Hypnogram and Sleep Parameter Computation From Activity and Cardiovascular Data

Alexandre Domingues*, Teresa Paiva, and J. Miguel Sanches, *Senior Member, IEEE*

***Abstract*—The automatic computation of the hypnogram and sleep Parameters, from the data acquired with portable sensors, is a challenging problem with important clinical applications. In this paper, the hypnogram, the sleep efficiency (SE), rapid eye movement (REM), and nonREM (NREM) sleep percentages are automatically estimated from physiological (ECG and respiration) and behavioral (Actigraphy) nocturnal data. Two methods are described; the first deals with the problem of the hypnogram estimation and the second is specifically designed to compute the sleep parameters, outperforming the traditional estimation approach based on the hypnogram. Using an extended set of features the first method achieves an accuracy of 72.8%, 77.4%, and 80.3% in the detection of wakefulness, REM, and NREM states, respectively, and the second an estimation error of 4.3%, 9.8%, and 5.4% for the SE, REM, and NREM percentages, respectively.**

***Index Terms*—Hypnogram estimation, rapid eye movement (REM)/nonREM (NREM) percentage, sleep efficiency (SE), sleep parameters.**

I. Introduction

SLEEP disorders form a class of medical problems generally characterized by changes of physiological or behavioral sleep patterns [1]. Their impact on both young and adult populations is well documented [2] and can be related with a wide range of short- and long-term consequences for the health of the subjects, including anxiety, memory and cognitive impairments, high blood pressure, obesity, and psychiatric problems, among others.

The golden standard for the diagnosis of sleep disorders is the polysomnography (PSG) [3], which is by far the most reliable and accurate method. Several important measures for diagnosis are derived from PSG data, such as the hypnogram, a graphical representation of sleep stages (*wakefulness*, *rapid eyemovement* (REM) sleep, and the three nonREM sleep stages) as a function of time. This graphical representation is useful to analyze the sleep cycle and several parameters are usually computed to quantify and characterize sleep, such as the *sleep efficiency* (SE), *sleep onset latency*, *REM sleep percentage* ($\mathrm{REM_p}$), *NonREM sleep percentage* ($\mathrm{NREM_p}$), and *REM latency*.

However, PSG involves complex acquisition devices and long setup procedures. It is uncomfortable to the subject and is usually done in clinical facilities. These highly constrained conditions prevent its use in a nonintrusive way in normal daily life and limits the duration of the typical exam, which is usually performed over one or two nights.

Due to these constraints, simpler alternatives have been suggested to complement the information given by the PSG. Sleep and dream diaries [4], sleep questionnaires, and in particular, *actigraphy* (ACT) [5], are very efficient acquiring behavioral data over long periods, often revealing abnormal trends in the subject's behavior. An extensive review of the application of ACT in the scope of sleep disorders is presented in [6]. The authors conclude that, although ACT has a reasonable validity and reliability in individuals with normal sleep patterns, its validity in patients with poor sleep is more questionable, thus motivating the combination of ACT with other sources of data.

The advent of small portable devices with high storage and processing capabilities have allowed physiological and behavioral data to be acquired, outside clinical environments, in a reliable way, often across several days. This data includes ECG, *respiratory inductance plethysmography* (RIP), oxygen saturation, among others.

Sleep patterns are known to be intimately connected with the activity of the autonomous nervous system (ANS) [7]. This activity can be indirectly estimated from several physiological signals, such as the *heart rate variability* (HRV), RIP, the *peripheral arterial tone*, and the *galvanic skin response*.

HRV, extracted from ECG data, has received particular attention by the researchers and medical community, particularly after the the publication of the *standards of measurement, physiological interpretation, and clinical use of HRV* [8]. The HRV reflects the complex balance between the two branches of the ANS, the parasympathetic and sympathetic pathways.

Several studies establish the correlation between sleep stages and HRV [9], [10]. They show that REM sleep is associated with an increased sympathetic activity and nonrapid eye movement (NREM) sleep with a predominance of parasympathetic output. In [11], the authors present a brief retrospective of the study of HRV in the scope of sleep studies and show detailed data on the variation of the HRV across the different sleep stages. Recent studies focus on nocturnal HRV under confinement conditions [12], and in [13], an in depth review of the relationship between HRV and several sleep disorders is presented.

The respiration process is controlled by a cyclic stimulation of the diaphragm mediated by the phrenic nerve, which contains

Manuscript received September 5, 2013; revised November 10, 2013; accepted January 8, 2014. Date of publication January 20, 2014; date of current version May 15, 2014. *Asterisk indicates corresponding author.*

*A. Domingues is with the Institute for Systems and Robotics/Bioengineering Department—Instituto Superior Técnico/Technical University of Lisbon, 1649-004 Lisbon, Portugal (e-mail: adomingues@gmail.com).

T. Paiva is with the Centro de Electroencefalografia e Neurologia Clinica (CENC)/Faculdade de Medicina da Universidade de Lisboa (FMUL), 1649-028 Lisbon, Portugal (e-mail: teresapaiva0@gmail.com).

J. M. Sanches is with the Institute for Systems and Robotics/Bioengineering Department—Instituto Superior Técnico/Technical University of Lisbon, 1649-004 Lisbon, Portugal (e-mail: jmrs@ist.utl.pt).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2014.2301462

motor, sensory, and sympathetic nerve fibers. The involuntary breathing process is thus a direct reflection of the activity of the ANS.

In [14], the authors show that respiration is more irregular during REM states when compared to nonREM, and in [15], the authors show that different sleep stages lead to distinct autonomic regulation of breathing.

The automatic extraction of useful indicators for sleep disorders diagnosis, using data acquired in mobile environments, is still an open issue that poses many challenges. The problems to solve have different degrees of complexity and include the accurate estimation of sleep and wakefulness periods, detection of REM and NREM sleep, and the automatic computation of *sleep parameters*. Several approaches have been proposed to address these issues. Spectral analysis of the HRV plays a major role in many publications, where the frequency bands described in [8] have become the standard for HRV spectrum analysis.

The low frequencies (LF—[0.015 – 0.15] Hz) are thought to reflect the balance between the activity of the two branches of the ANS and the high frequencies (HF—[0.15 – 0.4]Hz) the activity of the parasympathetic branch, highly modulated by the breathing pattern [16].

Some authors have proposed variations to these standards with relevant results. In [17], the authors propose an algorithm that adaptively extracts features from HRV for sleep and wake classification. They show that the adaptive frequency bands improve the discriminative power of the frequency based features. In [18] and [19], the authors propose the use of *time-variant autoregressive models* (TVAM) to extract the spectral features, they show that using TVAM the algorithm becomes more sensitive to fast variations in the sleep state.

