Apnea Detection Based on Respiratory Signal Classification

Laial Almazaydeh, Khaled Elleithy, Miad Faezipour and Ahmad Abushakra
Department of Computer Science and Engineering
University of Bridgeport
Bridgeport, CT 06604, USA
{lalmazay, elleithy, mfaezipo, aabushak}@bridgeport.edu

Abstract— Obstructive sleep apnea (OSA) is the most common form of different types of sleep-related breathing disorders. It is characterized by repetitive cessations of respiratory flow during sleep, which occurs due to a collapse of the upper respiratory airway. OSA is majorly undiagnosed due to the inconvenient Polysomnography (PSG) testing procedure at sleep labs. This paper introduces an automated approach towards identifying the presence of sleep apnea based on the acoustic signal of respiration. The characterization of breathing sound was carried by Voice Activity Detection (VAD) algorithm, which is used to measure the energy of the acoustic respiratory signal during breath and breath hold. The performance of our classification algorithm is tested on real respiratory signals and the experimental results show that the VAD is useful as a predictive tool for the segmentation of breath into sound and silence segments. Moreover, the system we developed can be used as a basis for future development of a tool for OSA screening.

Keywords: sleep apnea, OSA, PSG, VAD, respiration signal.

I. INTRODUCTION

Sleep apnea (SA) in the form of Obstructive sleep apnea (OSA) is becoming the most common respiratory disorder during sleep, which is characterized by cessations of airflow to the lungs. These cessations in breathing must last more than 10 seconds to be considered an apnea event. Apnea events may occur 5 to 30 times an hour and may occur up to four hundred times per night in those with severe SA [1].

The most frequent night symptoms of SA can include snoring, nocturnal arousals, sweating and restless sleep. Moreover, like all sleeping disorders, symptoms of sleep apnea do not occur just during the night. Daytime symptoms can range from morning headaches, depression, impaired concentration and excessive sleepiness which cause mortality from traffic and industrial accidents. However, these symptoms are not definitive to detect SA syndrome [2] [3].

In fact, SA is not a problem to be taken lightly, since it is associated with a major risk factor of health implications and increased cardiovascular disease and sudden death. It has been linked to irritability, depression, sexual dysfunction, high blood pressure (hypertension), learning and memory difficulties, in addition to stroke and heart attack [2] [3]. Several treatment options for OSA patients include weight loss, positional therapy, oral appliances, surgical procedures and continuous positive airway pressure (CPAP). CPAP is a common and effective treatment especially for patients with moderate to severe OSA. CPAP devices are masks worn during sleep that improves oxygen saturation and reduces sleep fragmentation [4].

Statistics show that around 100 million people worldwide, and in the US from 18 to 50 million people, are suspected to have OSA. This is while more than 80% of which remain undiagnosed [5]. The trouble of having examinations discourages patients prone to OSA undergo at the overnight clinical research through polysomnographic data. Polysomnography (PSG) is a complicated procedure and certain way of assessing the OSA problem. Complete PSG includes the monitoring of the breath airflow, respiratory movement, oxygen saturation (SpO₂), body position, electroencephalography (EEG), electromyography (EMG), electrooculography (EOG), and electrocardiography (ECG) [6]. However, PSG has received many criticisms from some researchers. This is due to several reasons, including first, the inconvenience since it requires the patient to be connected to numerous sensors and to stay in hospital for one night. Second, it is expensive. The average cost for a PSG is $2,625 due to the need for the study to take place in a specially equipped medical facility, in addition to the requirement of having a sleep lab staff overnight. Trained in ‘scoring’ the resultant measurements manually. Third, a long wait list of up to 6 months is caused by limited availability of PSG [7].

According to the American Academy of Sleep Medicine (AASM), the Apnea-Hypopnea Index (AHI) is used to describe the number of complete and partial apnea events per hour of sleep and it is calculated to assess OSA syndrome severity. OSA severity is usually determined as follows: AHI 5-15 indicates mild, 15-30 indicates moderate and over 30 indicates severe OSA syndrome. Therefore, patients are diagnosed with OSA if they have five or more apnea events per hour of sleep during a full night sleep period [8].

However, new simplified methods for diagnosis and screening of OSA are needed, in order to have a major benefit of the treatment on OSA outcomes. In this work, we develop an efficient algorithm for automatic classification of respiratory signal to detect abnormalities in breathing or breathing cessations. We use Voice Activity Detection (VAD) to classify respiratory signals into normal respiration and sleep apnea. The detection process would check for apnea attack for a time period of fifteen seconds or more.
In the following sections, we glance at a variety of sleep apnea detection methods. Section III, gives a general description of VAD. In section IV, the methodology of our proposed system is described. Section V demonstrates the results of our system. Then, we conclude our paper in section VI, and highlight some directions for future research.

