Journal of Neuroscience Methods 235 (2014) 130–137

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Clinical Neuroscience

Automatic sleep classification using a data-driven topic model reveals latent sleep states

Henriette Koch[a,*], Julie A.E. Christensen[a,b,c], Rune Frandsen[b,d], Marielle Zoetmulder[e], Lars Arvastson[c], Soren R. Christensen[c], Poul Jennum[b,d], Helge B.D. Sorensen[a]

[a] Technical University of Denmark, Department of Electrical Engineering, Ørsteds Plads, Building 349, 2800 Kgs. Lyngby, Denmark
[b] Danish Center for Sleep Medicine, Department of Clinical Neurophysiology, Glostrup Hospital, Nordre Ringvej 57, 2600 Glostrup, Denmark
[c] H. Lundbeck A/S, Ottiliavej 9, 2500 Valby, Denmark
[d] Center for Healthy Ageing, University of Copenhagen, Blegdamsvej 3B, 2200 Copenhagen N, Denmark
[e] Bispebjerg Hospital, Department of Neurology, Bispebjerg Bakke 23, 2400 Copenhagen NV, Denmark

HIGHLIGHTS

- We use a data-driven approach to develop a general and automatic sleep classifier.
- The model uses spectral EEG and EOG as well as eye correlation measures as input.
- Six latent sleep states with concordances to the golden standard are revealed.
- Sleep state transitions are expressed as continuous processes.
- We report good performance on subjects with and without neurodegenerative diseases.

ARTICLE INFO

Article history:
Received 25 January 2014
Received in revised form 29 June 2014
Accepted 1 July 2014
Available online 9 July 2014

Keywords:
Automatic sleep classification
Topic modeling
Electroencephalography (EEG)
Electrooculography (EOG)
Sleep state switching
Neurodegenerative diseases

ABSTRACT

Background: The golden standard for sleep classification uses manual scoring of polysomnography despite points of criticism such as oversimplification, low inter-rater reliability and the standard being designed on young and healthy subjects.
New method: To meet the criticism and reveal the latent sleep states, this study developed a general and automatic sleep classifier using a data-driven approach. Spectral EEG and EOG measures and eye correlation in 1 s windows were calculated and each sleep epoch was expressed as a mixture of probabilities of latent sleep states by using the topic model Latent Dirichlet Allocation. Model application was tested on control subjects and patients with periodic leg movements (PLM) representing a non-neurodegenerative group, and patients with idiopathic REM sleep behavior disorder (iRBD) and Parkinson's Disease (PD) representing a neurodegenerative group. The model was optimized using 50 subjects and validated on 76 subjects.
Results: The optimized sleep model used six topics, and the topic probabilities changed smoothly during transitions. According to the manual scorings, the model scored an overall subject-specific accuracy of 68.3 ± 7.44 (% $\mu \pm \sigma$) and group specific accuracies of 69.0 ± 4.62 (control), 70.1 ± 5.10 (PLM), 67.2 ± 8.30 (iRBD) and 67.7 ± 9.07 (PD).
Comparison with existing method: Statistics of the latent sleep state content showed accordances to the sleep stages defined in the golden standard. However, this study indicates that sleep contains six diverse latent sleep states and that state transitions are continuous processes.
Conclusions: The model is generally applicable and may contribute to the research in neurodegenerative diseases and sleep disorders.

1. Introduction

The golden standard for sleep analysis is defined by the American Academy of Sleep Medicine (AASM) (Iber et al., 2007) and uses manual scoring of polysomnography. AASM divides sleep

* Corresponding author. Tel.: +45 2513 6271/+1 650 716 8353.
E-mail address: henriette.koch@hotmail.com (H. Koch).

http://dx.doi.org/10.1016/j.jneumeth.2014.07.002
0165-0270/

into wakefulness (W), rapid-eye-movement (REM) and non-REM (NREM) sleep, where NREM is further divided into three stages N1–N3 according to the level of drowsiness and pattern of EEG changes. However, manual scoring is criticized for using the predefined epoch length of 30 s and oversimplification of the sleep structures by discretizing sleep stages into abruptly defined stages. Other weaknesses are the low inter-rater reliability of 67–91% (Ahmed and Tafreshi, 2009), large time consumption and the standard suited to fit young healthy subjects even though sleep changes with age (Crowley, 2011) and in patients with neurodegenerative diseases (Santamaria et al., 2011).

Sleep is primarily controlled by neurons in the brainstem and midbrain. State switching is determined by complex neuron pathways which can be described by two interacting flip–flip models (Dyken et al., 2012). These models suggest that only W, REM and NREM sleep are stable sleep states, states switches are continuous processes and that degeneration of the neurons involved will affect the overall sleep pattern.