An accurate detection of sleep stages across the entire night has been presented by some authors. In [20], the authors present an algorithm, optimized for sleep-disordered breathing patients, which discriminates sleep stages based on a set of heuristic rules and a threshold based discriminative function. In [21] and [22], a hidden Markov model (HMM) classifier based on features extracted from TVAM is presented to discriminate Wake–REM–NREM and REM–NREM, respectively. A recent study by Willemen *et al.* [23] evaluates the discriminative capacity of a large set of cardio–respiratory and movement features in three classification tasks: Sleep–Wake, REM–NREM,and light-deep sleep achieving high agreement rates.

In [24], the authors present an interesting approach, also adopted in this paper, where parts of the data, ambiguous from a classification point of view, are discarded in order to improve the final estimation of the sleep and wakefulness periods.

The estimation of sleep parameters/stages from multimodal data is presented in some papers with promising results. In [25], the authors combine ACT and cardio–respiratory signals to achieve high accuracies in sleep and wakefulness detection, although no proper validation data, (i.e., the hypnogram from the PSG), is used. In [26] and [14], the authors present a sleep staging algorithm that combines HRV and RIP, and explore the influence of obstructive sleep apnea (OSA) in the performance of the algorithm.

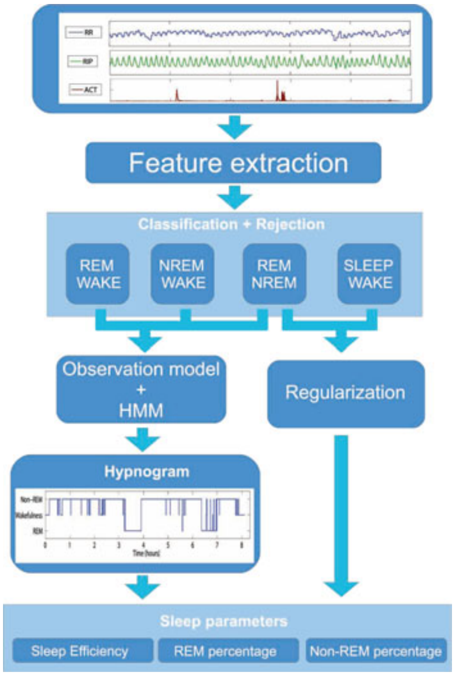

Fig. 1. Fluxogram of the proposed method. The hypnogram and sleep parameters are initially estimated from the output of three binary classifiers, fed to a HMM based algorithm (left). An alternative method for the estimation of the sleep parameters is also described based on the output of two binary classifiers and a regularization operation (right).

In [27], the authors combine ACT, respiratory effort, and HRV obtaining a high accuracy but relatively low sensitivity in the discrimination between sleep and wakefulness.

This paper deals with the problem of automatically estimating a simplified hypnogram (wakefulness, REM, and NREM) and three standard sleep parameters: 1) SE, 2) REM$_p$, and 3) NREM$_p$ from data easily acquired with portable sensors.

The sleep parameters are estimated using two different methods: first, the Hypnogram is estimated from the data and the sleep parameters computed. Then, an alternative method is described that eliminates the need for a hypnogram by combining the rejection of ambiguous samples and a regularization operation.

The two methods rely on an extended set of features, extracted from HRV, RIP, and ACT, and an ensemble of classifiers that include a rejection option.

II. Methods

In this section, the multimodal data is presented, followed by the description of the algorithm to estimate the Hypnogram and compute the three sleep parameters.

The complete estimation method, displayed in Fig. 1, is composed by the preprocessing and feature extraction procedures, presented in Sections II-B and II-C, respectively, followed by the classification stage (see Section II-D) designed to reject the ambiguous features.

The output of the set of classifiers is then used as the input for a HMM, as described in Section II-E and shown in Fig. 1, to estimate the Hypnogram, followed by the computation of the sleep parameters. An alternative method to compute the sleep parameters is finally described in Section II-F.

The performance of all the described methods is assessed with several *figures of merit* (FOM). These FOMs are computed in a *leave-one-patient-out* cross-validation basis, where each patient dataset is tested after training the algorithm (i.e., the classifiers and the HMM model) with the remaining data.

Besides the positive detection rate and global accuracy (Acc), the *Gmean*[1] [28] is also computed. This is motivated by the highly unbalanced nature of the classification tasks at hand, e.g., in the sleep versus wakefulness discrimination problem, up to 95% of the samples belong to the sleep class, in REM versus NREM typically around 80% of the samples belong to NREM class. The Gmean gives a global insight into the performance of the method, which is often masked in the Acc by the bias introduced by predominant classes. The Cohen kappa index[2] [29] is also computed, for performance comparison, when necessary.

A. Data

Each subject performed one standard nocturnal PSG exam at a sleep laboratory. The PSG data was jointly acquired with ACT using a *Somnowatch*™device, from Somnomedics, placed in the nondominant wrist of the subjects, acquiring with a sampling rate of 1 Hz. The core of these devices is a 3-D accelerometer that measures the acceleration along three orthogonal axes with a configurable output format. Here, the output of the ACT is the acceleration magnitude.

The hypnogram, obtained from the PSG by trained technicians, is used as a ground truth to identify *REM sleep*, *NREM sleep,* and *wakefulness* in epochs of 30 s.

Twenty adult subjects (age 42.1 ± 9 years, 12 Males, 8 Females), with no prediagnosed sleep disorders, participated in this study.

The SE was computed from the hypnogram for every patient, ranging from 75% to 95% with an average value of 86.1 ± 5.2%. This value is usually above 85% [30] in healthy patients, this suggests the occurrence of sleep disturbances in some of the subjects, although not necessarily pathological.

B. Preprocessing

Preprocessing operations are required to reduce the movement artifacts, normalize the data across different patients, and prepare it for feature extraction.

ECG filtering and QRS complex detection is performed according to the methods described in [31], the RR signal [8] is then constructed from the detected R peaks and downsampled to 2 Hz. The downsampling operation consists in an antialiasing filtering, using a 8th order Chebyshev low-pass filter, with 0.8-Hz cutoff frequency, followed by decimation. The 2-Hz sampling frequency is within the accepted range, as shown in [31], being above the Nyquist frequency for the considered frequency ranges.

Magnitude normalization and dc component removal are applied to both the RIP and ACT signals in a sliding window basis as follows:

$$\tilde{a}(n) = \frac{a(n) - \mu(n)}{\sigma(n)} \tag{1}$$

where $a(n)$ is the original sample, $\mu(n)$, and $\sigma(n)$ are the mean and standard deviation of the data within the 5-min window centred at the nth sample and $\tilde{a}(n)$ is the normalized sample.

