KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	Deviaene M., Lázaro J., Huysmans D., Testelmans D., Buyse B., Van Huffel S., Varon C. (2018), **Sleep Apnea Detection Using Pulse Photoplethysmography** in Proc. of the 45th annual Computing in Cardiology Conference (cinc2018), Maastricht, The Netherlands, Sep. 2018, pp. 1-4.
Archived version	Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher
Published version	http://www.cinc.org/2018/preprints/index.html
Journal homepage	http://www.cinc2018.org/
Author contact	Margot.deviaene@esat.kuleuven.be +32 16 32 89 60
Abstract	
IR	NA

(article begins on next page)

Sleep Apnea Detection Using Pulse Photoplethysmography

Margot Deviaene[1], Jesús Lázaro[2,3,4], Dorien Huysmans[1], Dries Testelmans[5], Bertien Buyse[5], Sabine Van Huffel[1], Carolina Varon[1]

[1]KU Leuven, Department of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, and imec, Leuven, Belgium
[2]Department of Biomedical Engineering, University of Connecticut, USA
[3]Aragon Institute for Engineering Research, University of Zaragoza, Zaragoza, Spain
[4]CIBER in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Madrid, Spain
[5]UZ Leuven, Department of Pneumology, Leuven, Belgium

Abstract

This study investigates the use of pulse photoplethysmography (PPG) for the detection of sleep apnea and its added value to oxygen saturation (SpO_2) based detection. PPG-time series known to be modulated by both respiration and the autonomous nervous system were derived: pulse rate, amplitude and width variability, slope transit time, maximal pulse upslope and the area under the PPG peak. Moreover, the instantaneous power in the high and low frequency band of the pulse rate was estimated using a point-process model. For all extracted time series, five features were computed over a 1 minute interval: the mean, minimum and maximum value, standard deviation and gradient. Feature selection resulted in the 6 most discriminative features for PPG based detection of apneic minutes. These features were used as input for a least-squares support vector machine classifier, which was applied on polysomnographic data of 102 subjects suspected of having sleep apnea-hypopnea syndrome. A classification accuracy of 68.7 % was achieved. When SpO_2 features were added to the classifier the accuracy increased to 83.4 %, which is only slightly higher than the 82.2 % obtained using only SpO_2. These results show the potential of PPG features for sleep apnea detection, however, their added value to SpO_2 is limited.

1. Introduction

Patients with sleep apnea-hypopnea syndrome (SAHS) suffer from complete or partial cessations of breathing during their sleep. These events are respectively called apneas and hypopneas. They are often accompanied with oxygen desaturations and/or arousals. Although these arousals restore normal breathing, they disturb the normal sleeping pattern and often cause excessive daytime sleepiness. Additionally, in the long term SAHS has been shown to increase the risk to develop cardiovascular diseases [1]. Therefore, timely diagnosis and treatment of SAHS is important.

SAHS is the most common sleep-related breathing disorder and recent studies have shown a high prevalence in the general population, up to 23.4 % of women and 49.7% of men are affected by SAHS [2]. However, a large portion of these people remain undiagnosed. This is mainly due to the fact that the standard SAHS diagnosis is based on an overnight assessment in the hospital, called polysomnography (PSG). This method is cumbersome, costly and not applicable to screen large populations.

Therefore many researchers have focused on the automatic detection of SAHS using signals which can easily be measured in a home environment. The most studied modalities are the Electrocardiogram (ECG) and the oxygen saturation (SpO_2), which obtain good performances [3-4]. The SpO_2, however, has as a disadvantage that only events accompanied with an oxygen desaturation can be detected. Other modalities should be added to detect events without desaturation.

Since the pulse oximeter used to measure the SpO_2 also records the pulse photoplethysmography (PPG) signal, this signal seems to be the first choice to improve the SpO_2 based detection. Previous studies have already used the PPG signal for the detection of sleep apnea [5-6], since parameters linked to autonomic arousals and respiration can be extracted from the signal.

In this study, a PPG based apnea detector will be developed using features from literature combined with newly defined features based on the point-process model of human heart beat interval dynamics to estimate the autonomic nervous system activity [7-8]. Moreover, the PPG based model will be combined with SpO_2 based apnea detection and the performances of the unimodal and multimodal classifiers will be compared.

2. Methodology

This section gives an overview of the used methods: the dataset will be described, followed by an explanation of the SpO_2 and PPG processing and ending with the feature selection and the construction of the classifier.

2.1. Data

The used dataset consists of PPG and SpO_2 signals extracted from full night PSGs recorded at the sleep laboratory of the University Hospitals Leuven. The signals were sampled at 500 Hz. 102 patients suspected of having SAHS were included in the study, they had an average age of 50 ± 12 years and a BMI of 30 ± 6 kg/m^2. 72 % of them were men and 55 % were former or current smokers.