II. RELATED WORK

Over the past few years most of the related research has focused on presenting methods for the automatic processing of different statistical features of different signals such as thorax and abdomen effort signals, nasal air flow, oxygen saturation, electrical activity of the heart (ECG), and electrical activity of the brain (EEG) for the detection of SA.

In our previous published research, we developed a Neural Network (NN) as a predictive tool for OSA using SpO2 signal and evaluated its effectiveness [9]. In addition, in [10] we further developed a model based on a linear kernel Support Vector Machines (SVMs) using a selective set of RR-interval features from short duration epochs of the ECG signal. The results show that our automated classification system can recognize epochs of SA with a high degree of accuracy, approximately 96.5%.

As OSA is generally caused by a blocked of the airflow airway and it is characterized by repetitive episodes of breathing cessation, the respiratory signal recording analysis during sleep becomes very valuable in order to estimate respiratory flow and distinguish the changes in the breathing pattern of the patient. Then, to provide additional and complementary information, other biological signal data measurements such as ECG and SpO2 could be bridged to analyze sleep data, as clinical experience indicates that an apneic event is frequently accompanied by a fall in the blood oxygen saturation (SpO2) [11], and cyclic variations in the duration of a heartbeat (ECG); this consists of bradycardia during apnea followed by tachycardia upon its cessation [12].

Recently, based on the SpO2 and tracheal breathing sound recording analysis during sleep, the study in [13] reports a new fully automatic technology for OSA detection. Different parameters were investigated to distinguish the breathing level during each individual apnea event. Therefore, in the first step, the drops (more than 4%) and rises of the SpO2 signal were marked, then, the total energy of the tracheal sound segments within the periods between a drop and the following rise in the SpO2 were found. After collection of data, each parameter was then fuzzified with a sigmoid function and the fuzzy outputs were added together to classify the sound signals. The results show high sensitivity and specificity values of more than 90% in differentiating normal respiration from disordered breathing in patients.

Several studies for the non-invasive sensing of the respiration rate are based either on the measurement of the nasal airflow or on the evaluation of the respiration movements (abdominal or chest respiration signals). Using mean absolute amplitude analysis of the combination of the thoracic and the abdominal signals in [14] showed that both signals are able to indicate the occurrence of SA events. The analysis identification results achieved a sensitivity ranged from 70.29-86.25% and the specificity values ranged from 74.82 to 90.09%.

It has been reported that snoring is a common finding in people with OSA. OSA is generally caused by a blocked of the airflow airway. Therefore, the snoring must be due to the vibration of soft tissues when the airflow stimulates the ill structure in the upper airway during sleep [15]. Of all methods for diagnosing OSA, the formants estimation method is most widely used. The formants information contains the essential acoustic properties of the upper airway. It has been discovered by studies that there is a correlation between the state of the upper airway and the first formant frequency. A narrower upper airway is usually led to a higher first formant frequency. Andrew et al. [16] and [17] proposed fixed formant frequency thresholds to detect the hypopneic snores which must be higher than that of the typical ones.

Various portable monitor devices already exist in the market. SleepStrip™ is one of the carries available in home sleep test diagnostic devices. This device has to be worn for a minimum of five hours of sleep, and the actual device is placed on the individual’s face where the two flow sensors (oral and nasal thermistors) are placed just below the nose and above the upper lip to capture the breath of the patient. For all samples combined, sensitivity and specificity values ranged from 80-86% and 57-86% respectively [18].

Even though most of the related studies yielded promising initial results, more improvement is needed, as it either requires physical attachment to a user or may be unreliable. Therefore, since apnea is a condition where patient pauses breathing; this can be of great concern for detecting breathing through the sound signal where an appropriate alarm can be released upon its cessation.

III. VOICE ACTIVITY DETECTION – THE PRINCIPLE

In this work, the principle of voice activity detection (VAD) algorithm employed to detect the presence or absence of apnea on real breathing signals is described.

Voice Activity Detector plays an important role in speech processing techniques such as speech coding [19], speech enhancement, and speech recognition [20]. Other examples include cellular radio systems (GSM and CDMA based) [21], hands-free telephony [22], VoIP applications and echo cancellation.