C. Feature Extraction

This paper combines features extracted from the RR, RIP, and ACT signals and one synchronization measure between the RR and RIP.

After preprocessing, each dataset is divided in contiguous epochs of $T = 30$ s, synchronized with the ground-truth hypnogram provided by the medical staff. All the epochs corresponding to any of the three distinct NREM sleep stages were grouped into one single label.

Let $w_j = \{\mathrm{RR}_j\ \mathrm{RIP}_j\ \mathrm{ACT}_j\}$ represent a T dimensional window, containing the multimodal data from the jth epoch, where $j \in [1, \ldots, M]$ and M the total number of epochs. The extracted features and the extraction procedures are the following.

1) RR Features: The RR frequency domain features are computed, according to the guidelines from [8], in the LF and HF bands. In order to extract these features from each RR_j, an eight-order autoregressive model (AR) [32] is fitted to the extended window $\mathrm{RR}_j^* = [\mathrm{RR}_{j-3} \ldots \mathrm{RR}_j]$ and a set of optimal coefficients $\hat{\boldsymbol{a}}_{\mathrm{RR}}$ and a residual, E_{RR}, are obtained. The length of RR_j^*, 2 m, follows the standards set by [8], allowing to capture the low-frequency components of the RR signal. The power spectrum is computed from the estimated AR coefficients and the following features are extracted:

1) PM_j: Magnitude of the high-frequency pole of the filter *impulsive response filter* (IIR) described by the coefficients $\hat{\boldsymbol{a}}_{\mathrm{RR}}$.
2) PP_j: Phase of the high-frequency pole.
3) $\mathrm{E}_{\mathrm{RRj}}$: Residual of the AR model fitted to RR_j^*.
4) TP_j: Total power (LF+HF).
5) HF_j: Power in the HF range.
6) LF_j: Power on the LF range.
7) $\mathrm{LF/HF}_j$: Power ratio between the two frequency bands.
8) MHR_j: Mean heart rate on the considered RR_j.

2) RIP Features: The RIP related features are extracted by estimating the optimal parameters $\hat{\boldsymbol{a}}_{\mathrm{BR}}$ of the four-order AR model fitted to each RIP_j, and computing:

1) BV_j: Magnitude of the high-frequency pole of the filter, describing the variance in the breathing rate.
2) BPM_j: Phase of the high-frequency pole, reflecting the average breathing rate.

3) RR + RIP Features: The temporal interplay between oscillations of heartbeat and respiration, reflect information related to the cardiovascular and autonomic nervous system [33].

[1] Given the positive detection rates ($R_{1\ldots N}$) of a N class problem, the Gmean is given by $\prod_{i=1}^{N} R_i^{\frac{1}{N}}$.

[2] Values of kappa can range from −1.0 to 1.0, with −1.0 indicating perfect disagreement below chance, 0.0 indicating agreement equal to chance, and 1.0 indicating perfect agreement above chance.

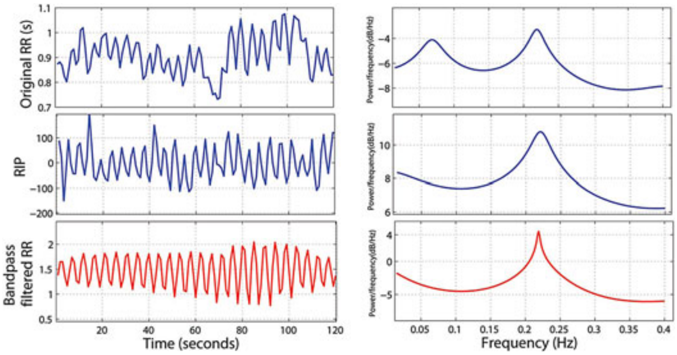

Fig. 2. Two minute window of the RR signal (top left) and the respective power spectrum (top right) showing two peaks centered in the LF and HF bands. The breathing signal (middle left) has its frequency response centred (middle right) in the breathing frequency. This response is used to bandpass filter the RR signal, resulting in the signal and power spectrum displayed in the bottom left and right, respectively.

Let $\mathrm{RR}_j^{\mathrm{Br}}$ denote the breathing component of RR_j. The phase synchronization between RIP_j and $\mathrm{RR}_j^{\mathrm{Br}}$ is quantified computing the *Phase-Locking Factor* (PLF) [34] given by

$$\theta_j = \left| \frac{1}{T} \sum_{n=1}^{T} e^{i(\phi_{\mathrm{RIP}}[n] - \phi_{\mathrm{RR}}[n])} \right| \tag{2}$$

where ϕ_{RIP} and ϕ_{RR} are the instantaneous phases of RIP_j and $\mathrm{RR}_j^{\mathrm{Br}}$, respectively, computed using the Hilbert transform. The value of θ_j is a measurement of the synchronization between the two oscillators, with 1 corresponding to perfect synchronization and 0 to no correlation between the phases.

The breathing component $\mathrm{RR}_j^{\mathrm{Br}}$ is obtained filtering RR_j with the bandpass IIR filter described by the set of optimal coefficients $\hat{\boldsymbol{a}}_{\mathrm{BR}}$. To compensate the nonlinearity of the phase, the signal is filtered in both forward and backward direction. Fig. 2 illustrates the steps in the computation of θ_j.

4) ACT Features: The features extracted from ACT are based on the work described in [35]. Each set of features is computed from a 3.5-min window $\mathrm{ACT}_j^* = [\mathrm{ACT}_{j-3} \ldots \mathrm{ACT}_{j+3}]$, centered on the jth epoch. The following features are extracted:

1) AR—Coefficients ($a_{\{1,\ldots,4\},j}$) and residue (E_{ARj}) of a four-order AR model fitted to ACT_j^*.
2) RMM—Weights ($w_{\{1,\ldots,3\},j}$), parameters ($r_{\{1,\ldots,3\},j}$), and the Kullback–Leibler (KL_j) divergence of the Rayleigh mixture model (RMM) [36] distribution fitted to ACT_j^*.
3) Mag_j—The energy of $x(k) = \mathrm{ACT}_j^*$ given by $\sum_k h(k)x(k)^2$, where $\mathbf{h} = \{h(k)\}$ is a Hanning window.

