Available online at www.sciencedirect.com
ScienceDirect
Biomedical Signal Processing and Control 2 (2007) 171–179
Biomedical Signal Processing and Control
www.elsevier.com/locate/bspc

Feature selection for sleep/wake stages classification using data driven methods

Lukáš Zoubek [a,b], Sylvie Charbonnier [b,*], Suzanne Lesecq [b], Alain Buguet [e], Florian Chapotot [c,d]

[a] VSB-Technical University of Ostrava, Department of Measurement and Control, 17. listopadu 15/2172, Ostrava-Poruba 708 33, Czech Republic
[b] GIPSA-Lab, Control Systems Department, BP 46, 38402 Saint Martin d'Hères Cedex, France
[c] Sleep, Chronobiology and Neuroendocrinology Laboratory, The University of Chicago, 5841 South Maryland Avenue, AMB M374 (MC 1027) Chicago, IL, USA
[d] PhiTools, 1 rue du Général de Castelnau, Strasbourg, France[1]
[e] Neurobiologie des états de vigilance, EA 3734, Université Claude-Bernard Lyon 1, Lyon, France

Received 27 February 2007; received in revised form 11 May 2007; accepted 18 May 2007
Available online 2 July 2007

Abstract

This paper focuses on the problem of selecting relevant features extracted from human polysomnographic (PSG) signals to perform accurate sleep/wake stages classification. Extraction of various features from the electroencephalogram (EEG), the electro-oculogram (EOG) and the electromyogram (EMG) processed in the frequency and time domains was achieved using a database of 47 night sleep recordings obtained from healthy adults in laboratory settings. Multiple iterative feature selection and supervised classification methods were applied together with a systematic statistical assessment of the classification performances. Our results show that using a simple set of features such as relative EEG powers in five frequency bands yields an agreement of 71% with the whole database classification of two human experts. These performances are within the range of existing classification systems. The addition of features extracted from the EOG and EMG signals makes it possible to reach about 80% of agreement with the expert classification. The most significant improvement on classification accuracy is obtained on NREM sleep stage I, a stage of transition between sleep and wakefulness.

Keywords: Decision making; Diagnosis; Medical applications; Pattern recognition; Signal processing

1. Introduction

Polysomnography (PSG) consists in study of sleep and wakefulness from the concurrent recording of multiple bioelectric signals including the electroencephalogram (EEG), electro-oculogram (EOG) and electromyogram (EMG). A system of standardized rules established in the conventional Rechtschaffen and Kales (R&K) human sleep/wake stage scoring manual [1] enables the visual recognition by medical and technical experts of up to six different vigilance stages: wakefulness, non-rapid eye-movement (NREM) sleep stages I, II, III and IV, and REM or paradoxical sleep (PS). NREM stages III and IV represent the slow wave sleep (SWS). The successive visual interpretation, by 20 or 30 s epochs, of 8–24 h PSG recordings leads to the representation of the temporal distribution of sleep/wake stages called a hypnogram, an example of which is presented in Fig. 1. A hypnogram reveals the internal architecture of sleep and the alternation of NREM and REM sleep phases, which makes the discrimination between normal and abnormal sleep much simpler. PSG is thus a powerful tool in the diagnosis of sleep disorders, which are rather common with about 5% of the general population affected [2].

Since 1970 and the development of computerized methods, automated systems have emerged in order to automatically score PSG recordings, so as to avoid the expert to spend too much time to this tedious and time-consuming work. The visual interpretation of PSG recordings is a typical pattern recognition task. Physicians look at the signals and classify successive epochs from the shape of their traces. Two problems must be

* Corresponding author. Tel.: +33 476826415; fax: +33 476826388.
E-mail address: Sylvie.Charbonnier@inpg.fr (S. Charbonnier).
[1] www.phitools.com.

1746-8094/$ – see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bspc.2007.05.005

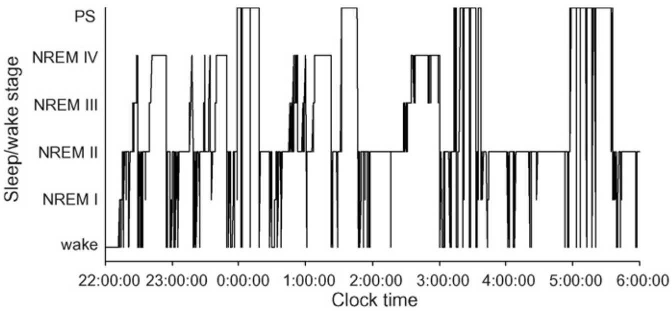

Fig. 1. An example of hypnogram obtained from a night-time PSG recording.

solved to obtain automatic classifiers able to obtain results similar to human experts: to choose the classification function that will give the best results on this problem and to process the signals with adequate techniques so as to obtain inputs to the classifier (features) which are the most similar to the visual information used by the expert.

