Sleep Apnea Detection Using Time-Delayed Heart Rate Variability

Marina-Marinela Nano, Xi Long, *Member, IEEE*, Jan Werth, Ronald M. Aarts, *Fellow, IEEE*, and Richard Heusdens

Abstract— **Sleep apnea is a sleep disorder distinguished by repetitive absence of breathing. Compared with the traditional expensive and cumbersome methods, sleep apnea diagnosis or screening with physiological information that can be easily acquired is needed. This paper describes algorithms using heart rate variability (HRV) to automatically detect sleep apneas as long as it can be easily acquired with unobtrusive sensors. Because the changes in cardiac activity are usually hysteretic than the presence of apneas with a few minutes, we propose to use the delayed HRV features to identify the episodes with sleep apneic events. This is expected to help improve the apnea detection performance. Experiments were conducted with a data set of 23 sleep apnea patients using support vector machine (SVM) classifiers and cross validations. Results show that using eleven HRV features with a time delay of 1.5 minutes rather than the features without time delay for SA detection, the overall accuracy increased from 74.9% to 76.2% and the Cohen's Kappa coefficient increased from 0.49 to 0.52. Further, an accuracy of 94.5% and a Kappa of 0.89 were achieved when applying subject-specific classifiers.**

I. INTRODUCTION

Sleep apnea (SA) is a highly prevalent sleep disorder characterized by repetitive cessation of breathing during sleep time. It can be classified as obstructive, central, and mixed SA based on whether an effort to breathe is present or not during the apneic event [1]. Among these classes, the obstructive SA is the most prevalent one, where a loss of pharyngeal dilator muscle tone at sleep onset causes recurrent pharyngeal collapse and temporary cessation of breathing. SA can yield daytime sleepiness and compromised alertness, due to which behavioral effects, such as moodiness and petulance, may arise. Accompanied with that are increased chances of traffic and work accidents. Ultimately severe SA can lead to long term clinical impacts such as diabetes or even death [2] [3] [4].

Polysomnography (PSG) is the "gold standard" for the diagnosis of sleep-related disorders in general. It measures multi-channel electrophysiological signals where respiratory effort, airflows, snoring, and blood oxygen saturation are often used for apnea diagnosis during sleep in clinical practice, where the apnea-hypopnea index (AHI) is usually derived as an indicator of SA severity [5]. Yet PSG is expensive and cumbersome, where patients are required to be hospitalized and many electrodes are connected to the patient's body during overnight recording. Recent wearable or unobtrusive measurement technologies have made it possible to capture some physiological activity. For example, the heart rate variability (HRV) has been shown to vary accompanying the presence of sleep apneic events, where bradycardia during the apneic periods, followed by tachycardia during the ensuing ventilatory phases, is consistently found [6] [7].

Due to the cessation of breathing during apnea, gradual and unsynchronized changes in cardiac rhythm have been found compared to the presence of apneas [8]. Another cause of the hysteretic cardiac variations is the end-apneic arousals, clearly leading to changes in (autonomic) cardiac activity after sleep apneic events [9]. Figure 1 illustrates the time difference between the changes in heart rate and the presence of apneic events. Therefore, we were motivated to use the later HRV activity to detect each SA episode/epoch.