In order to minimize the interpatient variability, a normalization operation was performed. Let $\boldsymbol{f}_{ij}$ denote the vector containing all the samples from feature i and subject j, the normalization is performed according to,

$$\tilde{f}_{ij}(n) = \frac{1}{1 + e^{-\frac{f_{ij}(n) - \mu_{ij}}{\sigma_{ij}}}} \tag{3}$$

where $\tilde{f}_{ij}(n)$ is the nth normalized sample and μ_{ij} and σ_{ij} the mean value and standard deviation of $\boldsymbol{f}_{ij}$, respectively. This normalization step ensures that all features fall in the range $[0, \ldots, 1]$.

The discriminative power of each feature was computed using the *Mahalanobis Distance* (MD) [37] and the statistical significance was assessed performing a one-way ANOVA test. Table I shows the MD and the result of the p-value test (for a significance level of 0.05) obtained for each feature in four different tasks: 1) *REM versus wakefulness*, 2) *REM versus NREM*, 3) *NREM versus wakefulness,* and 4) *sleep versus wakefulness* discrimination.

D. Classification and Feature Selection

The discrimination between the considered classes, wakefulness, REM, and NREM, falls within a common multiclass classification problem. Several approaches are possible to solve this kind of problem, they include: 1) the design of a All-versus-All classifier, where each sample is classified into one of the three possible classes, 2) a hierarchical classifier, with an initial classification separating wakefulness and sleep and a second classifier discriminating the former class into REM and NREM states, and 3) a combiner classifier composed by three One-versus-All classifiers and the final score given by a specific combining rule.

The main limitation of approach 1) is that the same group of features is used to discriminate the three different classes. From Table I, it is clear that distinct features are optimal to discriminate different classes thus motivating the use of binary classifiers.

The hierarchical classification approach, 2), enables the use of distinct features to discriminate between different classes, but the classification error in sleep/wakefulness propagates into the second stage, REM/NREM. The results using this approach are presented in Section III for comparison purposes.

The solution adopted in this paper, is an extension of approach 3). The estimation of the hypnogram, described in Section II-E, is based on three binary classifiers that independently classify all the samples into 1) REM/wakefulness (RW), 2) REM/NREM (RN), and 3) NREM/wakefulness (NW).

The estimation of the sleep parameters, described in Section II-F, uses a fourth binary classifier that maps all samples into sleep and wakefulness (SW) classes.

Each classifier is designed to take into account a rejection factor (RF), rejecting a specified percentage of samples, whose classification is ambiguous. In large biomedical datasets, such as the one considered, the systematic rejection of unreliable segments and/or samples has been shown to increase the accuracy of the classification procedures without compromising the overall result [24]. The rejection works by computing the true or estimate posterior probability of the winning class for each sample and rejecting those which are below the specified percentage.

Therefore, each classifier maps each sample into one of three labels: $\mathrm{RW} \in \{rs, wk, r\}$, $\mathrm{RN} \in \{rs, ns, r\}$, $\mathrm{NW} \in \{ns, wk, r\}$, and $\mathrm{SW} \in \{sl, wk, r\}$, where rs, ns, wk, sl, and r refer to *REM*, *NREM*, *wakefulness*, *sleep,* and *rejected sample*, respectively.

TABLE I
MD AND THE RESULT OF THE SIGNIFICANCE TEST, ($p = 1$ MEANS THAT THE NULL HYPOTHESIS IS REJECTED) FOR ALL THE EXTRACTED FEATURES ON THREE BINARY CLASSIFICATION TASKS: I) RW—*[REM / WAKEFULNESS]*, II) RN—*[REM/NREM]*, AND III) NW—*[NREM/WAKEFULNESS]*

	MD	p	MD	p	MD	p	MD	p	MD	p	MD	p
	PM		**PP**		**E**		**TP**		**HF**		**LF**	
RW	**0.53**	1	0.02	0	**0.37**	1	0.01	0	0.39	1	0.16	1
RN	**1.69**	1	**0.63**	1	**0.17**	1	**0.28**	1	**1.75**	1	**1.09**	1
NW	**0.51**	1	**0.43**	1	**1.48**	1	**0.21**	1	**0.67**	1	**0.47**	1
SW	0.24	1	0.21	1	**1.13**	1	0.14	1	**0.35**	1	0.26	1
	LF/HF		**MHR**		**BV**		**BPM**		**PLF**		$\mathbf{a_1}$	
RW	0.36	1	**0.38**	1	0.01	0	0.04	0	0.19	1	**2.6**	1
RN	**1.62**	1	**0.32**	1	**0.61**	1	**0.72**	1	**1.19**	1	**0.2**	1
NW	**0.65**	1	**1.9**	1	**0.42**	1	**0.32**	1	**0.47**	1	**2.7**	1
SW	**0.35**	1	**1.42**	1	0.23	1	**0.14**	1	**0.23**	1	**2.68**	1
	$\mathbf{a_2}$		$\mathbf{a_3}$		$\mathbf{a_4}$		$\mathbf{E_{AR}}$		$\mathbf{w_1}$		$\mathbf{w_2}$	
RW	**3.8**	1	**1.0**	1	0.62	1	**3.5**	1	**0.52**	1	0.28	1
RN	0.01	0	0.001	0	0.03	1	0.03	1	**0.02**	1	0.01	0
NW	**3.5**	1	**1.34**	1	0.4	1	**4.6**	1	0.9	1	0.4	1
SW	**3.7**	1	**1.41**	1	0.53	1	5.2	1	0.89	1	0.46	1
	$\mathbf{w_3}$		$\mathbf{r_1}$		$\mathbf{r_2}$		$\mathbf{r_3}$		**KL**		**Mag**	
RW	0.75	1	1.05	1	0.74	1	1.4	1	**1.1**	1	**2.5**	1
RN	0.02	1	0	0	0.03	1	0	0	0.08	1	0.01	0
NW	0.63	1	**2.1**	1	**0.5**	1	**3.4**	1	**1.89**	1	**4.60**	1
SW	0.69	1	2.27	1	0.58	1	3.5	1	**1.75**	1	**5.08**	1

Bold md values mark features selected during the feature selection procedure.

During the training step, the RW, RN, and NW binary classifiers are trained only with data from the two considered classes, respectively. However, during the test, they map samples belonging to three classes. Any sample from a class not predicted by the classifier will either be misclassified or rejected.

The four binary classification tasks were tested with several different classifiers, the support vector classifier with a second-degree polynomial kernel yielded the highest accuracy for all tasks with the exception of the REM versus NREM classification, where the Parzen classifier performed better. All the described classification tasks were implemented using PRtools [38].

Feature selection was performed for each classifier, considering only the statistically significant features and a floating feature selection algorithm [39], without any constraint on the number of selected features. The evaluation criteria is the accuracy of the classifier, used on each classification task. The selected features are displayed in Table I with the respective MD marked in bold.

