COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE 124 (2016) 180–192

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

ISRUC-Sleep: A comprehensive public dataset for sleep researchers

Sirvan Khalighi[a,*], *Teresa Sousa*[a], *José Moutinho Santos*[b], *Urbano Nunes*[a]

[a] *Institute of Systems and Robotics (ISR-UC), Department of Electrical and Computer Engineering, University of Coimbra, Portugal*
[b] *Sleep Medicine Centre, The Central Hospital of University of Coimbra (CHUC), Portugal*

ARTICLE INFO

Article history:
Received 6 May 2015
Received in revised form
6 September 2015
Accepted 5 October 2015

Keywords:
Sleep dataset
Automatic sleep stage classification
Polysomnographic signals
Effects of sleep disorder
Medication effects
Feature selection

ABSTRACT

To facilitate the performance comparison of new methods for sleep patterns analysis, datasets with quality content, publicly-available, are very important and useful.

We introduce an open-access comprehensive sleep dataset, called ISRUC-Sleep. The data were obtained from human adults, including healthy subjects, subjects with sleep disorders, and subjects under the effect of sleep medication. Each recording was randomly selected between PSG recordings that were acquired by the Sleep Medicine Centre of the Hospital of Coimbra University (CHUC). The dataset comprises three groups of data: (1) data concerning 100 subjects, with one recording session per subject; (2) data gathered from 8 subjects; two recording sessions were performed per subject, and (3) data collected from one recording session related to 10 healthy subjects. The polysomnography (PSG) recordings, associated with each subject, were visually scored by two human experts.

Comparing the existing sleep-related public datasets, ISRUC-Sleep provides data of a reasonable number of subjects with different characteristics such as: data useful for studies involving changes in the PSG signals over time; and data of healthy subjects useful for studies involving comparison of healthy subjects with the patients, suffering from sleep disorders.

This dataset was created aiming to complement existing datasets by providing easy-to-apply data collection with some characteristics not covered yet. ISRUC-Sleep can be useful for analysis of new contributions: (i) in biomedical signal processing; (ii) in development of ASSC methods; and (iii) on sleep physiology studies. To evaluate and compare new contributions, which use this dataset as a benchmark, results of applying a subject-independent automatic sleep stage classification (ASSC) method on ISRUC-Sleep dataset are presented.

1. Introduction

Sleep is an active and regulated process with an essential restorative function for physical and mental health [1]. Quality of sleep and sleep disorders have an important effect on the health and quality of life. The study of individual behaviors during sleep through all-night PSG recordings has consistently been an important research topic.

* *Corresponding author*. Tel.: +351 239796201.
E-mail address: skhalighi@isr.uc.pt (S. Khalighi).

http://dx.doi.org/10.1016/j.cmpb.2015.10.013
0169-2607/

Numerous methods have been developed for automatic detection of arousals, apnea, and sleep stages [2–5]. These methods often use PSG recordings, including electrophysiological signals (electrocardiographic activity, brain-wave patterns, eye movements, and activation signal of muscles), pneumological signals (airflow, blood oxygen level, and movement of respiratory muscles), and other contextual information (body position, lights, snore recording, etc.) [6]. These signals have been collected from human individuals using noninvasive surface electrodes. To evaluate the efficiency of automatic sleep pattern analysis methods, non-public and few existing public datasets have been used. Rigorous comparisons between the developed methods cannot be done since the used datasets differ in recording conditions, physiological conditions of subjects and number of assessed subjects. Datasets with quality content, publicly available, are an important vehicle for accelerating research, since they facilitate the performance comparison of new approaches and methods. This paper presents three main contributions:

- We introduce a publicly-available comprehensive sleep dataset, called ISRUC-Sleep, which comprises three subgroups as illustrated in Fig. 1. The subgroups of the dataset contain PSG signals of different adult individuals, including healthy subjects, subjects with sleep disorders, and subjects under the effect of sleep medication. Sleep stages were labeled by two sleep experts. Furthermore, for 8 subjects (subgroup-II), two sets of PSG data, which have been recorded at different time dates, are provided.
- Aiming to help sleep researchers in their analysis and inferences using this dataset, for each subject, useful and complementary information related to sleep disorders, used medications, and their effect on sleep patterns, are presented.
- Aiming to evaluate and compare of new contributions, which will use this dataset as a benchmark, results of applying a subject-independent ASSC method on ISRUC-Sleep dataset are presented. This supervised-learning based method, detailed in Khalighi et al. [7], is henceforth named SSM4S.[1]

2. Terminology and definitions

Background material, terms definition (Table 1), and effects of sleep disorders and medications on sleep patterns are summarized in the next subsections.