In recent years, methods for SA detection based on HRV features have been investigated extensively [10] [11] [12] [13]. In contrast to previous work, this paper proposes to detect SA using delayed HRV features. To the best of the authors' knowledge, this time-delay scheme was not studied. It is important to note that the purpose of this work was not to beat the SA detection performance of previous studies; instead we aimed at demonstrating the effectiveness of using the time-delayed features in helping improving the detection. In addition, due to the existence of individual differences in cardiac pathophysiology for SA patients, we expected that using a subject-specific approach would result in a better SA detection performance than a subject-independent approach. A support vector machine- (SVM-) based classifier, which has been successfully used in previous studies [11] [12] [14], was employed in this work as well.
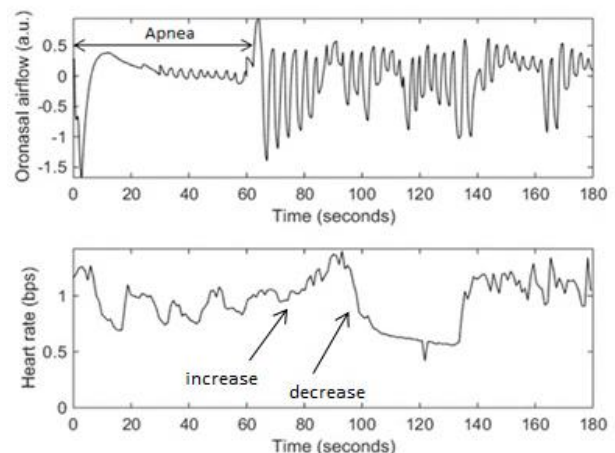

Figure 1. An example of the nasal airflow and heart rate of three minutes during sleep, where the first minute was annotated by human experts as sleep apnea and the next two minutes as normal breathing. The delayed changes in heart rate can be clearly seen.

II. Materials and Methods

A. Data

The Apnea-ECG database used in this work is available from PhysioNet for the *CinC Challenge 2000* [15]. Here we used only patients with an AHI greater than 4, corresponding to the mild, moderate and severe cases. This was to include sufficient epochs with apneic events for training for each patient when using the subject-specific approach. Therefore 23 single-lead ECG recordings (sampled at 100 Hz with a 12-bit resolution) were used and the other twelve (b01, b03 and c01-c10) were excluded. The recordings had varying lengths (from slightly less than 7 hours to nearly 10 hours each) and were acquired from 23 SA patients (men, age 51.7 ± 6.5 years ranging from 38 to 63 years, weight 96.5 ± 17.8 kg). The apneas were annotated on a one-minute basis by human experts.

B. R-peak localization

As mentioned before, cyclic variations in HRV have been reported to be associated with the presence of SAs. To extract HRV features, heartbeat or RR interval (RRI, the time interval between consecutive heartbeats or R peaks) is required. Here we simply adopted the Pan and Tompkins' QRS detector [16] to locate the R peaks, resulting in RRI series used for extracting features for SA detection.

C. Feature extraction

In this work, HRV features were extracted from RRIs for each 30-sec epoch. The reason we used 30-sec epochs was due to the fact that the hysteretic cardiac fluctuation (usually tachycardia) caused by apneic events last from a few seconds up to over a minute [17]. Using a higher resolution in HRV features can help better see the cardiac changes. Accordingly, the annotations were modified by assigning each minute of recording into two consecutive 30-sec epochs. Here a sliding window of nine epochs (or 4.5 minutes) was used for computing each HRV feature for each 30-sec epoch, where the central epoch was the epoch of interest for SA detection. This aimed at capturing the changes in (autonomic) cardiac activity as recommended in [18].

A pool of eleven existent HRV features in the time domain (known in literature [12] [19]) was extracted for SA detection. As mentioned, this work focused on showing the applicability of the time-delayed scheme rather than achieving best detection results. We thus chose a subset from a large number of the existent features presented in literature. Moreover, the choice of those time-domain features was also because of their relatively low computation complexity compared with the other features, for example, in the frequency domain. This is important from an implementation point of view when using a wearable/unobtrusive device for SA detection. The eleven features include:

- RR_{mean}, mean of RRIs
- RR_{median}, median of RRIs
- MAD, mean absolute deviation of RRIs
- SDNN, standard deviation of RRIs
- SDSD, standard deviation of the differences between successive RRIs

- RMSSD, root mean square (RMS) of the differences between successive RRIs
- IQR, inter-quartile range, the difference between the 75th and 25th percentiles of the RRI distribution
- $NN50_{first}$, number of pairs of successive RRIs where the first RRI exceeds the second RRI by more than 50 ms
- $NN50_{second}$, number of pairs of successive RRIs where the second RRI exceeds the first RRI by more than 50 ms
- $pNN50_{first}$, $NN50_{first}$ divided by the total number of RRIs
- $pNN50_{second}$, $NN50_{second}$ divided by the total number of RRIs.