E. Hypnogram and Sleep Parameter Estimation

The hypnogram estimation is based on a HMM, with three hidden states $x \in \{w, \text{REM}, \text{NREM}\}$. The HMM combines the output of the three binary classifiers (RW, NW, RN) producing a final estimate of the hypnogram. The HMM was chosen as the combiner of the three classifiers due to its ability to incorporate the information regarding the rejected samples in the observation model. Furthermore, HMMs are particularly useful in this kind of problem since they are able to model the temporal correlation between states which is the case on the sleep cycle dynamics.

Let us consider $\boldsymbol{O}$ a $N \times 3$ observation matrix, where each row/observation $o_n = [RW_n, RN_n, NW_n]$ contains the output of the three binary classifiers for the nth sample. The observation space is thus composed by 27 different possible observations.

The wmission matrix (EM) is a 3×27 matrix, representing the probabilities $P(o_{1\ldots27}|x_{1\ldots3})$, and is computed from the relative frequencies observed in the training data.

The transition matrix (TM) is a 3 × 3 matrix with the state transition probabilities, expressed as $P(x_n|x_{n-1})$. The TM is also computed from the relative frequencies observed in the training data.

The hidden state, $x(n)$, is estimated along the time from the observations, $\boldsymbol{o}(n)$, and the model parameters, EM and TM. The initial probabilities of *REM* and *NREM* sleep are set to 0 and the initial probability of *wakefulness* is set to 1 since all patients were awake in the beginning of the exam. The optimal solution, the most probable state sequence, is computed using the *Viterbi Algorithm* [40].

The three considered sleep parameters are computed from the estimated hypnogram as

$$\text{SE} = \frac{N(s)}{N(s)+N(w)} \tag{4}$$

$$\text{NREM}p = \frac{N(nr)}{N(nr)+N(r)} \tag{5}$$

$$\text{REM}p = \frac{N(r)}{N(r)+N(nr)} \tag{6}$$

where $N(.)$ is a counting operator for s, w, r, and nr, corresponding to *sleep*, *wakefulness*, *REM,* and *NREM* epochs, respectively.

F. Alternative Sleep Parameter Estimation

The estimation of the sleep parameters, as described in the previous section, follows the standard procedure, where the computation is performed directly from the hypnogram. The error associated with the estimated hypnogram will thus be directly reflected in the estimated parameters.

In this section, an alternative method is proposed that computes the sleep parameters directly from the output of the SW and RN classifiers. This method improves the accuracy of the estimated sleep parameters by taking into account 1) the higher accuracy of the binary classifiers, compared to the full hypnogram estimation, 2) a correction factor, computed in the training step, that takes into account the percentage of misclassified samples, and 3) an estimation of the number of samples that were rejected on each class.

Let us consider a binary classifier C, with a reject option, which maps each sample into one of three labels $l \in \{p, n, r\}$ where p, n, and r denote positive, negative, and reject.

The confusion matrix[3] is represented as

$$\boldsymbol{A} = \begin{bmatrix} \text{Tp} & \text{Fn} & \text{Rp} \\ \text{Fp} & \text{Tn} & \text{Rn} \end{bmatrix} \tag{7}$$

where Tp, Fn, Fp, Tn, Rp, and Rn are the *true positives*, *false negatives*, *false positives*, *true negatives*, *rejected positives*, and *rejected negatives*, respectively.

TABLE II
PERFORMANCE OF THE HYPNOGRAM ESTIMATION ALGORITHM

	Wake.(%)	REM(%)	Non-REM(%)	Acc.(%)	Gmean
0%	70.8	76.7	78.6	76.7	75.3
5%	70.0	**79.6**	76.0	75.4	75.0
10%	**72.8**	77.4	**80.3**	**78.3**	**76.8**
20%	68.5	73.4	78.7	76.0	73.4

Detection rates for the three considered states with different percentages of rejection.

TABLE III
HIERARCHICAL CLASSIFICATION

	Wake. %	REM %	NREM %	Acc. %	Gmean
Wake/Sleep	71.1	81.5		80.2	76.1
REM/NREM	-	80.4	79.5	79.7	79.9
Wake/REM/NREM	71.1	58.2	67.9	66.4	65.5

The positive ($\theta_{p,i}$) and negative ($\theta_{n,i}$) correction factors and the fraction of rejected samples per class ($\omega_{p,i}$ and $\omega_{n,i}$) are computed for each training dataset as

$$\theta_{p,i} = \frac{\text{Tp}_i + \text{Fp}_i}{\text{Tp}_i + \text{Fn}_i} \tag{8}$$

$$\theta_{n,i} = \frac{\text{Fn}_i + \text{Tn}_i}{\text{Fp}_i + \text{Tn}_i} \tag{9}$$

$$\omega_{p,i} = \frac{\text{Rp}_i}{\text{Rp}_i + \text{Rn}_i} \tag{10}$$

$$\omega_{n,i} = \frac{\text{Rn}_i}{\text{Rp}_i + \text{Rn}_i} \tag{11}$$

with $i \in [1, \ldots, M]$, and M as the number of training datasets. The final values are obtained averaging over $\boldsymbol{\theta}_{\{p,n\}}$ and $\boldsymbol{\omega}_{\{p,n\}}$.

The countingoperation can thus be improved by correcting the number of predicted samples in each class as

$$N(\hat{p}) = \frac{N(p)}{\theta_p} \tag{12}$$

$$N(\hat{n}) = \frac{N(n)}{\theta_n} \tag{13}$$

and estimating the number of rejected samples from each class as

$$N(r_p) = \omega_p N(r) \tag{14}$$

$$N(r_n) = \omega_n N(r). \tag{15}$$

The expressions for the three sleep parameters can now be rewritten as

$$\text{SE} = \frac{N(\hat{s})}{N(s)+N(w)} \tag{16}$$

$$\text{NREM}_p = \frac{N(\widehat{ns}) + N(r_{ns})}{(N(ns)+N(rs)+N(r)) \times \text{SE}} \tag{17}$$

$$\text{REM}_p = \frac{N(\widehat{rs}) + N(r_{rs})}{(N(ns)+N(rs)+N(r)) \times \ \text{SE}} \tag{18}$$

[3] The positive detection rate is computed as $\frac{\text{Tp}}{\text{Tp+Fn}}$ and the global accuracy as $\frac{\text{Tp+Tn}}{\text{Tp+Fn+Fp+Tn}}$

