

Multi-Feature Probabilistic Detector Applied to Apnea/Hypopnea Monitoring

Di Ge¹, Alfredo Hernández¹

¹ Univ Rennes, Inserm, LTSI - UMR 1099, F-35000 Rennes, France

Abstract

Robust, real-time apnea and hypopnea detection for monitoring patients suffering from sleep apnea syndrome (SAS) still represents an open problem due to the effect of noise artifacts, the complexity of respiratory patterns and inter-subject variability. We propose in this study the application of an original multi-feature probabilistic detector (MFPD) for SAS event detection during long-term monitoring recordings on three SAS patients. The nasal pressure signal is used as input to derive a set of respiratory features (variance, peak-to-peak amplitude and total respiration cycle) which are statistically characterized during time and used to provide a mono-feature detection probability in realtime. A centralized fusion approach based on the Kullback-Leibler divergence (KLD), optimally combines these mono-feature distributions in order to produce a final detection. While the optimal feature set selection lies beyond the scope of our study, we illustrate the ability to adapt each feature's weight dynamically to make centralized fusion decisions. The method can be directly applied to data acquired from multiple sensors as long as features are synchronized. Our proposed fusion method achieves a very high sensitivity (94%) as compared with reference thresholding based methods in the literature.

1. Introduction

Full polysomnography (PSG) with overnight recordings are normally annotated manually or semi-automatically for offline diagnosis of sleep breathing disorders (SBD). The AHI index is often calculated as the number of apnea/hypopnea events diagnosed per hour and is used to measure the SBD severity to determine the corresponding treatment [1]. In clinical practice, the continuous positive airway pressure (CPAP) is the most regularly deployed treatment by the re-opening of the upper airway of SBD patients.

Galetke *et al.* [2] studied the treatment adherence by calculating the built-in run time counter of the PAP device and found that 27% of the patients had definitively discontinued PAP treatment whereas long-term adherence to positive airway pressure (PAP) treatment is essential for obstructive sleep apnea syndrome (OSAS) patients. We pro-

pose in the present paper, an automatic online detection of apnea/hypopnea events in the perspective of further development of personalized, closed-loop devices for overnight monitoring and treatment [3]. The detection method is based on the individual learning of breathing features per cycle from the nasal air pressure data and we compared the performances with classical thresholding-based detector reported in [4].

2. Method

The proposed method is based on our previous work on robust, multi-feature QRS detection [5]. The flowchart for the SAS event detection is presented in Fig. 1. The method consists in 1) preprocessing respiration waveforms by a cycle detector, 2) extracting features with each cycle, 3) learning two competitive distributions for each feature based on previous decisions, 4) measuring the distribution distances as the discriminating power of each feature, 5) making a binary decision with the KLD as weights. The whole process is initialized by a fixed prior knowledge in terms of probability distributions of the selected features.

In the following, we will detail the realization of each step with real data processing results. We focus in this paper on the enhancement of the fusion method and its advantage in choosing the best discriminating feature and its performance comparison with other threshold-based SAS event detection methods.

2.1. Preprocessing and feature description

For breathing waveform, peaks and troughs are detected using the algorithm first developed by [6] and later adapted to its filtering version in [7], as shown in Fig. 2. A maximum (peak) is validated only if its relative difference with the previous minimum (trough) exceeds the product of the peak-trough amplitude (noted as A in Fig. 2) and a threshold coefficient. The validation of a minimum follows the same principle. This technique is more robust with regard to low frequency fluctuations (baseline shifts due to relative movements) and the abrupt oscillations of high frequency artifacts. The difference in time between two consecutive maxima corresponds to a respiratory cycle (T_{tot} in seconds) as a key feature in apnea/hypopnea events while its inverse multiplied by 60 is the instant breathing

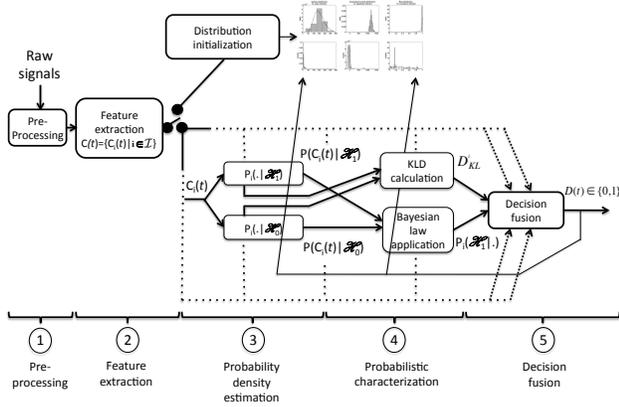


Figure 1: Flowchart of the proposed approach

rate in bpm.