Sleep specialists manually annotated the PSG studies according to the AASM 2012 rules [9], all apneas and the hypopneas accompanied by either an oxygen desaturation (≥ 3 %) or an arousal were labelled. For the remainder of the text, 'apneic events' will be used as a general term containing both apneas and hypopneas, unless otherwise stated. The apnea-hypopnea index (AHI) is used as severity index for SAHS and was calculated as the amount of apneic events per hour of sleep. The average AHI in the dataset was 25 ± 22, 57 % of the subjects had an AHI larger than 15, and are thus diagnosed with SAHS. The data was split into 1 minute segments, that were labelled apneic if they contained at least 10 seconds of apneic events, which corresponds to the minimum duration to be annotated [9].

A training set of 71 recordings was selected as centers of k-medoids clustering on the patient characteristics. The remaining 31 recordings were used as independent test set.

2.2. SpO_2 processing

For the SpO_2 processing and feature extraction, the workflow described in [4] was used. Oxygen desaturations were detected and six features were extracted from each of them: the SpO_2 variance, the amplitude of desaturation, the length of the resaturation period, the amplitude of the upward and downward phase rectified signal averaging (PRSA) [10] curves, and the timing of the first peak in the autocorrelation of the SpO_2. These features were then used to classify the desaturations as apneic or not.

In this study 1 minute segments were analyzed, so each desaturation needs to be linked to a segment, taking into account the delay of about 30 seconds between the start of the apnea and the start of the desaturation [4]. Therefore, desaturations were linked to the segment if they start earlier than 30 s after the end of the segment and end later than 30 s after the start of the segment. The segment got the SpO_2 feature values of the linked desaturation with the highest probability to be apneic. When no desaturation was linked to a segment, the SpO_2 features were set to NaN.

2.3. PPG processing

The PPG signals were preprocessed by applying a low-pass filter with a cut-off frequency of 35 Hz [6]. Next, the PPG peaks were detected by applying a linear filtering transformation to emphasize the peak slopes, followed by an adaptive thresholding [5]. From each detected peak, several points of interest were extracted which will be used to compute the PPG time series, see Figure 1. The peak apex point (n_A) was extracted, as well as its basal (n_B) and mid-point (n_M). Moreover, the first derivative of the signal was used to compute the peak onset (n_O), end point (n_E), and the slope end point (n_{Es}) [6].

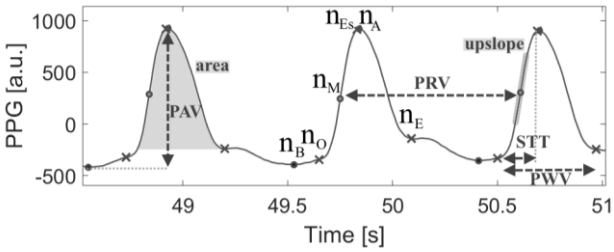

Figure 1. Overview of the PPG peak point delineation and feature extraction.

From these peak points 6 PPG time series were computed. The pulse rate variability (PRV) was computed as the inverse of the time between consecutive n_M's; the pulse amplitude variability (PAV) as the amplitude difference between n_B and n_A; and the pulse width variability (PWV) as the time between n_E and n_O. The slope transit time (STT) was defined as the time in between n_{Es} and n_O. Additionally, the maximal pulse upslope and the area under the peak were computed, as can be seen on Figure 1. These time series were all resampled to 4 Hz.

The autonomic nervous system can be assessed by studying the low frequency (0.04-0.15 Hz, LF) and high frequency power (0.15-0.4 Hz, HF) of the PRV. During apneic events, however, there is a close succession of parasympathetic and sympathetic changes [1]. Therefore it is interesting to study these parameters on short time intervals. To do so, the point-process model of human heart beat interval dynamics developed by Barbieri et al. [7] was applied to the PPG peak mid-points n_M, as surrogate of the R-peaks in the ECG. This model uses a history-dependent inverse Gaussian distribution to model the RR-interval time series and provides an instantaneous estimate of the heart rate LF power, the HF power and the LF/HF index [7]. Again a sampling frequency of 4 Hz was used.

The 9 computed PPG-time series were post-processed using a median absolute deviation (MAD)-based outlier detection rule [5]. Outliers longer than five times the average heartrate interval were removed, those shorter were interpolated. After outlier detection, 1-min averaged features were computed from the time series. The minimum, maximum and mean value, the standard

deviation and the gradient during the 1-min segment were extracted from each of the time series. This resulted in a total of 45 PPG features.