VAD relies on measurement of features from speech which yield highly in differentiating between voiced and unvoiced segments, where the regions of voice information within a given audio signal are referred to as ‘voice-active’ segments and the pauses between talking are called ‘silence’ or ‘voice-inactive’ segments. Therefore, the performance trade-offs of VAD algorithm are made by maximizing the detection rate of active speech while minimizing the false detection rate of inactive segments [23].

The most important part in VAD classifier is feature extraction, from which different regions in the audio signal can be separated. Common features used in the VAD detection process are cepstral coefficient [24], spectral entropy [19], zero-crossing rate [20, 25], least square periodicity measure
[26], and average magnitude difference function [27]. Another important and widely used feature in this regard is signal energy, which is presented in this work, and compared with the dynamically calculated threshold.

IV. APNEA DETECTION USING VAD BASED - ENERGY

The general VAD block diagram used in our methodology is shown in figure 1.

The assumptions on the VAD algorithm here is that the speech is quasi-stationary and its spectral changes quickly over short periods like 20-30ms, but the background noise is relatively stationary and changes very slowly with time. In addition, the energy of the active speech level is usually higher than background noise energy [28].

In the first step, the respiratory signal is filtered to remove the undesired low frequency components. Then, the power with different window sizes of the Fast Fourier Transform (FFT) is calculated for the filtered signal [29].

Let \( x(t) \) be the input signal samples, and \( X(n) \) be Fast Fourier Transform (FFT) samples. The VAD algorithm begins with the energy computation within the smallest integer range of frequency values \( n_1 \) and \( n_2 \):

\[
Energy = \sum_{n=n_1}^{n_2} |X(n)|^2
\]

(1)

The energy of the signal is computed in two window frames; short window and long window for every window number \( i \):

\[
E_{\text{short}}(i) = \varphi_{\text{short}} E_{\text{energy}} + (1 - \varphi_{\text{short}}) E_{\text{short}}(i)
\]

(2)

\[
E_{\text{long}}(i) = \varphi_{\text{long}} E_{\text{energy}} + (1 - \varphi_{\text{long}}) E_{\text{long}}(i)
\]

(3)

The number of frames used is \( N/L \), where \( N \) represents the number of samples in the signal and \( L \) represents the window size in frequency domain. The coefficients \( \varphi_{\text{short}} \) and \( \varphi_{\text{long}} \) refer to the window-length factors, where \( \varphi_{\text{short}} = 1/16 \), and \( \varphi_{\text{long}} = 1/128 \), and \( L = 528 \) were used in this study.

At this point, since VAD aims to differentiate voice and silence, where silence is mostly referring to background noise, the noise level at every frame needs to be computed. For this purpose, a threshold \( THR \) value needs to be determined for comparing the signal value against noise:

\[
THR = \frac{K_f}{1-\varphi_{\text{long}}} + M
\]

(4)

In the above formulation, \( K_f \) is the \( K \)-th frame and \( M \) is a margin value that can be considered to separate voice and silence in the event that noise level is flat.

The VAD technique eventually makes a decision by comparing every frame of signal energy against the \( THR \) value. It is important to note that transitional periods from active voice to silence may also affect the decision. Based on the above steps and discussions, the decision on the VAD identifier (ID) values is made as follows:

\[
\text{VAD-ID} = \begin{cases} 
1, & \text{if } Energy > THR \\
1, & \text{if } Energy \leq THR \text{ and in transitional period} \\
0, & \text{if } Energy \leq THR \text{ and not in transitional period}
\end{cases}
\]

(5)

The outcome of the VAD technique is the separated speech and silence phases which can be fine-tuned for identifying breath versus breathing cessations for apnea detection.

In our work, there is a second threshold (\( Tr \)) that would be used to decide whether a silence phase corresponds to apnea or not. According to the sleep apnea literature, a breathing cessation (silence) of 15 seconds or more would be classified as apnea, as shown below:

\[
\text{If } (\text{VAD-ID}_j = 0) \text{ and } (T_{\text{VAD-ID}} \geq 15 \text{Sec}) \Rightarrow \text{VAD-ID}_j \text{ is a SA event period}
\]

(6)

In the above relationship, \( T_{VAD-ID} \) corresponds to the duration of silence phase \( j \) detected by the VAD technique.

The testing procedure and experimental results to evaluate the described VAD algorithm will be discussed through the next section.

V. EXPERIMENTAL RESULTS

MATLAB environment was used to perform our methodology on various samples of breathing signals during breathing and breath hold in 50 normal people. The volunteers were asked to breath 20 cycles. They were asked to hold their breath before, during and after the 20 cycles. The breath recording was done using the SONY VAIO VPCEB42FM microphone (Realtek High Definition Audio [30] professional microphone with the accompanying Audacity software).

