Expert Systems with Applications 42 (2015) 7825–7830

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Fast and accurate PLS-based classification of EEG sleep using single channel data

Temel Kayikcioglu, Masoud Maleki *, Kubra Eroglu

Karadeniz Technical University, Faculty of Engineering, Department of Electrical and Electronics Engineering, 61080 Trabzon, Turkey

ARTICLE INFO

Article history:
Available online 10 June 2015

Keywords:
Partial least squares regression
Auto-regressive model
Electroencephalograph
k-Nearest neighborhood
Bayes
Linear discriminant classifier

ABSTRACT

Since speed of classification is important to real-time applications, this study proposed fast classification of sleep and wake stages using a single electroencephalograph (EEG) channel. Changes in the sleep and wake stages are accompanied by changes in the frequency spectrum of the EEG signals; so, the features extracted from the 5-s epoch of the EEG using auto-regressive (AR) coefficients were used to represent EEG signals of different sleep and wake stages. The proposed fast classification method was based on partial least squares regression (PLS), which was used to classify these features by finding an optimum beta using K-fold cross validation. The Physionet database was used to confirm accuracy and speed of the proposed classification system. This system could be used in real-time implementations because of its high classification rate, speed and capability to be implemented on hardware owing to be very comfortable. Finally, results of the PLS were compared with those of other classifiers such as *k*-nearest neighborhood (*k*-NN), linear discriminant classifier (LDC) and Bayes. We achieved 91% classification accuracy by selecting PLS as the classifier. These comparisons revealed that the proposed algorithm could recognize an emergency situation in less than 1 s with high accuracy.

1. Introduction

1.1. Background

Visual sleep scoring is a difficult process because of requiring a great deal of time and being a subjective procedure. In response to these challenges, automatic sleep-staging methods based on multichannel signals, including EEG, EMG and EOG (Kuwahara et al., 1988; Park, Park, & Jeong, 2000; Schaltenbrand et al., 1996; Smith, Negin, & Nevis, 1969; Smith and Karacan, 1971), have been developed. Two important items of sleep scoring are feature extraction, which helps researchers to analyze recording epoch, and classification, which helps researchers to recognize sleep stage of the epoch. A few features that adhere to the Rechtschaffen and Kales (R&K) standard have been proposed for sleep staging, which include alpha ratio (Agarwal and Gotman, 2001), spindle ratio (Duman, Erdamar, Erogul, Telatar, & Yetkin, 2009) and SWS ratio (Berthomier, Prado, & Benoit, 1999). Spectral power, power ratio and spectral frequency (Schaltenbrand et al., 1996) have been also used in previous studies. In addition, many methods have been proposed for classification, among which linear discriminant analysis (LDA) (Šušmáková & Krakovská, 2008), artificial neural network (Schaltenbrand et al., 1996), fuzzy system (Berthomier et al., 2007) and decision tree (Anderer et al., 2005) can be mentioned. Success of these methods has been in the range of 80–85%. One recent study (Sheng-Fu, Kuo Hu, Pan, & Wanga, 2012) proposed an automatic sleep-scoring method that combined multi scale entropy (MSE) and autoregressive models for a single-channel EEG. This work also recommended comparatively assessing performance of the method with the manual scoring based on full polysomnograms. Indeed, EEG data have been used for sleep scoring; but, using a system that is fast, accurate and comfortable with an implemented algorithm would be even more beneficial. Most of the previously proposed approaches are not suitable for implementation in real-time systems and many of them are not comfortable for the subjects. Some of the reasons for these shortcomings are low accuracy, infeasibility for hardware implementation, computational complexities and lack of generalization.

The objective of the present analysis was to test and compare performance of the PLS algorithm for sleep scoring with a single-source EEG (a single electrode) to test its feasibility in future works. This article is organized as follows. The following section presents sleep and sleep frequencies. Section 2 introduces the methods. Data acquisition, feature extraction and classifications

* Corresponding author. Mobile: +90 538 5787035.
E-mail addresses: tkayikci@ktu.edu.tr (T. Kayikcioglu), masoud.maleki1361@yahoo.com (M. Maleki), eroglu.kubra@gmail.com (K. Eroglu).

http://dx.doi.org/10.1016/j.eswa.2015.06.010
0957-4174/

are also described in this section. Finally, in Sections 3 and 4, the results and conclusions are respectively presented.

1.2. Sleep

The standard (R&K) for sleep stage classification defines two groups of stages for determining sleep depth. The first stage is non-rapid eye movement (NREM) stage, which is then sub-divided to four stages. Rapid eye movement (REM) stage is the second stage, which is characterized by high ocular activity in the EOG recordings. Four stages of NREM sleep are called Stage 1, Stage 2, Stage 3 and Stage 4. Recently, Stages 3 and 4 have been combined to form new slow-wave sleep stage (SWS) because they exhibit many similar characteristics.