In this study a general sleep classifier was designed using a data-driven approach to ensure that no biased definitions were used. The number of sleep states was predefined but this study's data-driven approach revealed the hidden and system immanent sleep states. Through a number of mathematical steps raw electroencephalography (EEG)/electrooculography (EOG) signals were converted into text strings and the topic model Latent Dirichlet Allocation (LDA) was applied to these text strings. LDA (Blei et al., 2003) analyzes the presence of words included in each text document in the text corpus. This is used to calculate the mixture of probabilities for each latent topic in each individual text document. In the context of sleep classification, raw EEG/EOG was analyzed in 1 s windows and converted into "words" with each 30 s epoch representing a document and topics representing latent sleep stages. The 1 s resolution ensures that micro-sleep events, which typically have a duration of 0.5–3 s, contribute to the epoch classification (Iber et al., 2007). The sleep model was trained on control subjects and applied to iRBD and PD patients, representing a neurodegenerative disease (NDD) group, and control subjects and patients with periodic leg movement (PLM), representing a non-NDD group. PLM patients were included to stress the model with consecutive limb movements and arousals but these patients show no signs of neurodegeneration.

2. Materials and methods

2.1. Data

Sleep recordings from 126 subjects were enrolled in this study. The Danish Center for Sleep Medicine in the Department of Clinical Neurophysiology, Glostrup University Hospital provided all data and carried out one manual sleep scoring according to the AASM standard for each polysomnography (PSG). Control subjects had no history of diagnosed sleep disorders, dream-enacting or movement disorder. Patients treated with anti-depressant drugs including hypnotics were excluded but dopaminergic treatments were continued. The dose of dopa may affect the vigilance states (Micallef et al., 2009) but it is clinical practice that patients continue the intake of dopa in parallel with the sleep recording. This study included clinical data to ensure a model suitable for clinical use was developed, and the dopaminergic treatments were therefore continued. The demographic data and leg movement (LM) index (number of leg movements per hour) are summarized in Table 1. Disease durations for the PD patients were 4.0 ± 4.0 years ($\mu \pm \sigma$) and 7.6 ± 4.0 years in the base term matrix and validation dataset respectively. One PD patient was later diagnosed with multipel system atrofi and one developed dementia with lewy bodies. Each dataset was designed to match age with no further knowledge about the subjects.

Table 1
Demographics and leg movement (LM) index.

Patient class	Subjects (F/M)	Age (years, $\mu \pm \sigma$)	LM-index ($\mu \pm \sigma$)
Train sleep model			
Control	10 (5/5)	57.2 ± 8.1	38.6 ± 41.3
Base term matrix			
Control	10 (5/5)	59.8 ± 8.0	21.3 ± 14.1
PLM	10 (6/4)	57.8 ± 9.3	48.0 ± 31.5
iRBD	10 (2/8)	59.0 ± 13.5	55.9 ± 34.2
PD	10 (4/6)	63.2 ± 6.0	52.8 ± 54.2
Validation			
Control	13 (11/2)	54.4 ± 9.1	15.5 ± 11.1
PLM	15 (6/9)	56.3 ± 12.6	62.2 ± 37.7
iRBD	22 (3/19)	63.6 ± 5.7	31.8 ± 23.3
PD	26 (8/18)	66.4 ± 6.7	31.2 ± 28.7

PSG was carried out as outpatient for control subjects, and for patients PSG was carried out as outpatient or inpatient according to the AASM standard (Iber et al., 2007). Recordings with artefacts such as electrode disconnections and continuous noise were excluded through visual inspection. Two or more PSGs were performed only if the prior recording did not meet the acceptable quality for clinical use. First night recordings may impact the sleep stage distribution but this was considered not to affect the micro-sleep distribution within the different sleep stages. Further, first night recordings were included in all subject groups in the model throughout the optimization and validation.

The raw sleep data and hypnograms were extracted from Nervus (V5.5, Cephalon DK, Nørresundby, Denmark) using the built-in export data tool. Further analysis was carried out in MATLAB (R2011b, MathWorks, Inc., Natick, MA, USA) with a sampling frequency of 256 Hz.

2.2. Automatic sleep classification

The optimized sleep classifier used two EEG channels (C3-A2 and O1-A2) and two EOG channels (EOGL-A2 and EOGR-A1 placed one cm out/up and out/down, respectively, from the canthus) in between lights-off and lights-on. The steps involved in the classifier are shown in Fig. 1.

Initially, the signals were filtered forward and reversed in time with 4th order Butterworth filters (3 dB) using cut-off frequencies 0.3 and 35 Hz for EEG and 0.3 and 10 Hz for EOG. These filters were chosen due to narrow transition bands and no ripples. Double filtering ensured a linear phase-response and the squared magnitude was equalized later on. The cut off frequencies for EEG and EOG were chosen according to the AASM standard and to focus the signals on eye movements, respectively. No other removal of intermittent artefacts was performed because the study aimed at using a data-driven approach and should be able to handle such cases. The filtering was followed by word creation and word counting, which involves contiguous mathematical steps. The word count output describes the signal characteristics and is given as input to the LDA model. It is common to view data as text documents in the setting of LDA. This is no limitation to the actual type of data analyzed.

2.2.1. Signal characteristics

2.2.1.1. EEG words. The single-sided amplitude spectrum was calculated by fast Fourier transform in non-overlapping 1 s segments by using MATLAB's built-in *fft* function, a rectangular window and zero-padding. For EEG channels, the clinical frequency bands ($\delta < 4$, $4 \leq \theta < 8$, $8 \leq \alpha < 14$, $14 \leq \beta \leq 30$ Hz) were used, and within each frequency band the frequency content was summed. Symbolization was used to define the amplitude level in each 1 s segment and this