Automatic artifact detection for whole-night polysomnographic sleep recordings

D. Coppeters ‘t Wallant1, S.L. Chellappa1, G. Gaggioni1, M. Jaspar1, C. Meyer1, V. Muto1, G. Vandewalle1, P. Maquet1,2,3, C. Phillips1,4

1Cyclotron Research Centre, University of Liège, Belgium; 2Dep. of neurology, CHU University of Liège, Belgium; 3Walloon Excellence in Lifesciences and Biotechnology (WELIBIO); 4Dep. of Electrical Engineering and Computer Science, University of Liège, Belgium.

INTRODUCTION

Analyses of sleep electro-encephalographic data (EEG) have first to detect artifacts in order to reject the corresponding time points before further examination. Manual artifact detection has two main shortcomings. It is

- A very time consuming and tedious task
- A subjective procedure leading to disagreements between experts

Ideally artifact detection should be automatic, fast, reproducible and accurate.

There are currently no such method. Proposed approaches face different issues:

- Lack of specificity (detection of some but not all artifacts)
- Methodological weaknesses (e.g., need of a training set, arbitrarily fixed thresholds)
- Computational burden (huge computing time for a whole night).

See review for artifacts processing in sleep EEG in [1].

The aim of this project was to develop an automatic artifact detection method for whole-night polysomnographic sleep recordings.

METHODS

Principle

Artifacts are marked either:

- per short (1 second) epoch, over all channels
- per channel, over a "scoring window" (20s [2] or 30s [3])

Processing line

The raw data are processed through different modules (Figure 1) that are applied successively and have a specific task: "pre-processing", "bad channel detection", and "artifact detection" (Figure 2).

Figure 1: Processing pipeline

- Filtering: EEG (0.5-30Hz) – EOG (0.1-5Hz) – EMG (10-125Hz)
- Mean correction for each "scoring windows"
- Power spectrum computed in 4 frequency bands for each 1sec epoch: 60-5, 5-4Hz, o(6-12Hz), (12-16Hz) and (16-30Hz)
- Noisy and flat EEG channels removed from each "scoring window".
- Reconstruction of a good EMG channel from EMG available.

Figure 2: Modules details

Output

- list to each of bad channels per scoring window and
- list to each of artifacted episodes defined by 1 second epochs (Figure 3).

Figure 3: Bad channels (D) and bad epochs (E). In this case, D is empty, no short artifacts whereas E is composed of two bad channels: ‘O’ and ‘P’.

Threshold definition

The thresholds used in the two detection modules are directly derived from data, making the automatic method:

- robust in front of inter- and intra- subject variability
- expert independent and reproducible

RESULTS

Data

Data consist in whole night sleep multichannel EEG recording (10-20 system)

Table 1: Dataset Characteristics

<table>
<thead>
<tr>
<th>Dataset</th>
<th>#SR</th>
<th>#SE</th>
<th>Age</th>
<th>Gender</th>
<th>Parameters fixation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset 0</td>
<td>6</td>
<td>1</td>
<td>21</td>
<td>31</td>
<td>R&amp;K</td>
</tr>
<tr>
<td>Dataset I</td>
<td>35</td>
<td>2</td>
<td>21</td>
<td>31</td>
<td>R&amp;K Evaluation: Phase I</td>
</tr>
<tr>
<td>Dataset II</td>
<td>4</td>
<td>6</td>
<td>21</td>
<td>31</td>
<td>AASM Evaluation: Phase II</td>
</tr>
</tbody>
</table>

### Phase I: Robustness through assessment with 35 sleep recordings (Dataset I)

S1: Artefacts scored independently by an expert (VM) or the “automatic detection” (AD):

- AD is compared to VM (gold standard), Figure 5a (S1).
- AD detects 81.7% of the artefactepisodes considered by VM

S2: All the artefacts detected by AD but not VM have been reviewed and reassessed (false detection or oversight.) by another expert (GG), Figure 5b (S2).
- AD reached finally 91.8% sensitivity
- FDR decreased from 37.6% to 22.2%

### Phase II: Robustness in front of six different sleep experts (Dataset II)

- Gold standard created by the union of 6 experts scoring.
- Each expert scored and the AD compared to the gold standard, Figure 5b.
- AD’s artefact detection is similar to that of the best two experts with smaller standard deviation over the 4 recordings (SAD = 62.21% +/- 8.1%, SAD = 63.20% +/- 15.25% and SAD = 62.95% +/- 17.12%)
- Episode overlap is smaller as AD is more conservative (over time) than the experts

REFERENCES


ACKNOWLEDGEMENTS & SPONSORS

Cyclotron Research Centre (CRC); Belgian National Funds of Scientific Research (FNRS); Actions de Recherches Concertées (ARC, ULg) – Fondation Médicale Reine Elisabeth (FMRE); Walloon Excellence in Lifesciences and Biotechnology (WELIBIO)

CYCLOTRON RESEARCH CENTRE | http://www.cyclotron.ulg.ac.be | Contact | d.coppeters@ulg.ac.be