The human respiratory signal is given to the classification system as the input, and the coding is developed in such way that it calculates the fundamental feature of the respiratory signal, which is the energy. The threshold is then applied to the extracted energy feature and the binary decision is made. VAD=1 is declared if the energy feature exceeds the threshold. Otherwise, VAD=0 is for no breath or when silence (cessations of breathing) is present.

Figure 2 shows the results obtained from the segmentation technique of the input signal which splits the acoustic signal of respiration into silence and voiced phases. The start point and end point of a respiratory signal which contains breathing

![Figure 1. A block diagram of VAD design.](Image)

![Figure 2. Result of the VAD classification system.](Image)
phases is determined in this work. Hence, the apnea events that are silence phases lasting 15s or longer can be detected.

VI. CONCLUSIONS AND FUTURE WORK

This work sought to determine the effectiveness of VAD based - energy in distinguishing the apnea in breathing signal.

The provided respiratory signal is classified successfully with the help of the formulated algorithm with more than 97% accuracy.

In order to detect sleep apnea in real time, the proposed algorithm could be improved and adjusted by adding calibration procedures to run on an FPGA [31]. After the successful design and implementation of the OSA system, it is planned to be experimentally tested in order to evaluate its accuracy and practicality. The tests will take place in a local hospital for a set of patients who have symptoms of OSA. In addition, a different set of other subjects without SA symptoms will test the system to verify its false positive accuracy.

REFERENCES

[1] Sleep Disorders Guide. www.sleepdisorderguide.com.

[2] N. Israel, A. Tarasiuk, and Y. Zigel, “Nocturnal Sound Analysis for the Diagnosis of Obstructive Sleep Apnea,” in Proceedings of the 32nd IEEE International Conference on Engineering in Medicine and Biology Society (EMBS 2010), pp. 6146-6149, Sep. 2010.

[3] A. Yilmaz, and T. Dundar, “Home Recording for Pre-Phase Sleep Apnea Diagnosis by Holter Recorder Using MMC Memory,” in Proceedings of the 2010 IEEE International Conference on Virtual Environments Human-Computer Interfaces and Measurement Systems (VECMIS), pp. 126-129, Sep. 2010.

[4] “Choosing a CPAP,” American Sleep Apnea Association, www.sleepapnea.org/resources/pubs/cpap.htm.

[5] SleepMedInc. www.sleepmed.md.

[6] D. Baraglia et al., “Automated Sleep Scoring and Sleep Apnea Detection in Children,” in Proceedings of SPIE 6039, 2005.

[7] “Sleep Study Cost”, Shop & Compare Healthcare Facilities and Cost with NewChoiceHealth.com. Web. 03 Dec. 2010. www.newchoicehealth.com/SleepStudyCost.

[8] P. Chazal, C. Heneghan, and W. Mcnicholas, “Multimodal Detection of Sleep Apnoea Using Electrocardiogram and Oximetry Signals,” Philosophical Transactions of The Royal Society A: Mathematical, Physical and Engineering Sciences, vol. 367, no. 1887, pp. 369-389, 2009.

[9] L. Almazaydeh, M. Faezipour, and K. Elleithy, “A Neural Network System For Detection Of Obstructive Sleep Apnea Through SpO2, Signal Features,” (IJACSA) International Journal of Advanced Computer Science and Applications, vol. 3, no. 5, pp. 7-11, Jun. 2012.

[10] L. Almazaydeh, K. Elleithy, and M. Faezipour, “Detection of obstructive sleep apnea through ECG signal features,” in Proceedings of the IEEE International Conference on Electro Information Technology (IEEE ei2012), pp. 1-6, May. 2012.

[11] M. Canosa, E. Hernandez, and V. Moret, “Intelligent Diagnosis of Sleep Apnoea Syndrome,” In IEEE Engineering in Medicine and Biology Magazine, vol. 23, no. 2, pp. 72-81, 2004.

[12] P. Chazal, T. Penzel, and C. Heneghan, “Automated Detection of Obstructive Sleep Apnoea at Different Time Scales Using the Electrocardiogram,” Institute of Physics Publishing., vol. 25, no. 4, pp. 967–983, Aug. 2004.

[13] A. Yadollahi, and Z. Moussavi, “Acoustic Obstructive Sleep Apnea Detection”, In Proceedings of IEEE Conference on Engineering in Medicine and Biology Society (EMBC 2009), pp. 7110-7113, Sep. 2009.

