Heart rate analysis for sleep detection

Ronald M. Aarts

The 18th edition of the annual international clinical symposium Update@Kempenhaeghel
Friday 11 March 2016 in Heeze
Sleep monitoring
- clinical practice

- Polysomnography (PSG)
  - gold standard for sleep monitoring and diagnosis of sleep disorders
Electro Cardio Gram (ECG) measurements

Willem Einthoven (1860-1927)
Unobtrusive sleep monitoring

- scope

• Enabling sleep monitoring
  – at the *convenience* of the personal bedroom
  – unobtrusively and for *prolonged* periods of time

• Measure
  – body movement
  – cardio-respiratory activity

• To extract
  – sleep – wake
  – sleep stages
    ▪ REM – light – deep sleep
  – arousals
Benchmarking - healthcare

- ASAA+IBM Watson SleepHealth app
- S+ by ResMed
- Sleep Profiler by Advanced Brain Monitoring
- SleepImage
- WatchPAT by Itamar
Benchmarking - consumer

Basis

Zeo

UP by Jawbone

Sleep Monitor by Applied Radar Technology

FitBit
Sleep staging with ECG & Respiratory inductance plethysmography (RIP) and Body movement
From sensor to sleep information
- algorithms

Sensor options
- ECG + RIP
- Wrist-worn PPG
- (In-bed BCG)

Physiology
- body movement
- respiration interval
- respiration depth
- heartbeat interval

Sleep
- What information?
  - in/out of bed
  - sleep – wake
  - REM – Light - Deep
  - arousals
  - ...

- Performance evaluation
  - agreement with hypnogram
Unobtrusive sleep monitoring
- sensing: body movements

Moderate performance *sleep/wake* classification, insufficient for other sleep stages
Unobtrusive sleep monitoring
- sensing: respiratory effort
Unobtrusive sleep monitoring
- sensing: cardiac activity
Dataset and reference
- ECG & RIP

• SIESTA data set
  – Full PSG
  – > 100 healthy subjects, 2 nights
  – Reference hypnogram scored by sleep technicians

• Includes ECG and Respiratory Inductance Plethysmography (RIP) thorax belt
  – Used for sleep stage classification
# Sleep stage classification
- ECG+RIP performance

<table>
<thead>
<tr>
<th>Classifier</th>
<th>N</th>
<th>Avg. kappa</th>
<th>Avg. accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>3 classes: Wake / non-REM / REM</td>
<td></td>
</tr>
<tr>
<td>Literature [1]</td>
<td>31</td>
<td>0.45 ± ---</td>
<td>0.76 ± ---</td>
</tr>
<tr>
<td>Literature [2]</td>
<td>18</td>
<td>0.44 ± 0.19</td>
<td>0.79 ± 0.10</td>
</tr>
<tr>
<td><strong>Our classifier</strong></td>
<td><strong>61</strong>*</td>
<td><strong>0.66 ± 0.13</strong></td>
<td><strong>0.85 ± 0.06</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 classes: Wake / Light sleep / Deep sleep / REM</td>
<td></td>
</tr>
<tr>
<td><strong>Our classifier</strong></td>
<td><strong>61</strong>*</td>
<td><strong>0.60 ± 0.12</strong></td>
<td><strong>0.76 ± 0.07</strong></td>
</tr>
</tbody>
</table>

- Features from ECG and respiration belt, N=61 healthy subjects (42 females), first night, ages 20-86, results with 10-fold cross-validation

---

[3] Poor: .21-.40
[4] Reasonable: .41-.60
[5] Good: .61-.80
Arousal detection
- During N3

• Feature extraction
  – derived from **RR-intervals** on ECG
  – 1-second resolution

• Data set
  – **15 healthy subjects**, full PSG including ECG (lead II)
  – Annotated by (external) sleep technician, AASM rules
    ▪ Arousal length > 3 and < 15 seconds