2.1. *Background*

- The Rechtschaffen and Kales standard (R&K) rules are the basis of a consensus scoring procedure for adults [8]. The American academy of sleep medicine (AASM) defined new criteria for sleep scoring based on the R&K rules. In adults, sleep-wake cycle is categorized in awake, non-rapid eye movement (NREM) and rapid eye movement (REM) sleep stages. NREM sleep is further divided into three stages: N1 (drowsiness/transitional sleep), N2 (light sleep) and N3 (deep sleep) [9], the last of which is also called delta sleep or slow wave sleep (SWS). The 2007 AASM visual scoring rules recommend a frontal electrode for best detecting K-complexes, a central electrode for spindles, and an occipital electrode for alpha waves [10]. Based on both scoring rule-sets (R&K and AASM), epochs of 30 s (more rarely 20 s) are defined for the PSG signals scoring [11]. Difficulties in sleep scoring arise when sleep does not behave in accordance with the *normal*/expected way as a consequence of sleep disorders, medication or in face of individual specific characteristics of sleep Electroencephalogram (EEG).
- Since the collected PSG signals are characterized by low signal-to-noise ratio (SNR), a preprocessing stage is applied to improve the quality of the signals; i.e. some channels of the recorded signals are filtered to eliminate noise and undesired background EEG, aiming to enhance the PSG signal quality and increase the SNR. The filtering stage comprises: (1) a notch filter to eliminate the 50 Hz electrical noise; (2) a bandpass Butterworth filter with a lower cutoff of frequency 0.3 Hz and higher cutoff of frequency 35 Hz for EEG and EOG channels, and a lower cutoff of frequency 10 Hz and higher cutoff of frequency 70 Hz for EMG channels. More details are presented in Table 3.
- The common EEG frequency bands are: low delta 0.3–1 Hz, delta 1–4 Hz, theta 4–8 Hz, alpha 8–12 Hz, sigma 12–15 Hz, and beta 15–30 Hz. Different EEG waves (alpha, beta, sigma, delta, and theta) characterize different sleep stages. Low amplitude, mixed EEG frequency, saw-tooth pattern, low amplitude Electromyogram (EMG) and high level Electrooculogram (EOG) signals from both eyes, are apparent during the REM stage. In stage N1, waves with high amplitude and frequency range of 2–7 Hz together with the existence of alpha waves are found in EEG signal. Still regarding N1, EMG level is lower when compared to the awake stage. Sleep spindles (12–14 Hz) and K-complexes are observed during N2. N3 (deep sleep) consists of low-frequency high-amplitude waves with frequencies of 2–4 Hz.

2.2. *Effect of sleep related disorders and sleep pathology on sleep stage patterns*

Sleep apnea is the most frequent sleep disorder seen in sleep medicine centers. The syndrome is characterized by repetitive episodes of upper airway obstruction that occur during sleep and are usually associated with a reduction in blood oxygen saturation. These nocturnal respiratory disturbances result in brief arousals in sleep, which promotes sleep fragmentation that typically disturbs sleep architecture with reduction or even complete deprivation of REM sleep and N3 sleep. An increase of arousals of different length together with an increase in sleep stage changes is a feature of the syndrome. This fragmentation of sleep, inhibiting cortical synchronization, would be responsible for the lower amount of slow wave sequences of the deep sleep [12]. On the other hand, transient experimental hypoxia induced abnormal posterior resting state delta and alpha rhythms in healthy volunteers, and EEG slowing during awake with an increase in relative theta and delta power in occipital, temporal and parietal areas

[1] Sirvan Supervised Method for Sleep Staging.