Note that all the features were normalized with a Z-score normalization (zero mean and unit variance) for the purpose of mitigating the effects caused by individual differences.

D. Support vector machines

SVM-based classifiers have been widely used for SA detection [11] [12] [14]. SVM is a machine learning method allowing for obtaining the optimal boundary of two data sets or classes which are non-linearly mapped in a high-dimension feature space, independently on the probabilistic distributions of training vectors in the data sets.

The fundamental idea of SVM is to locate the boundary that is the most distant to the vectors nearest to the boundary in both of the data sets, referred to as the maximum margin hyperplane. Figure 2 illustrates the basic idea of SVM for a binary-class problem. The circles and squares represent samples of two classes "-1" and "+1", respectively; x1 and x2 represent two features. SVM handles the class separation by using a kernel function to map the data into a different space with a hyperplane. In this work, we used the radial basis function (RBF) as a kernel function. More details about SVM can be found elsewhere [20].
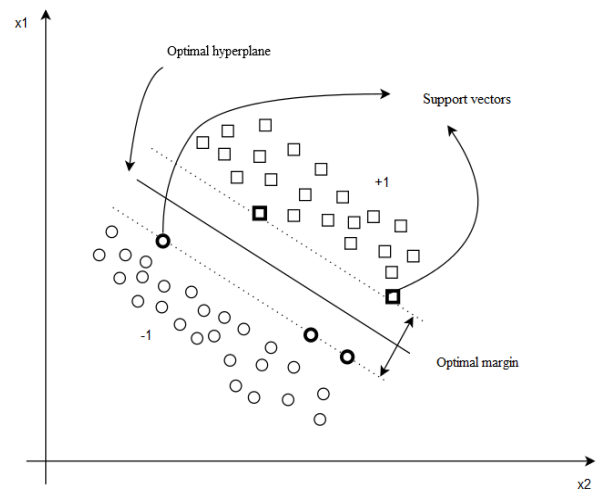

Figure 2. Description of the basic idea of SVMs in a two-dimensional space, where the two classes "-1" and "+1" can be separated by the optimal hyperplane (with support vectors) using two features x1 and x2.

E. Time-delayed detection of SA

As mentioned, we propose to use the HRV features with a certain time delay to perform SA detection. For the i-th epoch ($i = 1, 2, …, N$) from a given recording with N 30-sec epochs, we used the features computed from the ($i+j$)-th epoch (with a delay of j epochs) to identify the class (SA or non-SA) of the i-th epoch.

Introducing the time delay was expected to improve the features by capturing the gradual and delayed changes in cardiac rhythm (Figure 1) in distinguishing between apnea and non-apnea (or normal) epochs, thus enhancing the overall performance of SA detection. A positive time delay ($j > 0$) corresponds to that the cardiac changes are later/slower than the presence of apneas while a negative time delay ($j < 0$) means that the cardiac changes precede the apneic events.

III. EXPERIMENTS AND EVALUATION

A subject-independent approach was first considered where the classifier to detect SA epochs for each patient was trained by the data from the other patients. With this approach, a leave-one-out cross validation (LOOCV) was conducted. During each iteration of the LOOCV procedure, data from 22 recordings were used for training and the remaining one was used for testing. After the cross validation, classification results obtained for each patient in each iteration's testing set were collected, and performance metrics (aggregated over all patients) were then computed to evaluate the classifier.