• Performance
  – Sensitivity: 70.2%, PPV=60.5%
    ▪ Most “false detections” are shorter than 3 seconds – ground-truth limitation
Sleep staging with PPG
Sleep staging with Philips PPG sensor
- In-house PPG sensor: Wearable Sensing Technologies (WeST)

• Reflective photoplethysmography (PPG) used to accurately detect inter-beat intervals

• Can be used to extract cardiac features needed for sleep staging

• Also has an accelerometer for measuring body movements
Dataset and reference

- PPG

  - HHS data set
    - Full PSG
    - Wrist-worn PPG and accelerometer
    - 16 subjects, 2 nights
    - 40+ years old, BMI > 25
    - Reference hypnogram scored by external sleep technician
Sleep stage classification
- ECG/PPG performance

<table>
<thead>
<tr>
<th>Classifier</th>
<th># recordings (#subjects)</th>
<th>Avg. kappa</th>
<th>Avg. accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIESTA (ECG only)*</td>
<td>135 (101)</td>
<td>0.53</td>
<td>0.68</td>
</tr>
<tr>
<td>HHS (PPG)**</td>
<td>26 (17)</td>
<td>0.50</td>
<td>0.65</td>
</tr>
</tbody>
</table>

No respiratory information used, only cardiac inter-beat interval features
*Results with 10-fold cross-validation
**Trained with SIESTA, validated on HHS (hold-out validation set)

.21-.40 poor
.41-.60 reasonable
.61-.80 good
.81-.99 excellent
Sleep stage classification

Next steps

- Explore use of additional information in morphology of PPG signal
  - Influence of respiration in intensity, pulse width and pulse rate
  - Influence of sudden increasing distal skin temperature, e.g. during sleep onset

Figure 2. Respiratory-induced intensity variations (RIIVs) in the baseline of the photoplethysmographic (PPG) signal. a.u. = arbitrary units.

Derived respiration signal (continuous line) and amplitude-scaled reference (dashed line), using pulse-rate variability.
Fig. 1. Typical examples of respiratory time series (a) during sleep and (b) during wake in a period of 1 min, and respiratory PSD series (c) during sleep and (d) during wake.

Long et al, Sleep and Wake Classification With Actigraphy and Respiratory Effort Using Dynamic Warping, IEEE JOF BIOMEDICAL AND HEALTH INFORMATICS, VOL. 18, NO. 4, JULY 2014
Sleep/wake: dynamic warping on respiratory effort

• Dynamic warping (DW)
  
  – A robust measure of similarity in terms of signal waveform that allows signal *shifting* and *offset*

<table>
<thead>
<tr>
<th>Results</th>
<th>Features</th>
<th>Accuracy</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sleep/wake classification</td>
<td>Existing features</td>
<td>91.5%</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>DW features</td>
<td>94.3%</td>
<td>0.59</td>
</tr>
</tbody>
</table>

Sleep/wake: spectral boundary adaptation on HRV

• Time-frequency analysis
  – Traditional boundaries
    ▪ VLF (0.003 – 0.04 Hz)
    ▪ LF (0.04 – 0.15 Hz), sympathetic
    ▪ HF (0.15 – 0.4 Hz), parasympathetic
    ▪ LF/HF ratio, sympathetic-parasympathetic balance
  – Adaptive boundaries
    ▪ Better express autonomic activity
    ▪ Adapted by LF and HF peak frequencies.

<table>
<thead>
<tr>
<th>Results</th>
<th>Boundaries</th>
<th>Accuracy</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sleep/wake classification</td>
<td>Traditional</td>
<td>89.9%</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>Adaptive</td>
<td>93.1%</td>
<td>0.48</td>
</tr>
</tbody>
</table>

Sleep staging: respiratory amplitude

- Respiratory “depth” and “volume” features

<table>
<thead>
<tr>
<th>Results</th>
<th>Resp. features</th>
<th>Accuracy</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wake/REM/light/deep classification</td>
<td>Existing</td>
<td>61.7%</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td><strong>Existing + Amplitude</strong></td>
<td><strong>63.8%</strong></td>
<td><strong>0.38</strong></td>
</tr>
<tr>
<td>Deep sleep detection</td>
<td>Existing</td>
<td>84.9%</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td><strong>Existing + Amplitude</strong></td>
<td><strong>86.1%</strong></td>
<td><strong>0.43</strong></td>
</tr>
</tbody>
</table>