2.4. Feature selection and classification

From the 45 PPG features, the most discriminative ones to separate apneic minutes from normal breathing minutes were selected. A similar feature selection workflow as in [4] was applied: first highly correlated features were removed, then the minimal redundancy maximal relevance algorithm was used to further reduce the feature set to 10 features [11]. Finally any remaining redundant features were removed using a backwards wrapper combined with the classifier explained below.

Using the selected features, a fixed-size least-squares support vector machines (FS LS-SVM) classifier using an RBF kernel was built to classify segments as apneic or normal breathing [12]. In order to balance the training set, an equal number of segments from the apnea and normal class were included per training patient. All segments from the smallest class were included, the segments from the largest class were selected as the centers of k-medoids clustering on the feature set. The LS-SVM hyper parameters were optimized using 10-fold cross validation on this training subset.

This classifier was first trained on both the PPG and SpO_2 features separately, to obtain the classification performance per modality. Afterwards the classifier was trained on a feature set containing both the PPG and SpO_2 features.

3. Results

The PPG feature selection resulted in a subset of 6 features: the minimum of the PAV, standard deviation and mean of the PRV, standard deviation of the PWV and the minimum and mean value of the LF/HF power of the pulse rate. Boxplots of four of them are depicted in Figure 2, comparing the normal and apneic segments.

The obtained classification performances on the test set of the PPG, SpO_2 and combined classifier are presented in Table 1.

4. Discussion

When observing the selected PPG features in Figure 2, a decrease in PAV is seen during apneic minutes. This decrease has been linked to vasoconstriction and arousals due to apneic events [5]. Moreover, an increased standard deviation of the PRV is observed indicating autonomic nervous system activations and the increased LF/HF power points towards an increased sympathetic activation.

Using the six PPG features, the PPG-based apnea classifier was trained. The PPG classification performance,

Table 1. Classification performances on the test set. Accuracy (Acc), sensitivity (Se), specificity (Sp), area under the curve (AUC) and positive predictive value (PPV) are all given in percentages.

Classifier	Acc	Se	Sp	AUC	PPV
PPG	68.7	69.1	68.6	74.8	42.4
SpO2	82.2	**75.3**	84.5	87.6	62.6
PPG+SpO2	**83.4**	73.7	**86.6**	**88.4**	**64.8**

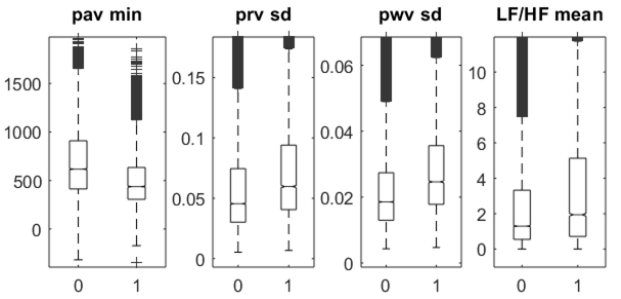

Figure 2. Four of the selected PPG features for normal breathing segments (0) and apneic segments (1). Minimum PAV (a.u.), standard deviation of the PRV (1/s) and PWV (s), and the mean LF/HF power (a.u.) are depicted. Due to space limitations, the other features are not shown.

given in Table 1, is lower than in previous studies in literature. We achieve an accuracy of 68.7 %, whereas in [6] an accuracy of 72.66% was obtained. In this study, however, only obstructive apneas were considered. Our study considers all types of apneic events, including central events and also hypopneas which only have a partial obstruction of the airway and are thus more difficult to detect. Only 11.9 % of the apneic minutes in this study include obstructive apneas. If only these event minutes are taken into account, the accuracy, sensitivity, specificity and AUC of PPG based detection rise to, respectively, 76.2 %, 67.9 %, 76.5 % and 78.0%. The sensitivity and AUC further increase when central and mixed apneas are added, probably due to the increase in available training data. An accuracy of 74.8 % is then achieved with a sensitivity, specificity and AUC of, respectively, 73.8 %, 74.9 % and 80.4 %.

Moreover, the performance parameters vary a lot between patients. If a patient has only few apneic events, they are often more isolated, have a less severe autonomic response and are therefore more difficult to detect. When apneas are grouped together, they are more easily to detect and thus higher sensitivities are obtained. Another observation that can be made is the low positive predictive value (PPV) of 42.4 %. Previous studies also encountered the same problem. A lot of sympathetic activations are detected, which are not linked to annotated apneic events. These could be due to subclinical apneas which are too short to be scored according to the AASM rules, or other causes such as movement of the patient. In [6], it was

investigated whether these false positives occurred more often during REM sleep, which was not the case. In this study the patients were on average in REM sleep for 14 % of the recordings. During this time 20 % of all apneic minutes and 20 % of all false positives occured. There is thus no excessive amount of false positive detections during REM sleep. Therefore, REM sleep autonomous nervous system variations cannot account for the low PPV.