In addition, we considered a subject-specific approach, performing the SA detection for each patient based on the classifier trained by the data from the same patient. A 10-fold cross validation (10-fold CV) was conducted to evaluate this approach. First, the data of each patient was randomized so that both training and testing sets would contain SA and non-SA (normal) epochs. During each iteration of the 10-fold CV procedure, 90% of the data of each patient was used to train the classifier, while the remaining 10% was used for testing. After that, classification results obtained for each patient in the testing set of each iteration were collected, and performance metrics (aggregated over all patients) were then computed.

The classification performance was first evaluated using a traditional metric of overall accuracy. However, a strong imbalance between classes observed in our data made this metric less appropriate for evaluation. The Cohen's Kappa coefficient is a measure that takes into account the prior probability of a specific class occurring. This means that it offers a better understanding of the general classification performance in correctly identifying both classes.

IV. RESULTS AND DISCUSSION

As shown in Figure 3, the detection performance using the subject-independent approach reached a maximum at a time delay of 1.5 minutes (or 3 epochs). This confirms that the cardiac changes are slower than the presence of apneic events, suggesting that they are not a predictor of apnea, but a consequence. For the globally optimized time delay of 1.5 minutes, we reached an overall accuracy of 76.2% and a Cohen's Kappa coefficient of 0.52 (see Table I), which were better than those obtained without time delay (at $j = 0$).

For the subject-specific approach, SA detection results using the HRV features with 1.5-min delayed were also found to be better than those produced without time delay. For this time delay, we achieved a slightly increased overall accuracy of 94.5% and a Kappa coefficient of 0.89 in SA detection. These results were much better than those obtained using the subject-independent approach, which confirms the previous results presented by Varon *et al.* [11]. This means that the pathophysiology of SA expressed by cardiac activity exhibits large differences between patients.

TABLE I. SUMMARY OF SA DETECTION PERFORMANCE

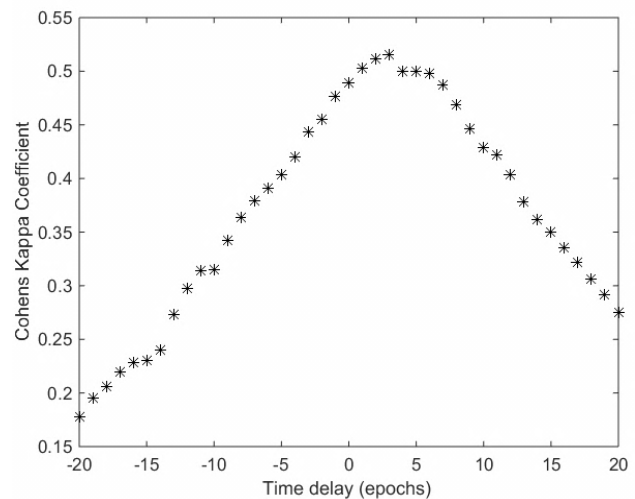| Metric	Subject-independent		Subject-specific	
	Without time delay	*With time delay*	*Without time delay*	*With time delay*
Accuracy (%)	74.9±0.09	76.2±0.08	94.1±0.03	94.5±0.03
Kappa	0.49±0.21	0.52±0.20	0.88±0.11	0.89±0.09

Figure 3. SA detection performance (Cohen's Kappa coefficient) using the HRV features with different time delays.

To compare the detection results between patients with different SA severity levels, Figure 4 plots the Kappa value versus AHI. For the subject-specific approach, the Cohen's Kappa coefficient varies from 0.55 to 0.94, while for the subject-independent approach it varies from 0.02 to 0.78. The SA detection performance for the subject-specific approach seems to be independent of the AHI value. However, for the subject-independent approach the Kappa values for patients with a lower AHI seem higher than those for patients with a higher AHI, as shown in the figure. The reason is not yet clear and needs further investigation with a larger data set including more SA patients.