An important number of publications can be found in the literature on automatic sleep/wake stages analysis. Many of these papers focus on the choice of an adequate type of classifier to achieve accurate classification. The authors use either classical algorithms or artificial intelligence methods, such as neural networks [3–7]. Features used as inputs to the classification systems are extracted from PSG signals at constant intervals (epoch) using various signal processing techniques operating in the time domain and/or in the frequency domain. Though several propositions were made to process the PSG signals, only few studies have been performed to determine the optimal set of features achieving an accurate sleep/wake stage classification [8–14].

Thanks to the development of computerized methods and in parallel to automated systems, a research field has emerged, known as data mining or knowledge discovery. This research field proposes methods that enable the extraction of knowledge from large sets of examples [15]. The aim of the study presented in this paper was to apply data mining methods to extract knowledge about sleep/wake stages classification. Knowledge extraction was performed from a large database composed of 47 night sleep recordings from 41 healthy subjects. Feature selection algorithms and systematic statistical assessment were performed to determine which signals and processing methods are the most relevant and accurate for sleep/wake stage automated classification.

The outline of the paper is the following. The whole database and the techniques used to process the signals and extract the features are presented in Section 2. The features selection methods used are described in detail in Section 3. The results are presented in Section 4 and discussed in Section 5.

2. Materials

2.1. Presentation of the PSG recording database

In this study, a large database of PSG recordings was used. The full database contains 47 night-time PSG recordings obtained from 41 healthy adult subjects (19–47 years old, 39 males and 2 females). Recordings were made continuously during the night (8 h between 22:00 h and 06:00 h). Four EEG channels (C3-A2, P3-A2, C4-A1, and P4-A1), one transversal EOG and one chin EMG were registered and digitized at a sampling frequency $f_s = 128$ Hz. The EEG leads were attached onto the scalp according to the International 10-20 EEG System of Electrodes Placement [16].

All the 47 PSG recordings were visually interpreted by two independent sleep physicians. Visual sleep/wake stage scoring was performed with constant epoch duration of 20 s according to the conventional rules of the R&K manual [1]. Each epoch was thus classified into one of five different stages: wakefulness, NREM sleep stage I, NREM sleep stage II, slow wave sleep (SWS or NREM stages III and IV), and paradoxical sleep. To avoid the introduction of expert inaccuracies in the database, only the epochs classified in the same stage by both experts were considered in this project. They represent 84% of the original PSG recording database and only that subset was used to form our study database. The total number of epochs included was 63,254. As it can be seen in Table 1, the number of epochs classified in each sleep/wake stage is different. NREM stage II lasts a long time, whereas NREM stage I is rather short. To avoid classification errors related to differences in the sample size of each class, the database was further reduced to a smaller one where each class is composed of about the same number of epochs. The numbers of epochs classified in each sleep stage for the database reduced are presented in the second row of Table 1.

The database used in this study thus consisted of 10,000 randomly selected epochs classified into one of the five sleep/wake stages. The set S of 10,000 epochs was split in 10 subsets $S = \{S_1, S_2, \ldots, S_{10}\}$, each subset S_k containing 1000 epochs, equally distributed in the five classes. The size of the subsets (1000 epochs) was chosen from a previous study whose goal was to analyze the effect of the number of examples on the classification error [17]. Its main conclusion was that a minimal number of 500 examples was required to train and validate a classifier on a sleep/wake classification problem and get an unbiased evaluation of the classification accuracy.

2.2. Features extracted from the PSG recordings

Each epoch stored in the database consists of a 20 s recording of six signals (four EEG, one EOG and one EMG) [18]. Since the PSG recordings were sampled at 128 Hz, each time series contains 2560 samples. Various features describing different signal characteristics were extracted from each signal using multiple processing techniques. The PRANA software for PSG analysis (PhiTools, Strasbourg, France) was used to

Table 1
Description of the database used in this study (number of epochs in the sleep/wake stages)

	Wake	NREM I	NREM II	SWS	PS
Full database	5232	1989	32966	7701	15,366
Reduced database	1914	1879	2206	1902	2,099