The SpO_2 based classifier clearly outperforms the PPG based classifier with an accuracy of 82.2 %. The SpO_2 signal, however, is not always available in wearable PPG measurement systems. If both modalities are available, they can be combined, then an accuracy of 83.4 % is obtained, which is only slightly better than taking only the SpO_2 features into account. These results show that the PPG features are promising for the detection of sleep apnea events, however, when the SpO_2 signal is available, the added value of the PPG is limited.

One of the main objectives was to improve the sensitivity of the SpO_2 based apnea detector by adding the PPG features, as extra apneic events which are not linked to oxygen desaturations could then be detected. In this dataset, during 8.8 % of all apneic minutes no desaturation of at least 1 % could be observed, these apneic minutes could therefore not be detected using the SpO_2 signal. The low PPV of the PPG based classifier, however, makes it impossible to improve the sensitivity without greatly decreasing the specificity of the classifier.

These results might still be improved if extra PPG features are considered, investigating for example the dynamics and nonlinear properties of the PPG signal. Moreover, the full night recording is now processed, including periods where the patient lies awake during the night. Implementing some PPG-based sleep staging could avoid false positives during wake, which now represent 22.4 % of the false positives for PPG based detection and 49.2 % for SpO_2 based detection.

5. Conclusion

This study investigated the use of PPG derived features for the detection of apneic minutes in overnight recordings and their added value compared to SpO_2 based features. The PPG based results are promising, but the positive predictive value should still be improved to be useful in practice. Moreover, the PPG features have shown to have a limited added value compared to the SpO_2 based features.

Acknowledgements

This work is supported by: Agentschap Innoveren en Ondernemen (VLAIO) Project #:

SWT 150466 - OSA+. imec funds 2017. imec ICON projects: ICON HBC.2016.0167. European Research Council: The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Advanced Grant: BIOTENSORS (n° 339804). This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information. Carolina Varon is a postdoctoral fellow of the Research Foundation-Flanders (FWO). This project has received funding from the European Union's Framework Programme for Research and Innovation Horizon 2020 (2014-2020) under the Marie Sklodowska-Curie Grant Agreement No. 745755.

References

[1] Bradley T.D. and Floras J.S. Obstructive sleep apnoea and its cardiovascular consequences. The Lancet 2009; 373.9657;82-93.

[2] Heinzer R. et al. Prevalence of sleep-disordered breathing in the general population: the HypnoLaus study. Lancet Respir Med 2015;3.4;310-318.

[3] Varon C. et al. A novel algorithm for the automatic detection of sleep apnea from single-lead ECG. IEEE TBME 2015;62.9;2269-2278.

[4] Deviaene M. et al. Automatic screening of sleep apnea patients based on the SpO_2 signal. IEEE JBHI 2018; Early Access; DOI:10.1109/JBHI.2018.2817368.

[5] Lázaro J. et al. Pulse rate variability analysis for discrimination of sleep-apnea-related decreases in the amplitude fluctuations of pulse photoplethysmographic signal in children." IEEE JBHI 18.1 (2014): 240-246.

[6] Lázaro J. et al. Pulse photoplethysmography derived respiration for obstructive sleep apnea detection. Computing in cardiology 44 2017;1.

[7] Barbieri R. et al. A point-process model of human heartbeat intervals: new definitions of heart rate and heart rate variability. Am J Physiol Heart Circ Physiol 2005;288.1;H424-H435.

[8] Barbieri R. and Brown E. N. Analysis of heartbeat dynamics by point process adaptive filtering. IEEE TBME 2006;53.1;4-12.

[9] Berry RB et al. Rules for scoring respiratory events in sleep: update of the 2007 AASM manual for the scoring of sleep and associated events. J Clin Sleep Med 2012;8.5; 597-619.

[10] Bauer A. et al. Phase-rectified signal averaging detects quasi-periodicities in non-stationary data. Physica A: Statistical Mechanics and its Applications 2006;364;423-434.

[11] Peng H. et al. Feature selection based on mutual information criteria of max-dependency, max-relevance and min-redundancy. IEEE TPAMI 2005;27.8;1226-1238.

[12] Suykens J et al. Least squares support vector machines. World Scientific, 2002.

Address for correspondence.

Margot Deviaene
ESAT-STADIUS, KU Leuven
Kasteelpark Arenberg 10, bus 2446, 3001 Leuven, Belgium.
margot.deviaene@esat.kuleuven.be