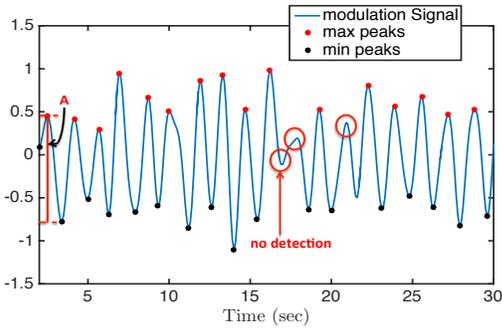


Figure 2: Example of cycle detection: low-amplitude oscillations pointed above should not yield peak/trough detections due to the thresholding mechanism while being robust to baseline shifts since only relative differences in amplitudes are taken into account.

It is noteworthy that due to the amplitude variations in apnea / hypopnea events (red and blue annotations respectively in the upper panel of Fig. 3) and during normal breathing periods, our method aims to detect only normal breathing cycles (peaks with red dot) and typical apnea/hypopnea would yield higher $Ttot$ values (red trace in the lower panel of Fig. 3), which is the main rationale to keep it as a key feature in apnea/hypopnea feature as in most SAS event detection methods [4]. Other features such as variance, peak-to-peak amplitude (blue and black traces in the lower panel of Fig. 3), however should also contribute to performance enhancement to eliminate false SAS positives, if sufficient variance or min-max difference within the cycle is achieved. These features can also be valuable in differentiating between apnea and hypopnea

though not tested in the current study. The output here can thus be noted as $C(t) = \{C_i(t) | i \in \mathcal{I}\}$, where t corresponds to each detected breathing cycle, and $\mathcal{I} = \{t, v, a\}$ the feature label set including $Ttot$, variance and peak-to-peak amplitude of the cycle.

2.2. Statistical learning

The choice of probability distribution function (pdf) of features depends on: 1) the support of feature, e.g. $(0, \infty)$ for $Ttot$, variance and peak-to-peak amplitude, 2) the feasibility of the parameter estimations, and 3) the tractability of K-L distance calculation with analytical results. In this paper, we implemented the Gamma distribution with one-sided open support:

$$P_s(x; k, \theta) = \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\Gamma(k)\theta^k} \mathbb{1}_{x>0}, \quad (1)$$

where $k \in \mathbb{R}^+$ is the shape parameter, and $\theta \in \mathbb{R}^+$ the scale parameter. The indicator function $\mathbb{1}_{x>0}$ typically limits the function support to \mathbb{R}^+ .

For initialization of $\{k, \theta\}$ for each feature (cf block *distribution initialization* in Fig. 1), they are chosen according to the sleep apnea definitions of the American academy of sleep medicine [1]. Typically, for the $Ttot$ feature, the mean ($k\theta$) is set to 4s and 10s, and the variance ($k\theta^2$) to $4s^2$ and $9s^2$ for normal and apnea/hypopnea breathing respectively. Large enough variance values should typically allow for successful first detections in the heating-up period while converge in distribution with feature samples labelled as either apnea/hypopnea (\mathcal{H}_1) or normal (\mathcal{H}_0). The maximum likelihood estimator (MLE) is used by first solving the numerically very well-behaved equation for \hat{k}_{ML} :

$$\ln(k) - \psi(k) = \ln\left(\frac{1}{N} \sum_{t=1}^N C_i(t)\right) - \frac{1}{N} \sum_{t=1}^N \ln C_i(t) \quad (2)$$

with, for example, Newton's method. $\psi(\cdot)$ is the digamma function, a.k.a. the first derivative of the log-gamma function. Analytic expression exists for $\hat{\theta}_{ML}$:

$$\hat{\theta}_{ML} = \frac{1}{kN} \sum_{t=1}^N C_i(t) \quad (3)$$

Once initialized, the MFPD detector shifts to the decision fusion mode (steps 3-5) of Fig. 1) whose final detection results feed the pdf parameter updating (step 3).

2.3. Centralized fusion

Based on the pdf model for each feature, two probabilistic markers are calculated for each feature: the event detection probability and the KLD. The detection probability for

