31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Normal Probability Testing of Snore Signals for Diagnosis of Obstructive Sleep Apnea

H. Ghaemmaghami, U. R. Abeyratne and C. Hukins

Abstract—**Obstructive Sleep Apnea (OSA) is a highly prevalent disease in which upper airways are collapsed during sleep, leading to serious consequences. The standard method of OSA diagnosis is known as Polysomnography (PSG), which requires an overnight stay in a specifically equipped facility, connected to over 15 channels of measurements. PSG requires (i) contact instrumentation and, (ii) the expert human scoring of a vast amount of data based on subjective criteria. PSG is expensive, time consuming and is difficult to use in community screening or pediatric assessment. Snoring is the most common symptom of OSA. Despite the vast potential, however, it is not currently used in the clinical diagnosis of OSA.**

In this paper, we propose a novel method of snore signal analysis for the diagnosis of OSA. The method is based on a novel feature that quantifies the non-Gaussianity of individual episodes of snoring. The proposed method was evaluated using overnight clinical snore sound recordings of 86 subjects. The recordings were made concurrently with routine PSG, which was used to establish the ground truth via standard clinical diagnostic procedures. The results indicated that the developed method has a detectability accuracy of 97.34%.

I. Introduction

THERE are three syndromes of sleep apneas: obstructive [1], central [2] and hypopnea [3]. Obstructive sleep apnea (OSA) is a sleep disorder that is commonly diagnosed by means of overnight sleep studies or Polysomnography (PSG). Sleep apnea refers to episodes of non-breathing events lasting equal to or more than 10 seconds [4]. A PSG test can determine the number of obstructive breathing events per hour. This calculation is commonly known as the Respiratory Disturbance Index (RDI). It is a representation of the total number of apneas, hypopneas and respiratory arousals that occur per hour during sleep. The RDI value is utilised to diagnose and evaluate the degree of a subject's OSA disorder. Typically, patients with (RDI<10) are considered normal while those with (RDI≥10) are diagnosed with OSA [1]. Obstructive apnea episodes, at times, awaken the patient, resulting in sleep fragmentation. During apnea episodes, the level of oxygen may drop to dangerously low levels, which may then result in cardiac arrhythmias, and in turn, can be fatal. Hence, individuals with sleep apnea are more prone to heart attacks and strokes. Additionally, other effects of sleep apnea may include depression, irritability, sexual dysfunction, learning and memory difficulties, and fatigue [5].

Many patients who have treatable OSA go undiagnosed by both the primary care physician and the specialist [6, 7, 8]. OSA is a syndrome that pervades the general population and is currently under-diagnosed [9]. Due to this, several approaches for finding simpler and more practical methods, than PSG, for the detection and classification of OSA have been studied in the literature. These studies all conduct diagnosis of OSA through utilising an extractable feature of the snore sounds. The features employed include: energy [10], zero-crossing rate [10, 11], power spectrum [12], Linear Predictive Coding (LPC) coefficients [13], formants [14] and, pitch and jitter [15] of snore signals. In addition, more computationally complex methods such as Higher-Order Statistics (HOS) of snore signals have also been employed [16]. These techniques have provided strong evidence that snoring carries sufficient evidence to characterize OSA. However, the performance of these methods can be further improved in terms of the sensitivity/specificity of detection and the computational complexity.

In this paper, we develop a feature based on the non-Gaussianity of snore episodes as a means of characterizing the severity of OSA. We call the proposed feature the *Non-Gaussianity-Index* (NGI). The NGI is computed for each individual snore episode and then an overall measure is computed for each patient to serve as an index of the severity of OSA. The proposed method is fully automated and free from subjective interpretation.

The ability to diagnose OSA reliably, based on a single channel of non-contact snore measurement, has the potential to change the way OSA epidemic is managed throughout the world. The low-cost snore acquisition instrumentation does not need expert humans to operate, and will be suitable for population screening and pediatric use. The large amount of time a sleep expert has to devote to the manual scoring of overnight data can be saved, due to the availability of reliable automated technology.

II. Database

The database consisted of 86 subjects, both male and female, with RDI's ranging from 0.5 to 106.7. The subjects

Manuscript received April 23, 2009. This work was supported by the Australian Research Council (ARC) under Grant No. ARC DP0773687 to Dr. Abeyratne.