TABLE IV
AVERAGE OF THE ESTIMATED PARAMETERS (EST.) AND ESTIMATION ERROR (ERR.) OF THE SLEEP PARAMETERS, COMPUTED USING THE ESTIMATED HYPNOGRAM AND THE ALTERNATIVE METHOD FOR SEVERAL VALUES OF THE RF

	SE		REM_p		NREM_p	
	T.V. - 86.1 ± 5.2%		T.V. - 17.4 ± 2.9%		T.V. - 82.5 ± 3.1%	
RF	Est.%	Err.%	Est.%	Err.%	Est.%	Err.%
	Sleep Parameters from estimated Hypnogram					
0%	78.1 ± 7.2	10.9 ± 5.8	21.7 ± 4.7	27.3 ± 21.7	71.8 ± 5.6	17.3 ± 9.2
5%	78.4.4 ± 7.2	11.2 ± 5.3	22.8 ± 4.8	32.7 ± 22.4	71.1 ± 9.1	18.6 ± 9.9
10%	78.1 ± 7.5	10.6 ± 5.1	20.5 ± 5.3	27.3 ± 14.7	73.1 ± 8.4	15.5 ± 9.7
20%	78.3 ± 6.8	10.6 ± 5.5	20.6 ± 5.1	28.2 ± 15.5	73.2 ± 6.9	16.2 ± 10.4
	Alternative parameter estimation					
0%	85.9 ± 5.0	4.9 ± 4.0	17.8 ± 2.5	12.3 ± 9.1	81.8 ± 6.1	5.7 ± 4.2
5%	86.8 ± 4.7	**3.7 ± 3.4**	17.9 ± 2.4	11.2 ± 8.8	82.2 ± 4.2	5.5 ± 3.6
10%	**86.2 ± 4.3**	4.3 ± 3.4	**17.7 ± 2.2**	**9.8 ± 7.7**	**82.3 ± 5.1**	**5.4 ± 3.9**
20%	86.7 ± 4.7	4.6 ± 4.0	18.2 ± 1.7	13.4 ± 8.3	82.3 ± 6.4	5.7 ± 2.3

The average of the true values (tv) are displayed next to each parameter.

where SE is computed from the output of the SW classifier and NREM_p and REM_p from the RN classifier.

III. RESULTS

A. Hypnogram Estimation

The algorithm for hypnogram estimation was tested with several RFs, the obtained results are listed in Table II. The RF of 10% yields the highest values for almost all the FOMs, achieving a detection ratio of 72.8%, 77.4%, and 80.3% for wakefulness, REM, and NREM, respectively, and a global accuracy of 78.3%. This result (obtained with data previously unseen by the classifiers) corresponds to a gmean of 76.8% and a kappa index of $k = 0.58$.

For performance comparison purposes, the hierarchical classification method, with no data rejection, discussed in Section II-D was also implemented. Table III shows that the two classifiers wake/sleep and REM/NREM have relatively good performances (Acc $\approx$ 80%) which are in concordance with the performances reported in [27] for sleep/wake discrimination using cardiovascular data and ACT and [22] for REM/NREM. However, the hierarchical combination of the two classifiers (three class discrimination) leads to a poor Accuracy/Gmean, which are lower than the worst result from Table II.

B. Sleep Parameters Estimation

The three sleep parameters and the estimation error[4] were computed, for each dataset, using the estimated hypnogram and using the alternative method.

Table IV shows the average value and error for each parameter, computed for several different RFs. As expected, by incorporating the rejection information, the alternative parameter estimation outperforms the hypnogram method in all the metrics.

Using a RF of 10% and the alternative parameter estimation method, the average values are almost coincident with the real values. The estimation errors are 4.3% for the SE, 9.8% for the REM_p, and 5.5% for the NREM_p.

In order to test the influence of the training and test sets and to assess the generalization capability of the algorithm the following steps were performed:

1) Ten datasets were randomly selected from the pool of 20 available datasets.
2) From these ten datasets, five were randomly selected to train the algorithm.
3) The sleep parameters were estimated for the remaining five datasets and the average error computed.

This procedure was repeated ten times resulting in average errors of 5.9 ± 1.4, 11.8 ± 5.6, and 4.3 ± 2.6 for SE, REM_p, and NREM_p, respectively. These values are very similar to the ones reported in Table IV suggesting that the reported results should be extensible to other datasets.

IV. DISCUSSION

The automatic estimation of a hypnogram is often limited by noisy observations that need to be discarded. This is particularly relevant in real environments using data acquired from portable devices. In this paper, a HMM-based algorithm is described to overcome this limitation and compute sleep parameters from a limited set of observations.

[4] Let α represent a sleep parameter, the estimation error is given by $E_\alpha = \frac{|\alpha_{\text{true}} - \alpha_{\text{estimated}}|}{\alpha_{\text{true}}}$.

The hypnogram estimation algorithm achieves an accuracy of 78.3% with similar detection rates for all considered states. The corresponding k-index, $k = 0.58$, is, to the best of our knowledge, among the highest values reported in the literature for a three state discrimination task: [21] ($k = 0.42$), [26] ($k = 0.32$), [14] ($k = 0.45$), [41] (Acc = 76%), [42] ($k = 0.44$), and [23] ($k = 0.62$). In addition, it is important to stress that many of the cited methods discard noisy observations, preventing the estimation of a continuous hypnogram, and do not take into account the inherent temporal correlation between sleep states.

A recent study by Rosenberg *et al.* [43] shows that the interscorer agreement in the hypnogram estimation is approximately 83%. The accuracy reported in this paper is already close to this value.

The sleep parameter estimation method, designed to reject ambiguous samples, led to estimation errors of $\approx 5\%$ for SE and $\mathrm{NREM_p}$ and $\approx 10\%$ for $\mathrm{REM_p}$. These results are encouraging, suggesting that preliminary screenings for sleep disorders can be done using data acquired by noncumbersome and portable devices.

The data used in this study were collected from a heterogeneous group of subjects, having no described pathological condition. The heterogeneity of the group promotes the generalization ability of the method. However, the absence of pathologies in the dataset might lead to poor performance of the method with subjects presenting aberrant sleep patterns, like in OSAs or Insomnia. A possible approach to overcome this limitation is the use of a multimodel approach.

V. Conclusion

In this paper, we have presented a new method to estimate the Hypnogram from RR, RIP, and ACT data. The method relies on an ensemble of classifiers, trained with a rejection option and a HMM based regularization algorithm, which takes into account statistical information regarding the sleep cycle. The proposed method is able to estimate a three-state hypnogram with an acceptable accuracy, outperforming most of the state of the art algorithms. However, we have shown that the computation of sleep parameters from this hypnogram, particularly REM, and nonREM percentages, is strongly affected by the estimation error.