[14] A. Ng, J. Chung, M. Gohel, et al., “Evaluation of the Performance of Using Mean Absolute Amplitude Analysis of Thoracic and Abdominal Signals for Immediate Indication of Sleep Apnoea Events,” Journal of Clinical Nursing, vol. 17, no. 17, pp. 2380–2386, Sep. 2008.

[15] Y. Zhao, H. Zhang, W. Liu, and S. Ding, “A Snoring Detector for OSAHS Based on Patient’s Individual Personality,” In 3rd International Conference in Awareness Science and Technology (iCAST), pp. 24-27, 2011.

[16] K. Andrew, S. Tong, et al. “Could Formant Frequencies of Snore Signals Be an Alternative Means for the Diagnosis of Obstructive Sleep Apnea?” Else, Sleep Medicine, vol. 9, pp. 894-898, Dec. 2008.

[17] K. Andrew, T. Koh, E. Baey, et al., “Speech-Like Analysis of Snore Signals for the Detection of Obstructive Sleep Apnea”, In International Conference on Biomedical and Pharmaceutical Engineering 2006 (ICBPE 2006), pp. 99-103, Dec. 2006.

[18] T. Shochat, N. Hadas, M. Kerkhofs, et al., “The SleepStripTM: An Apnoea Screenr for the Early Detection of Sleep Apnoea Syndrome,” European Respiratory Journal, vol. 19, pp. 121-126, 2002.

[19] S. McClellan, J. Gibson,”Spectral entropy: an alternative indicator for rate allocation,” In IEEE International Conference on Acoustics, Speech, Signal Processing, pp. 201-204, Apr. 1994.

[20] B. Atal, L. Rabiner, “A Pattern Recognition Approach to Voiced-Unvoiced-Silence Classification with Applications to Speech Recognition,” IEEE Transactions on Acoustics, Speech, Signal Processing, vol. 24, no. 3, pp. 201-212, June 1976.

[21] ETSI TS 126 094 V3.0.0 (2000-01), 3G TS 26.094 version 3.0.0 Release 1999. Universal Mobile Telecommunications System (UMTS); Mandatory Speech Codec speech processing functions AMR speech codec; Voice Activity Detector (VAD), 2000.

[22] A. Benyassine, E. Shlomot, and H. Su, “ITU-T recommendation G.729 annex B: A silence compression scheme for use with G.729 optimized for V.70 digital simultaneous voice and data application,” IEEE Communications Magazine, vol. 35, no. 9, 1997.

[23] E. Verteleetskaya, K. Sakhnov, “Voice Activity Detection for Speech Enhancement Applications,” ACTA Polytechnica, vol. 50, no. 4, pp. 100-105, 2010.

[24] J. Haigh, and J. Mason, “Robust Voice Activity Detection Using Cepstral Features,” In Proceedings of IEEE Region 10 Conference on Computer, Communication, Control and Power Engineering (TENC’93), vol. 3, pp. 321-324, 1993, Beijing.

[25] A. Sangwan, C. MC, H. Jamadagni, et al., “VAD Techniques for Real-Time Speech Transmission on the Internet,” In Proceedings of 5th IEEE International Conference on High-Speed Networks and Multimedia Communications, pp. 46-50, 2002.

[26] S. Tanyer, and H. Ozer, “Voice Activity Detection in Nonstationary Gaussian Noise,” In Proceedings of International Conference on Signal Processing (ICSP’98), pp. 1620-1623, 1998.

[27] M. Orlandi, A. Santarelli, and D. Falavigna, “Maximum Likelihood Endpoint Detection with Time Domain Features,” In Proceedings of 8th European Conference on Speech Communication and Technology (eurospeech 2003), pp. 1757-1760, 2003, Geneva.

[28] K. Sakhnov, E. Verteleetskay, and B. Simak, “Approach for Energy-Based Voice Detector with Adaptive Scaling Factor,” IAENG International Journal of Computer Science”, vol. 36, no. 4, pp. 394, 2009.
[29] A. Abushakra and M. Faezipour, “Acoustic Signal Classification of Breathing Movements to Virtually Aid Breath Regulation,” IEEE Journal of Biomedical and Health Informatics, vol. 17, no. 2, pp. 493-500, March 2013.

[30] VAIO® Laptop ComputersVPCEB42FM/BJ: Available from: http://www.docs.sony.com/release/VPCEB3_Series.pdf, 2011.

[31] B. Marinkovic, M. Gillette, and T. Ning, “FPGA Implementation of Respiration Signal Classification Using a Soft-Core Processor,” In Proceedings of the IEEE 31st Annual Northeast Bioengineering Conference, pp. 54-55, April 2005.