ( $\delta\% \geq 50\%$ )				N3	
$\delta\% < 50\%$					N2
$\delta\% + \theta\% > 20\%$		N1			

between the  $\delta\%$ ,  $\theta\%$ ,  $\alpha\%$ , and  $\beta\%$  in the current epoch to the power percentages in the previous epoch. This resulted in a feature vector comprising 20 features (10 per ocular channel LOC and ROC).

An NN with one hidden layer was selected for this approach. The NN had then, 20 units in the input layer, 10 units in the hidden layer, and five units in the output layer. This approach was evaluated under the leave-one-participant-out modality in the sense that the data of 4 participants in the experiment was used to train the NN, and the data of the fifth participant was used to evaluate the staging accuracy against the hypnogram. Thus, five runs could be executed.

### III. RESULTS AND DISCUSSION

The sleep staging accuracy of both approaches, namely state machine and NN based was assessed from the confusion matrix and the Cohen's  $\kappa$  coefficient [14]. The confusion matrices (see Table II) were estimated by aggregating the data of all participants in the experiment.

A total of 673 epochs from all five participants were manually and automatically (using the two approaches in Section II-C) staged. The number of epochs corresponding to WO, WC, N1, N2, and N3 were 52, 81, 122, 259, and 159 respectively. The confusion matrices clearly show the superiority of the state-machine based approach for automatic sleep staging. The WO state can be recognized perfectly using the criteria from Table I. Interestingly N3 appears to be better identified using the NN-based approach. This is a motivation for the implementation of hybrid solutions combining the physiological knowledge summarized in Table I and machine learning algorithms.

In addition to the confusion matrices, the  $\kappa$  coefficient was also estimated because this provides a more robust estimate of the automatic staging performance as compared to the simple agreement percentage [14]. The  $\kappa$  estimates can be obtained from the confusion matrix using:

$$\kappa = \frac{N \sum_{i=1}^5 x_{ii} - \sum_{i=1}^5 \mathbf{x}_{i,:} \cdot \mathbf{x}_{:,i}}{N^2 - \sum_{i=1}^5 \mathbf{x}_{i,:} \cdot \mathbf{x}_{:,i}}, \quad (1)$$

where  $N$  is equal to 673,  $x_{ii}$  is the  $i$ -th diagonal term of the confusion matrix,  $\mathbf{x}_{i,:}$  is the sum of the elements on the  $i$ -th

TABLE II  
CONFUSION MATRIX OF THE STATE MACHINE BASED SLEEP STAGING

Actual stage ↓	Predicted stage				
	WO	WC	N1	N2	N3
WO	52	0	0	0	0
WC	0	73	8	0	0
N1	0	4	101	17	0
N2	0	2	7	227	23
N3	0	0	0	45	114

TABLE III  
CONFUSION MATRIX OF THE NN BASED SLEEP STAGING

Actual stage ↓	Predicted stage				
	WO	WC	N1	N2	N3
WO	30	7	8	1	6
WC	11	30	34	6	0
N1	0	14	81	26	1
N2	4	0	32	192	31
N3	0	0	1	23	135

row, and  $\sum x_{:,i}$  is the sum of the elements on the  $i$ -th column.

It is considered that a value of  $\kappa$  in the 0.40 to 0.60 (resp. 0.60 to 0.80) interval can be interpreted as a moderate (resp. good) performance [15]. The values of  $\kappa$  for the state-machine and NN confusion matrices are 0.79 and 0.59 respectively. This confirms the superiority of the method based on the state-machine. This is due to the fact that the state machine approach incorporates knowledge (endorsed by the AASM [4]) that is normally utilized by a human sleep scorer. An illustration of the good agreement between the state-machine based sleep staging and the manual scoring (referred to as hypnogram) is presented in Fig. 4.

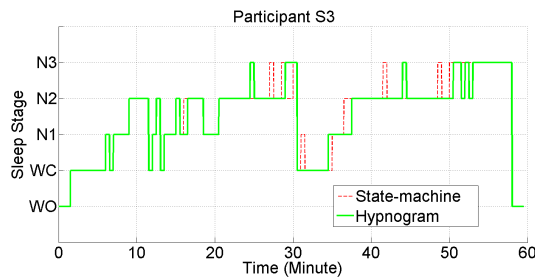


Fig. 4. Hypnogram versus state-machine based automatic sleep staging for participant S3.