In order to solve this problem, we describe a method that discards ambiguous samples and estimates the sleep parameters based on the information regarding classifiers performance and rejection patterns. With this new method the estimation errors are $\approx 5\%$ for SE and $\mathrm{NREM_p}$ and $\approx 10\%$ for $\mathrm{REM_p}$.

References

[1] A. of Sleep Disorders Centers and the Association for the Psychophysiological Study of sleep. (1979, Jan.). Diagnostic classification of sleep and arousal disorders. *Sleep* [Online]. *2(1)*, p. 1–154. Available: http://www.ncbi.nlm.nih.gov/pubmed/531417

[2] D. Lger, S. R. Pandi-perumal and I. Healthcare, "Review of sleep disorders : Their impact on public health," *Public Health*, vol. 30, no. 7, pp. 92161–92161, 2007.

[3] C. A. Kushida, M. R. Littner, T. Morgenthaler, C. A. Alessi, D. Bailey, J. Coleman, L. Friedman, M. Hirshkowitz, S. Kapen, M. Kramer, T. Lee-Chiong, D. L. Loube, J. Owens, J. P. Pancer, and M. Wise, "Practice parameters for the indications for polysomnography and related procedures: An update for 2005," *Sleep*, vol. 28, no. 4, pp. 499–521, Apr. 2005.

[4] D. J. Buysse. (2005, Mar.). Diagnosis and assessment of sleep and circadian rhythm disorders. *J. Psychiatr. Pract.* [Online]. *11(2)*, pp. 102–115. Available: http://www.ncbi.nlm.nih.gov/pubmed/15803045

[5] C. Acebo and M. K. LeBourgeois. (2006, Mar.). Actigraphy. *Respirat. Care Clin. North Amer.* [Online]. *12(1)*, pp. 23–30, viii. Available: http://www.ncbi.nlm.nih.gov/pubmed/16530645

[6] A. Sadeh. (2011, Aug.). The role and validity of actigraphy in sleep medicine: An update. *Sleep Med. Rev.* [Online]. *15(4)*, pp. 259–267. Available: http://www.ncbi.nlm.nih.gov/pubmed/21237680

[7] H. J. Burgess, J. Trinder, Y. Kim, and D. Luke. (1997, Oct.). Sleep and circadian influences on cardiac autonomic nervous system activity. *Amer. J. Physiol.* [Online]. *273(4 Pt 2)*, pp. H1761–8. Available: http://www.ncbi.nlm.nih.gov/pubmed/9362241

[8] T. Electrophysiology (1996). *Heart Rate Variability. Standards of Measurement, Physiological Interpretation, and Clinical Use.* Circulation. [Online]. *93(5)*, pp. 1043–1065. Available: http://www.ncbi.nlm.nih.gov/pubmed/8598068.

[9] E. Vanoli, P. B. Adamson, Ba-Lin, G. D. Pinna, R. Lazzara, and W. C. Orr. (1995). Heart rate variability during specific sleep stages: A comparison of healthy subjects with patients after myocardial infarction. *Circulation* [Online]. *91(7)*, pp. 1918–1922. Available: http://circ.ahajournals.org/cgi/content/abstract/91/7/1918

[10] S. Elsenbruch, M. J. Harnish, and W. C. Orr. (1999, Dec.). Heart rate variability during waking and sleep in healthy males and females. *Sleep* [Online]. *22(8)*, pp. 1067–1071. Available: http://www.ncbi.nlm.nih.gov/pubmed/10617167

[11] B. V. Vaughn, S. R. Quint, J. A. Messenheimer, and K. R. Robertson. (1995, Mar.). Heart period variability in sleep. *Electroencephalogr. Clin. Neurophysiol.* [Online]. *94(3)*, pp. 155–162. Available: http://www.ncbi.nlm.nih.gov/pubmed/7536150

[12] D. E. Vigo, B. Ogrinz, L. Wan, E. Bersenev, F. Tuerlinckx, O. Van den Bergh, and A. E. Aubert, "Sleep-wake differences in heart rate variability during a 105-day simulated mission to mars," *Aviat., Space, Environ. Med.*, vol. 83, no. 2, pp. 125–130, Feb. 2012.

[13] P. K. Stein and Y. Pu. (2012, Feb.). Heart rate variability, sleep and sleep disorders. *Sleep Med. Rev.* [Online]. *16(1)*, pp. 47–66. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/21658979

[14] S. J. Redmond, P. Chazal, C. OBrien, S. Ryan, W. T. McNicholas, and C. Heneghan. (2007, Oct.). Sleep staging using cardiorespiratory signals. *Somnologie—Schlafforschung und Schlafmedizin* [Online]. *11(4)*, pp. 245–256. Available: http://www.springerlink.com/index/10.1007/s11818-007-0314-8

[15] J. W. Kantelhardt, T. Penzel, S. Rostig, H. F. Becker, S. Havlin, and A. Bunde, "Breathing during rem and non-rem sleep: correlated versus uncorrelated behaviour," *Physica A: Statist. Mech. Appl.*, vol. 319, pp. 447–457, Mar. 2003. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0378437102015029

[16] F. Yasuma. (2004, Feb.). Respiratory sinus arrhythmia: Why does the heartbeat synchronize with respiratory rhythm? *Chest* [Online]. *125(2)*, pp. 683–690. Available: http://www.chestjournal.org/cgi/doi/10.1378/chest.125.2.683

[17] X. Long and P. Fonseca, "Time-frequency analysis of heart rate variability for sleep and wake classification," in *Proc. IEEE 12th Int. Conf. Bioinformatics Bioengineering*, 2012, pp. 85–90.

[18] A. M. Bianchi, L. Mainardi, E. Petrucci, M. G. Signorini, M. Mainardi, and S. Cerutti. (1993, Feb.). Time-variant power spectrum analysis for the detection of transient episodes in hrv signal. *IEEE Trans. Biomed. Eng.* [Online]. *40(2)*, pp. 136–144. Available: http://www.ncbi.nlm.nih.gov/pubmed/8319964

[19] G. Tacchino, S. Mariani, M. Migliorini, and A. M. Bianchi, "Optimization of time-variant autoregressive models for tracking rem—non rem transitions during sleep," in *Proc. IEEE Annu. Int. Conf. Eng. Med. Biol. Soc.*, 2012, pp. 2236–2239.