1.3. Sleep frequencies

EEG rhythms are closely related to sleep and wake stages. Characteristics and patterns of the EEG recordings associated with the wake stage and various sleep stages are: alpha-band (8–12 Hz with 20_60 micro volt _ Amplitude) in wake, stages 1 and REM, beta-band (13–49 Hz with 2_20 micro volt _ Amplitude) in wake and theta-band (4–7 Hz with 50_75 micro volt _ Amplitude) in Stages 1, 2, 3 and 4 and REM and delta-band (0–4 Hz with 75 micro volt _ Amplitude) in Stages 3 and 4.

2. Material and methods

2.1. Dataset acquisition

The database used in this study was provided by Physionet (http://www.physionet.org/cgi-bin/atm/ATM). EEG signals were obtained from seven subjects ranging from healthy to abnormal. These subjects included Caucasian males and females (21–35 years old) who were neither on prescription drugs nor on recreational drugs at the time of the study. Sleep EEG for 80 h was extracted from the recordings and sampled at 100 Hz. Format of the dataset was an EDF (European Define Format) so that it could be downloaded. The 10–20 standard electrode placement system was used for the EEG recordings, which contained horizontal EOG, Fpz-Cz EEG and Pz-Oz EEG, each one sampled at 100 Hz. The process of sleep scoring involves identifying EEG signal epochs according to the sleep stage using a graphical plot called a hypnogram, which shows the sleep profile. Hypnograms were manually scored according to the R&K scale based on the Fpz-Cz and Pz-Oz EEGs and then classified into the following stages: waking, NREM 1, NREM 2, NREM 3, NREM 4, REM and movement time (M). One of the hypnograms is shown in Fig. 1. The epochs were pre-processed by filtering all of the data. This filtering step eliminated unwanted artefacts from the EEG data and enhanced their accuracy (http://www.physionet.org/cgi-bin/atm/ATM).

2.2. System design to classification of EEG sleep

Fig. 2 shows the flowchart of the proposed classification of EEG sleep method that includes three parts: (1) pre-processing; (2) feature extraction; and (3) classification. Each of these parts will be explained in the following subsections.

2.3. Pre-processing

The alternations in the EEG signals across the sleep stages are very delicate and therefore require advanced signal processing techniques to extract the features. EEG has different specific frequency components, some of which contain this discriminative information, which includes energy of the delta, theta, alpha and beta bands. This energy is important for classifying different brain states. In this study, the alpha, theta and beta bands were used to classify the sleep and wake stages.

First, the epochs were normalized between $[-1\,1]$ so that they all possessed similar conditions. In the next step, the epochs were filtered by three Butterworth band pass filters in the order of ten for the alpha-band (8–12 Hz), in the order of eleven for the beta-band (13–49 Hz) and in the order of eight for the theta-band (4–8 Hz). After the filtration, all three types of data were ready for any epoch and they were placed as input to the AR model in this format.

2.4. Feature extraction

2.4.1. Auto-regressive coefficients

AR model is a powerful and useful tool for signal modeling. In this model, each sample of a given signal is considered a prediction of the previous weighted samples of that signal. The number of coefficients determines the model order. In this paper, autoregressive coefficients were estimated with Burg method (Stoica and Moses, 1997). The Burg method fits the p'th order AR model to the input signal, x, by minimizing (least squares) the forward and backward prediction errors while constraining AR coefficient, a_i, to satisfy the Levinson–Durbin recursion. Eq. (1) shows the AR model.

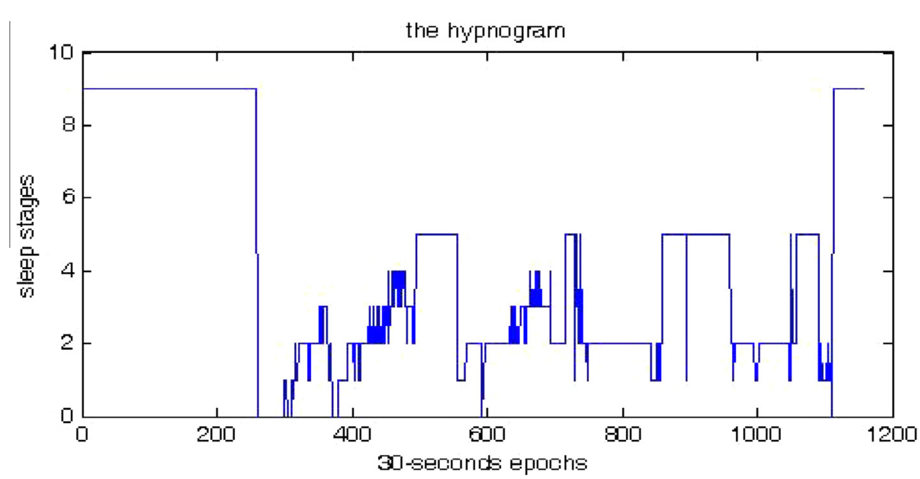

Fig. 1. Hypnogram of a subject; stages 0, 1, 2, 3, 4, and 5 are waking, NREM 1, NREM 2, NREM 3, NREM 4, and REM, respectively.