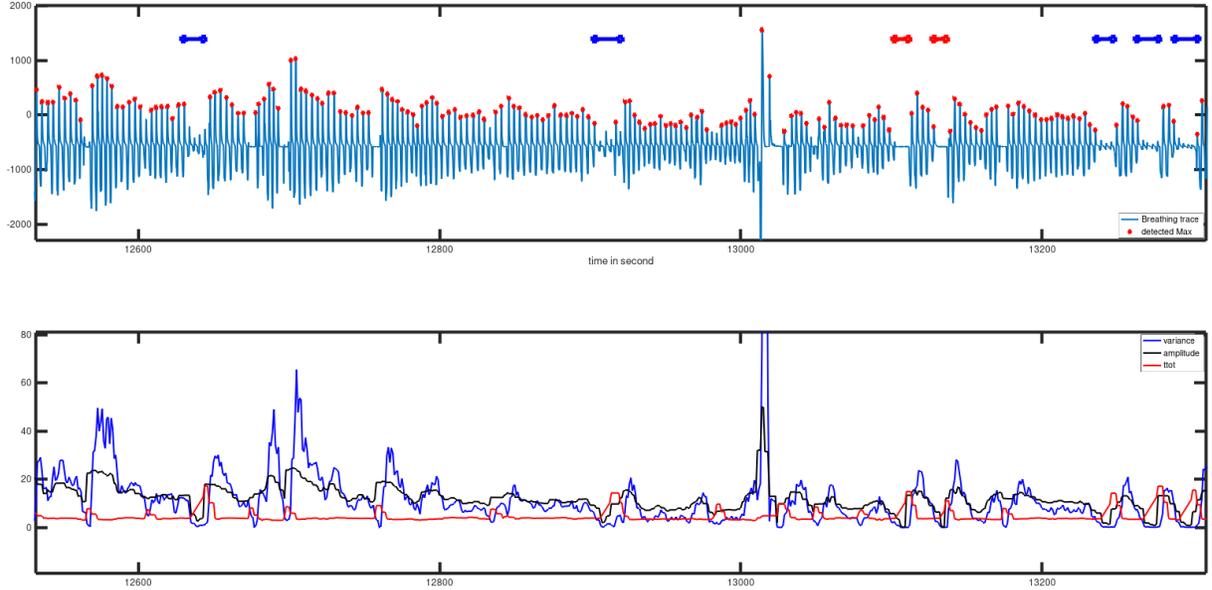


Figure 3: Nasal pressure with annotated apnea/hypopnea episodes (upper panel) and feature evolutions (lower panel) including variance (blue), amplitude (black) and T_{tot} (red).

feature $C_i(t)$ is given by

$$P_i(\mathcal{H}_1|C_i(t)) = \frac{P_i(C_i(t)|\mathcal{H}_1) P_i(\mathcal{H}_1)}{P_i(C_i(t)|\mathcal{H}_1)P_i(\mathcal{H}_1) + P_i(C_i(t)|\mathcal{H}_0)P_i(\mathcal{H}_0)}$$

using the Bayes rule.

We propose in this paper a centralized fusion by making a weighted sum of these posterior probabilities. The key here is to update dynamically a metric to measure the pertinence of each feature, or the power of separating two classes in the detection context by measuring the distance between the two competitive pdfs. The KLD is such a non-negative measure defined by:

$$D_{KL}(p||q) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)} dx. \quad (4)$$

to assess the distance between the distribution pair $P_i(C_i(t)|\mathcal{H}_{0/1})$. For the gamma distribution, its closed form [8, 9] writes:

$$D_{LK}(p||q) = (k_p - k_q)\psi(k_p) + \ln \frac{\Gamma(k_q)}{\Gamma(k_p)} - k_q \ln \frac{\theta_p}{\theta_q} + k_p \frac{\theta_p - \theta_q}{\theta_q}$$

Without loss of generality, the KLD is also one of the f -divergence functions that measures the difference between two probability distributions. They are non-negative, monotonous and jointly convex. For example, the reverse KLD $D_{KL}^R(p||q) = D_{KL}(q||p)$ is another f -divergence and can be easily calculated by reversing the role of the two distributions. However, it is less evident to derive tractable numerical methods for the Hellinger distance, or the χ^2 -divergence for the exponential pdf family.

Based on the probabilistic markers, the following decision rule is applied

$D(t) = 1$ if and only if

$$\sum_{i \in \mathcal{I}} \frac{\bar{D}_{KL}^i \cdot P_i(\mathcal{H}_1|C_i(t))}{\sum_{j \in \mathcal{I}} \bar{D}_{KL}^j} > \lambda \quad (5)$$

where \bar{D}_{KL}^i is calculated by letting $i^* = \operatorname{argmax}\{D_{KL}^i\}$:

$$\bar{D}_{KL}^i = \begin{cases} \min \left\{ D_{KL}^i, 2 \sum_{h \neq i^*} D_{KL}^h \right\}, & i = i^* \\ D_{KL}^i, & i \neq i^* \end{cases} \quad (6)$$

such that the most significant KLD should not exceed $2/3$ after normalization. Intuitively, the decision rule represents a weighted sum of detection probabilities from each feature, such that features that are better separated in distributions (between \mathcal{H}_0 and \mathcal{H}_1) have more weight in the final decision making.

3. Materials and results

3.1. Annotated PSG database

The proposed MFPD was tested on retrospective data from three obstructive sleep apnea (OSA) patients who underwent full PSG monitoring (age: 55 – 58, all male). Data were acquired at the Institut Rennais du Thorax et des Vaisseaux (IRTV), Rennes. The study was conducted in

accordance with applicable good clinical practice requirements in Europe, French law and the ethical principles of the Helsinki Declaration (1996 and 2000). The study was approved by an independent Ethics Committee and all patients provided informed consent. As part of the PSG data, a nasal pressure signal was recorded at 200Hz and annotated by experts for each SAS event (begin-end timestamps as well as apnea/hypopnea type).

