Paper No. 74

Classification of EEG waveforms by spectral clustering

Evangelia I. Zacharaki*, Evangelia Pippa*, Andreas Koupparis**, George K. Kostopoulos**, Vasileios Megalooikonomou*

*Department of Computer Engineering and Infomatics, University of Patras, Patras 26500, Greece
**Department of Physiology, Medical School, University of Patras, Patras 26500, Greece
ezachar@upatras.gr, pippa@ceid.upatras.gr, akoupparis@gmail.com, gkkostop@med.upatras.gr
vasilis@ceid.upatras.gr

Introduction

Pattern analysis has often been applied for automated extraction of the constituent brain waves and rhythms in EEG. These waveforms, such as the K-complexes, delta waves and spindles, form the sleep microstructure which is examined during diagnosis in sleep studies. Specifically K-complexes have been suggested to protect sleep and also to provide gating functions in idiopathic generalized epilepsies or sleep disorders 1.
Due to the extreme size of all night sleep EEG, visual marking of the EEG waveforms is almost prohibitive in a routine setting and also highly scorer dependent, mainly due to high intra- and inter-subject variability of the EEG characteristics. Thus, various automatic or semi-automatic extraction methods have been proposed in the literature, mostly for detection of waveforms that follow a typical pattern, such as the K-complexes. Most of these methods however suffer from a large number of false positives, when applied without knowledge of the sleep stage 234. In this paper we propose a pattern analysis scheme for reducing false detections of target EEG waveforms without using sleep stage information.
The method has been tested on the detection of K-complexes (KCs). First all possible KC waves are extracted based on empirical rules on fundamental features for the description of KC morphology, such as peak to peak amplitude. Then each candidate wave is classified as KC or outlier by evaluating its similarity, regarding both signal and frequency content, to a set of different patterns of annotated KCs (training set). The different patterns are constructed by applying graph partitioning on the training set using the spectral clustering algorithm.

Materials and Methods

Creation of the testing set: The data used in this work consist of whole-night sleep EEG recordings (13 excerpt files of 0.5 h sleep each) of a healthy volunteer using 10 (out of 58 acquired) EEG channels. The recordings were downsamled to reduce dimensionality and low-pass filtered to remove high frequency noise. Baseline correction was also performed by using a moving average filter in order to correct for slow shifts over time which happen during the recording and alter the zero level. Subsequently, the candidate KCs were automatically detected according to pre-defined expert-based rules 5. The rules are applied on fundamental morphological features which characterize the visual appearance of the signal and imitate the standardized scoring rules by the American Association of Sleep Medicine (AASM) 6. We extracted KC segments of 2sec duration (± 1sec around the peak of the negative phase). These KC segments formed the testing set.

Classification of the KC waveforms: The proposed method applies an outlier detection scheme in order to classify the waveforms in the testing set into true positives (TP) and false positives (FP). The feature representation of each KC segment was based on two temporal patterns: (i) the amplitude change over time (signal representation) and (ii) the frequency content over time, which was calculated as the power spectral density integrated in the frequency range of the waveform. Thus two feature vectors (time series) were produced for each KC segment. Both vectors were downsampled and z-score normalized. Then, each vector was assigned a pseudo-probability value of being a KC depending on its similarity with manually marked KC segments, as described next. The classification algorithm includes a training and a testing phase.

In the *training phase* annotated KCs were used to learn the different patterns of KC morphology. In order to extract these patterns, a graph partitioning technique known as spectral clustering 7 was applied to partition the training set into a set of clusters. Spectral clustering divides graph nodes into groups so that connectivity is maximized between nodes in the same cluster and minimized between nodes in different clusters. Connectivity is measured by some affinity (similarity) measure. We used the following quantity as affinity measure between sample i and j,

$$A_{ij} = e^{\frac{D_{ij}^2}{2\sigma^2}}, \qquad (1)$$

where D_{ij} is the (Euclidean) distance of sample i and j and σ some normalization constant. Each of the resulting clusters includes examples of a pattern of KC morphology regarding the amplitude change or power of signal over time. For each training cluster c, the cumulative histogram of all pairwise affinities is calculated, $p_c(A_i)$, as a probability measure of a sample i with affinity value A_i to be an inlier of the distribution. The higher p_c, the smaller the significance (p-value) and thus the possibility of the sample to be an outlier (false positive).