For the subject-specific approach the model has to be trained using the specific patients beforehand, and can therefore not be used for short-term SA screening. However, after the training phase, the subject-specific approach can result in an increased SA detection performance, which makes it suitable for long-term monitoring and follow-up of SA screening, for example, in a home environment.

For both approaches, the SA detection performance was improved when the 1.5-min delayed HRV features were used.
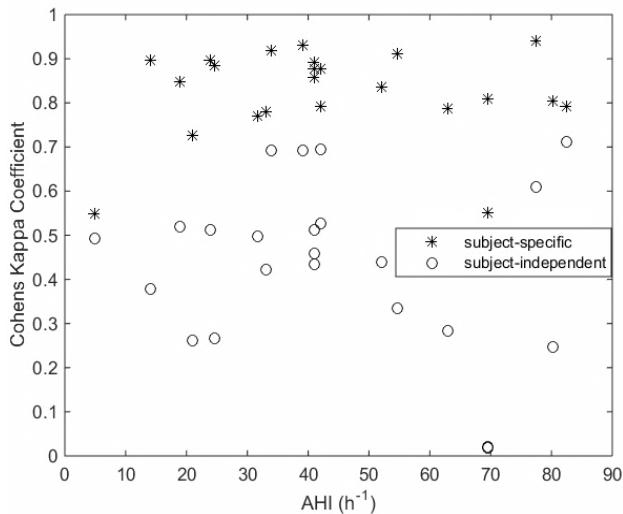
Figure 4. SA detection performance (Cohen's Kappa coefficient) versus AHI of the SA patients for the subject-independent and the subject-specific approach.

The time delay of 1.5 minutes was considered the optimal one that can improve the overall detection performance for all patients. However, it might be influenced by AHI value, varying from patient to patient so that the optimal time delay for each patient could be different. This might also explain the slight performance improvement. The optimal time delay might differ also over time throughout the night. Therefore, a subject-specific 'time-varying time delay' should be further studied. Additionally, to better understand the intra-subject variability for subject-specific approach, our SA detector trained based on the data from previous recordings will be investigated in the future.

Although this preliminary study shows promising results in general for SA detection when using the time-delayed HRV features, a larger-sized data set including more patients is required to test our method in future work.

V. CONCLUSION

This preliminary study has performed SA detection using eleven time-domain HRV features (derived from single-lead ECG). In practice, these features can be derived from the heartbeat intervals acquired using advanced unobtrusive or wearable devices. In order to further improve the detection performance by capturing the hysteretic changes in cardiac rhythm in comparison with the occurrence of apneic events, a novel method of using the features with a time delay of 1.5 minutes was proposed. It has been demonstrated that using the time-delayed features, the SA detection results can be clearly improved compared with when those obtained using the features without delay in time. In addition, the SA detection for the subject-specific approach performed much better than for the subject-independent detection approach. We also speculate that the SA detection performance is associated with the severity level of SA, requiring further study using a larger data set with different severity levels.

REFERENCES

[1] N. M. Punjabi, "The epidemiology of adult obstructive sleep apnea," *Proc. Am. Thorac. Soc.,* vol. 5, no. 2, pp. 136-143, 2008.

[2] S. M. I. Mary, B. Lam, M. M. T. Ng, W. K. Lam, K. W. T. Tsang and K. S. L. Lam, "Obstructive sleep apnea is independently associated with insulin resistance," *Am. J. Respir. Crit. Care Med.,* vol. 165, no. 5, pp. 670-676, 2002.

[3] N. Botros, J. Concato, V. Mohsenin, B. Selim, K. Doctor and H. K. Yaggi, "Obstructive sleep apnea as a risk factor for type 2 diabetes," *Am. J. Med.,* vol. 122, no. 12, pp. 1122-1127, 2009.

[4] H. K. Yaggi, J. Concato, W. N. Kernan, J. H. Lichtman, L. M. Brass and V. Mohsenin, "Obstructive sleep apnea as a risk factor for stroke and death," *New Engl. J. Med.,* vol. 353, no. 19, pp. 2034-2041, 2005.