Sleep staging: respiratory self-similarity

- Respiratory waveform similarity
  - Subseries with a certain number of breaths
  - Dissimilarity is the **uniform scaling** distance: *min-Euclidean distance between two subseries*

### Results

<table>
<thead>
<tr>
<th></th>
<th>Resp. features</th>
<th>Accuracy</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wake, REM, light, deep</td>
<td>Existing</td>
<td>63.8%</td>
<td>0.38</td>
</tr>
<tr>
<td>Existing + New</td>
<td><strong>64.9%</strong></td>
<td><strong>0.41</strong></td>
<td></td>
</tr>
</tbody>
</table>

Cardiac activity precedes brain activity during sleep transitions except for REM transitions

- Parameters
  - Brain: EEG mean frequency
  - Cardiac: HR, SDNN, LF, HF.

Overnight sleep

HRV seem changing earlier than EEG

Important: some sleep stages can be predicted just a few minutes' earlier!!!
Deep sleep detection with time-delayed features

- **Features**
  - HRV: SDNN, LF, DFA
  - Respiratory: frequency variation, breathing depth

### Results

<table>
<thead>
<tr>
<th></th>
<th>Time-delay</th>
<th>Accuracy</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deep sleep detection</td>
<td>No</td>
<td>87.1%</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>Yes (5 min)</td>
<td>88.7%</td>
<td>0.57</td>
</tr>
</tbody>
</table>
Cardiorespiratory interaction in complex networks

Cardiorespiratory interaction (CRI)

Visibility graph in complex network

Sleep Apnea Detection Using Time-Delayed Heart Rate Variability
Obstructive Sleep Apnea

• Involuntary cessation of breathing that occurs during sleep

• Estimates of disease prevalence are in the range of 3% to 7%

• Strongest risk factor is obesity reflected by several markers including BMI, neck circumference, and waist-to-hip ratio

• Untreated obstructive sleep apnea is associated with:
  ✓ an increased risk of fatal and nonfatal cardiovascular event
  ✓ a higher propensity of sudden death during sleep
  ✓ a greater risk for stroke
  ✓ all-cause mortality
Hysteretic changes in cardiac activity
Method

• Feature Extraction

✓ HRV features were extracted from RRIs for each 30-sec epoch
✓ Time-domain features chosen due to low-computation complexity
✓ HRV features with a certain time delay to perform SA detection

• Classification

SVM classifier
  ➢ the radial basis function (RBF) as a kernel function

Subject-Specific approach
  ➢ using leave-one-out cross-validation

Subject Independent approach
  ➢ Using 10-fold cross-validation
Results
Results

**TABLE I. SUMMARY OF SA DETECTION PERFORMANCE**

<table>
<thead>
<tr>
<th>Metric</th>
<th>Subject-independent</th>
<th>Subject-specific</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Without time delay</td>
<td>With time delay</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>74.9±0.089</td>
<td>76.2±0.078</td>
</tr>
<tr>
<td>Kappa</td>
<td>0.49±0.205</td>
<td>0.52±0.197</td>
</tr>
</tbody>
</table>
Overall Summary

• Assets:
  – Algorithms for 4-class sleep staging
  – Sensor agnostic: algorithms can be used with different sensors

• Benchmarking:
  – ECG/RIP: classification performance exceeds reported in scientific literature
  – PPG: comparable to ECG

  – Using the time-delayed features improved the SA detection results
Acknowledgment

- Pedro Fonseca
- Reinder Haakma
- Xi Long
- Mustafa Radha
  all of Philips Research, and
  Marina Nano (TU/e)

- All mentioned papers (incl. Xi Long’s PhD thesis) available at: http://www.extra.research.philips.com/hera/people/aarts/
Thank you!