#### IV. CONCLUSIONS

In this paper, it has been shown that EOG signals can be used for automatic sleep staging. The motivation behind the use of EOG is of practical nature in view of a potential consumer application. The advantage of using EOG only as opposed to full PSG recordings is evident given the latter's cumbersome procedures.

To achieve automatic sleep stage classification, two methods were tested based on a state-machine and on a neural-network based implementation respectively. The state-machine approach implemented rules for sleep staging as recommended by the AASM standard. Thus, it was necessary to implement the automatic detection of particular patterns

in the signal that characterize sleep such as: slow eye movements, eye blinks, K-complexes, spindles, and spectral features. Although most of these patterns were defined on the basis of the EEG, they are also visible (to a lesser extent) in the EOG. The neural-network based approach extracted ten features for the LOC and ROC and combined them into a feature vector. A multilayer perceptron was then trained on a leave-one-out-participant modality.

The results obtained in this paper as characterized by the well known  $\kappa$  coefficient show the superiority of the state-machine based approach over the neural-network based one. It is important to mention that the  $\kappa$  coefficient for the state-machine based approach was equal to 0.79. This constitutes a very promising result which shows the effectiveness of using EOG for automatic sleep staging.

#### REFERENCES

- [1] C. Cirelli and G. Tononi, "Is sleep essential?" *PLoS biology*, vol. 6, no. 8, p. e216, Aug. 2008.
- [2] C. Amlaner, J. Lesku, and N. Rattenborg, "The Evolution of Sleep," in *SRS Basics of Sleep Guide*, 2nd ed. Sleep Research Society, 2011, ch. 1, pp. 11–19.
- [3] I. de Andrés, M. Garzón, and F. Reinoso-Suárez, "Functional Anatomy of Non-REM Sleep." *Frontiers in neurology*, no. November, p. 2:70, Jan. 2011.
- [4] C. Iber, S. Ancoli-Israel, A. Chesson, and S. Quan, *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*, 1st ed. American Academy of Sleep Medicine, 2007.
- [5] T. Penzel and R. Conrads, "Computer based sleep recording and analysis." *Sleep medicine reviews*, vol. 4, no. 2, pp. 131–148, Apr. 2000.
- [6] A. Rechtschaffen and A. Kales, *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects*. Bethesda, Md.: U.S. Dept. of Health, Education, and Welfare, Public Health Services-National Institutes of Health, National Institute of Neurological Diseases and Blindness, Neurological Information Network., 1968.
- [7] Biosemi B.V., "Biosemi." [Online]. Available: <http://www.biosemi.com/>
- [8] T. Pander, T. Przybyla, and R. Czabanski, "An Application of Detection Function for the Eye Blinking Detection," in *Conference on Human System Interactions*, 2008, pp. 287–291.
- [9] L. De Gennaro, M. Ferrara, F. Ferlazzo, and M. Bertini, "Slow eye movements and EEG power spectra during wake-sleep transition." *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 111, no. 12, pp. 2107–2115, Dec. 2000.
- [10] E. Magosso, F. Provini, P. Montagna, and M. Ursino, "A wavelet based method for automatic detection of slow eye movements: a pilot study." *Medical engineering & physics*, vol. 28, no. 9, pp. 860–875, Nov. 2006.
- [11] F. Amzica and M. Steriade, "The K-complex: its slow ( $\approx$  1-Hz) rhythmicity and relation to delta waves." *Neurology*, vol. 49, no. 4, pp. 952–9, Oct. 1997.
- [12] H. Akaike, "Fitting autoregressive models for prediction," *Annals of the Institute of Statistical Mathematics*, vol. 21, pp. 243–247, 1969.
- [13] D. Petit, D. Lorrain, S. Gauthier, and J. Montplaisir, "Regional Spectral Analysis of the REM Sleep EEG in Mild to Moderate Alzheimer's Disease," *Neurobiology of Aging*, vol. 14, no. 11, pp. 141–145, 1993.
- [14] A. Ben-David, "Comparison of classification accuracy using Cohens Weighted Kappa," *Expert Systems with Applications*, vol. 34, no. 2, pp. 825–832, Feb. 2008.
- [15] D. G. Altman, *Practical statistics for medical research*. Chapman and Hall, 1991, vol. 1st ed.