The PSG annotated events are compared with the MFPD detection results after temporal realignment of the two systems. A detected event occurring less than 10 s or lying within the experts' annotation is considered true positive (TP), for either apnea or hypopnea events. While all other detections are considered false positive (FP). Experts' annotations without corresponding TP are counted as false negative (FN). Sensitivity (*se*) and positive predictive value (*ppv*) are calculated by:

$$se = \frac{TP}{TP + FN}, ppv = \frac{TP}{TP + FP} \quad (7)$$

In the following we report all these statistics for each tested patient. The calculation of specificity and negative prediction values requires typically the true negative (TN) counts that would depend on the selection of window length to segment the overall recording. For this reason, we consider it not adequate in measuring the detector's performance since it could be tuned arbitrarily high especially for patients with lower AHI scores (with less recurrent SAS events). Non-sleeping periods are also excluded from analysis thanks to the PSG annotations.

3.2. Performances

As a preliminary study, a total of 3 patients, with more than 20 hrs and 1000 OSA events are compared with reference experts' annotation based on full PSG acquisitions. The detector is highly sensitive (94%), with significant improvement than thresholding based apnea/hypopnea detection method [4], though important performance differences are observed among patients.

Rec ID	# Events	TP	FP	FN	SE	PPV
FRX10	335 (6.4 Hrs)	324	15	11	96	95
FRX09	476 (6.1 Hrs)	455	25	21	95	95
FRX08	250 (7.2 Hrs)	207	57	43	83	78
overall	1061 (19.7 Hrs)	986	97	75	92	91

Table 1: statistics per recording

4. Discussion and conclusion

The current study proposes a realtime multi-feature probabilistic detector for OSA events using nasal pres-

sure signal features only. The obtained performance levels encourages us towards the development of closed-loop, event-based therapies with individualized online learning. However, further evaluation is needed on a larger dataset. From a methods standpoint, further performance improvements might be achieved by exploiting features from other PSG signals, such as ECG or SatO2 without modifying the general structure of the MFPD detector.

References

- [1] Berry R, Budhiraja R, Gottlieb D, Gozal D, Iber C, et al. Rules for scoring respiratory events in sleep: update of the 2007 aasm manual for the scoring of sleep and associated events. deliberations of the sleep apnea definitions task force of the american academy of sleep medicine. *J Clin Sleep Med* Oct 2012;8:597–619.
- [2] Galetke W, Puzzo L, Priegnitz C, Anduleit N, Randerath W. Long-term therapy with continuous positive airway pressure in obstructive sleep apnea: Adherence, side effects and predictors of withdrawal a real-life study. *Respiration* 2011; 82:155–61.
- [3] Hernandez AI, Perez D, Feuerstein D, Liodice C, Graindorge L, Guerrero G, Limousin N, Gagnadoux F, Dauvilliers Y, Tamisier R, Prigent A, Mabo P, Amblard A, Senhadji L, Pepin JL. Kinesthetic stimulation for obstructive sleep apnea syndrome: An "on-off" proof of concept trial. *Scientific Reports* 2018;8:7. ISSN 2045-2322.
- [4] Feuerstein D, Graindorge L, Amblard A, Tatar A, Guerrero G, Christophe-Boulard S, Liodice C, Hernandez AI, Ppin J. Real-time detection of sleep breathing disorders. In 2015 Computing in Cardiology Conference (CinC). Sep. 2015; 317–320.
- [5] Doyen M, Ge D, Beuchée A, Carrault G, Hernandez A. Robust, real-time generic detector based on a multi-feature probabilistic method. *PlosOne* 2019;.
- [6] Navarro X. Analysis of cerebral and respiratory activity in neonatal intensive care units for the assessment of maturation and infection in the early premature infant. Ph.D. thesis, Université de Rennes 1, 2013.
- [7] Khreis S, Ge D, Abdul Rahman H, Carrault G. Breathing rate estimation using kalman smoother with electrocardiogram and photoplethysmogram. *IEEE Transactions on Biomedical Engineering* 2019, preprint;ISSN 0018-9294.
- [8] McCrimmon C. Distance metrics for gamma distributions 02 2018;.
- [9] Mathiassen JR, Skavhaug A, B K. Texture Similarity Measure Using Kullback-Leibler Divergence between Gamma Distributions. In *Computer Vision ECCV 2002*. Springer Berlin Heidelberg, 2002; 133–147.

Address for correspondence:

Di GE

LTSI INSERM UMR 1099, Université de Rennes 1

di.ge@univ-rennes1.fr