[5] C. Iber, S. Ancoli-Israel, A. J. Chesson and S. F. Quan, *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications*, American Academy of Sleep Medicine, 2007.

[6] K. Narkiewicz, N. Montano, C. Cogliati, P. J. van de Borne, M. E. Dyken and V. K. Somers, "Altered cardiovascular variability in obstructive sleep apnea," *Circulation,* vol. 98, no. 11, pp. 1071-1077, 1998.

[7] N. A. Collop, W. M. Anderson, B. Boehlecke, D. Claman, R. Goldberg, D. J. Gottlieb, D. Hudgel, M. Sateia and R. Schwab, "Clinical guidelines for the use of unattended portable monitors in the diagnosis of obstructive sleep apnea in adult patients," *J. Clin. Sleep Med.,* vol. 3, no. 7, pp. 737-747, 2007.

[8] J. A. Dempsey, S. C. Veasey, B. J. Morgan and C. P. O'Donnell, "Pathophysiology of sleep apnea," *Physiol. Rev.,* vol. 90, no. 1, pp. 47-112, 2010.

[9] R. J. Kimoff, "Sleep fragmentation in obstructive sleep apnea," *Sleep,* vol. 19, no. 9 Suppl., pp. S61-66, 1996.

[10] F. Roche, J. M. Gaspoz, I. Court-Fortune, P. Minini, V. Pichot, D. Duverney, F. Costes, J.-R. Lacour and J.-C. Barthelemy, "Screening of obstructive sleep apnea syndrome by heart rate variability analysis," *Circulation,* vol. 100, no. 13, pp. 1411-1415, 1999.

[11] C. Varon, D. Testelmans, B. Buyse, J. A. Suykens and S. Van Huffel, "Sleep apnea classification using least-squares support vector machines on single lead ECG," in *Conf. Proc. IEEE Eng. Med. Biol. Soc. (EMBS)*, Osaka, Japan, 2013, pp. 5029-5032.

[12] B. Yilmaz, M. H. Asyali, E. Arikan, S. Yetkin and F. Ozgen, "Sleep stage and obstructive apneaic epoch classification using single-lead ECG," *Biomed. Eng. Online,* vol. 9, p. 39, 2010.

[13] M. Drinnan, J. Allen, P. Langley and A. Murray, "Detection of sleep apnoea from frequency analysis of heart rate variability," in *Comput. Cardiol.*, Cambridge, MA, 2000, vol. 27, pp. 259-262, 2000

[14] L. Almazaydeh, K. Elleithy and M. Faezipour, "Detection of obstructive sleep apnea through ECG signal features," in *IEEE Int. Conf. Electro/Inf. Technol.*, Windsor, Ontario, Canada, 2012, pp. 1-6.

[15] T. Penzel, G. B. Moody, R. G. Mark, A. L. Goldberger and J. H. Peter, "The Apnea-ECG Database," in *Comput. Cardiol.*, Cambridge, MA, 2000, vol. 27, pp. 255-258.

[16] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.,* vol. 32, no. 3, pp. 230-236, 1985.

[17] C. A. Nigro and E. E. Rhodius, "Variation in the duration of arousal in obstructive sleep apnea," *Med. Sci. Monit.*, vol. 11, vol. 4, pp. CR188-192, 2005

[18] Task Force of the European Society of Cardiology and The North American Society of Pacing and Electrophysiology, "Heart rate variability: standards of measurement, physiological interpretation, and clinical use," *Eur. Heart J.,* vol. 17, pp. 354-381, 1996.

[19] P. de Chazal, T. Penzel and C. Heneghan, "Automated detection of obstructive sleep apnoea at different time scales using the electrocardiogram," *Physiological Measurement,* vol. 25, no. 4, pp. 967-983, 2004.

[20] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York: Springer-Verlag, 1995.