[20] K. Kesper, S. Canisius, T. Penzel, T. Ploch, and W. Cassel. (2012, Feb.). Ecg signal analysis for the assessment of sleep-disordered breathing and sleep pattern. *Med. Biol. Eng. Comput.* [Online]. *50(2)*, pp. 135–144. Available: http://www.ncbi.nlm.nih.gov/pubmed/22194020

[21] M. Mendez, M. Matteucci, S. Cerutti, F. Aletti, and A. Bianchi, *Sleep Staging Classification Based on HRV: Time-Variant Analysis*. Piscataway, NJ, USA: IEEE, Sep. 2009, pp. 9–12.

[22] M. Matteucci, V. Castronovo, L. Ferini-Strambi, S. Cerutti, and A. M. Bianchi, "Sleep staging from heart rate variability: Time-varying spectral features and hidden markov models," *Int. J. Biomed. Eng. Technol.*, vol. 3, no. 3/4, pp. 246–263, 2010.

[23] T. Willemen, D. V. Deun, V. Verhaert, M. Vandekerckhove, V. Exadaktylos, J. Verbraecken, S. V. Huffel, B. Haex, and J. V. Sloten. (2013, Sep.). An evaluation of cardio-respiratory and movement features with respect to sleep stage classification. *IEEE J. Biomed. Health Informat.* [Online]. *PP(99)*, p. 1. Available: http://www.ncbi.nlm.nih.gov/pubmed/24058031

[24] A. Lewicke, E. Sazonov, M. J. Corwin, M. Neuman, and S. Schuckers. (2008, Jan.). Sleep versus wake classification from heart rate variability using computational intelligence: Consideration of rejection in classification models. *IEEE Trans. Biomed. Eng.* [Online]. *55(1)*, pp. 108–118. Available: http://www.ncbi.nlm.nih.gov/pubmed/18232352

[25] W. Karlen, C. Mattiussi, and D. Floreano. (2008, Jan.). Improving actigraph sleep/wake classification with cardio-respiratory signals. in *Proc. IEEE Annu. Int. Conf. Eng. Med. Biol. Soc.* [Online]. *2008*, pp. 5262–5265. Available: http://www.ncbi.nlm.nih.gov/pubmed/19163904

[26] S. J. Redmond and C. Heneghan. (2006, Mar.). Cardiorespiratory-based sleep staging in subjects with obstructive sleep apnea. *IEEE Trans. Biomed. Eng.* [Online]. *53(3)*, pp. 485–496. Available: http://www.ncbi.nlm.nih.gov/pubmed/16532775

[27] S. Devot, R. Dratwa, and E. Naujokat. (2010, Jan.). Sleep/wake detection based on cardiorespiratory signals and actigraphy. in *Proc. IEEE Annu. Int. Conf. Eng. Med. Biol. Soc.* [Online]. *2010*, pp. 5089–5092. Available: http://www.ncbi.nlm.nih.gov/pubmed/21096033

[28] M. Kubat and S. Matwin, "Addressing the curse of imbalanced training sets: One-sided selection," in *Proc. 14th Int. Conf. Mach. Learn.*, 1997.

[29] C. C. Berry. The kappa statistic. *J. Amer. Med. Assoc.*, vol. 268, no. 18, p. 25134, Nov. 1992. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/1404812

[30] R. J. Salin-Pascual, T. A. Roehrs, L. A. Merlotti, F. Zorick, and T. Roth. (1992). Long-term study of the sleep of insomnia patients with sleep state misperception and other insomnia patients. *Amer. J. Psychiatry* [Online]. *149(7)*, pp. 904–908. [Online]. Available: http://ajp.psychiatryonline.org/cgi/content/abstract/149/7/904

[31] G. D. Clifford, "Signal processing methods for heart rate variability analysis," Ph.D. dissertation, St Cross College, Dept. Eng. Sci., Oxford, U.K., 2002.

[32] R. Takalo, H. Hytti, and H. Ihalainen, "Tutorial on univariate autoregressive," *J. Clin. Monitor. Comput.*, no. 2005, pp. 401–410, 2006.

[33] D. Widjaja, E. Vlemincx, and S. V. Huffel, "Stress classification by separation of respiratory modulations in heart rate variability using orthogonal subspace projection," in *Proc. IEEE Eng. Med. Biol. Soc.*, 2013, pp. 6123–6126.

[34] M. Almeida and R. Vigário, "Source separation of phase-locked subspaces," in *Proceedings of the 8th International Conference on Independent Component Analysis and Signal Separation,* (ser. ICA '09). Berlin, Germany: Springer-Verlag, 2009, pp. 203–210.

[35] A. Domingues, T. Paiva, and J. Sanches, "Sleep and wakefulness state detection in nocaturnal actigraphy based on movement information," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 2, pp. 426–431, Sep. 2013.

[36] J. C. Seabra, F. Ciompi, O. Pujol, J. Mauri, P. Radeva, and J. Sanches. (2011, May). Rayleigh mixture model for plaque characterization in intravascular ultrasound. *IEEE Trans. Biomed. Eng.* [Online]. *58(5)*, pp. 1314–1324. Available: http://www.ncbi.nlm.nih.gov/pubmed/21245004

[37] P. Mahalanobis. (1936). On the generalised distance in statistics. *Proc. Nat. Inst. Sci. India* [Online]. *2(1)*, pp. 49–55. Available: http://www.citeulike.org/user/ashleygeorge/article/4155812

[38] PRTools. (2012). The MATLAB toolbox for pattern recognition. [Online; accessed 20-September-2012]. [Online]. Available: http://www.prtools.org/

[39] P. Pudil, J. Novovicova, and J. Kittler, "Floating search methods in feature selection," *Pattern Recognit. Lett.*, vol. 15, pp. 1119–1125, 1994.

[40] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inf. Theory*, vol. 13, no. 2, pp. 260–269, Apr. 1967.

[41] S. Canisius, T. Ploch, V. Gross, A. Jerrentrup, T. Penzel, and K. Kesper. (2008, Jan.). Detection of sleep disordered breathing by automated ECG analysis. in *Proc. IEEE Annu. Int. Conf.Eng. Med. Biol. Soc.* [Online]. *2008*, pp. 2602–2605. Available: http://www.ncbi.nlm.nih.gov/pubmed/19163236

[42] J. M. Kortelainen, M. O. Mendez, A. M. Bianchi, M. Matteucci, and S. Cerutti, "Sleep staging based on signals acquired through bed sensor," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 3, pp. 776–785, May 2010.

[43] R. S. Rosenberg and S. Van Hout, "The American Academy of Sleep Medicine inter-scorer reliability program: Sleep stage scoring.," *J. Clin. Sleep Med. Offic. Publicat. Amer. Acad. Sleep Med.*, vol. 9, no. 1, pp. 81–87, Jan. 2013.

Authors' photographs and biographies not available at the time of publication.