H. Ghaemmaghami is with School of Information Technology and Electrical Engineering,, The University of Queensland, St. Lucia, Brisbane, Australia, (Phone: +61 73365 2378, Fax: +61 7 3365 4999, e-mail: houman@itee.uq.edu.au)

U. R. Abeyratne is with School of Information Technology and Electrical Engineering,, The University of Queensland, St. Lucia, Brisbane, Australia, (Phone: +61 73346 9063, Fax: +61 7 3365 4999, e-mail: udantha@itee.uq.edu.au)

C. Hukins is with the Sleep Diagnostic Unit, Princess Alexandra Hospital, Brisbane, Australia.

had undergone a PSG assessment at the Respiratory and Sleep Disorders Unit of The Princess Alexandra Hospital, Brisbane, Australia. The data was recorded during the overnight sleep study. This was carried out using a high fidelity, computerised sound acquisition system.

The PSG data was gathered from a Compumedics sleep acquisition system, and consisted of EEG,EOG, ECG, EMG, Leg movements, respiration nasal air flow, nasal pressure, respiratory movements, blood oxygen saturation, breathing sounds and the body position.

The sound acquisition system consisted of a pair of matched low noise microphones having a hypercardiod beam pattern (Model NT3, RODE, Sydney, Australia). The nominal distance from the microphone to the mouth of the patient was 50 cm, but could vary from 40 to 70 cm due to patient movements. A professional quality pre-amplifier and A/D converter unit (Model Mobile-Pre USB, M-Audio_, California, USA) was used for data acquisition, at a sampling rate of 44.1 kHz and a 16 bits/sample resolution. The recorded data was a collection of Snore Related Sounds (SRS).

III. METHOD

The method proposed in this paper assumes the following SRS signal model:

$$x[n] = s[n] + b[n] + y[n] \tag{1}$$

Where, $x[n]$ represents the discrete-time SRS signal, $b[n]$ the background and instrument noise with Gaussian distribution, and $y[n]$ the non-Gaussian background noise or activities, such as speech, equipment fumbling and testing, or thumping sounds. However, as the data acquisition process was carried out in a controlled environment the non-Gaussian processes modeled by $y[n]$ were later eliminated. Hence, (1) was simplified to:

$$x[n] = s[n] + b[n] \tag{2}$$

Furthermore, it can be said that the nature of the $s[n]$ component of the SRS signal varies; $s[n]$ at times represents snoring sounds made by the studied subject. These sounds have a periodic nature and thus posses a non-Gaussian distribution. However, $s[n]$ could also represent breathing sounds of the subject. In this case, $s[n]$ resembles a white process with a Gaussian distribution.

A. Signal Pre-Processing

Initially, the input SRS signal, $x[n]$, was passed through a decimator to obtain a lower sampling rate, as the bandwidth of snore signals is less than 10 kHz and the data required to carry out normal probability analysis is sufficient within this bandwidth. In addition, decimation of signals with high sampling rates could contribute to the removal of high frequency noise.

The similarities between the generation mechanism of human speech and snore signals led to the application of the pre-emphasis filter. In speech analysis, a -6 dB/octave roll-off in the spectrum of voiced speech exists. This is due to the voiced excitation source and the radiation from the mouth. The pre-emphasis filter is utilised to correct this roll-off. Hence, this filter was employed in the pre-processing stage of the SRS signal. The cut-off of the pre-emphasis filter can be set by altering α in (3), which is typically set to (α = 0.96) in speech processing applications [17]. Hence, this value was utilised for this application.

$$x[n]_{\text{pre}-\text{emphasised}} = x[n] - \alpha x[n-1] \tag{3}$$

B. Non-Gaussianity Index (NGI)

A new index, called the Non-Gaussianity Index (NGI), hereon referred to using the symbol ψ, was developed to conduct the SRS signal analysis. The developed index provides a measure of non-Gaussianity of a given segment of data. To obtain ψ, normal probability plot analysis was employed. The normal probability plot can be utilised to obtain a visual measure of the Gaussianity of a set of data.

To obtain ψ, for $x[n]$, the inverse of the normal Cumulative Distribution Function (CDF) for the data was first calculated:

$$\gamma = F^{-1}(p|\mu,\sigma) = \left\{\gamma : F(\gamma|\mu,\sigma) = p\right\} \tag{4}$$

Where,

$$p = F(\gamma|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\gamma} e^{\frac{-(t-\mu)^2}{2\sigma^2}}\, dt \tag{5}$$

$$\mu = \text{mean of } x[n]$$

$$\sigma = \text{standard deviation of } x[n]$$

The probabilities were then assigned to p in (4) to obtain the required γ values for plotting the normal probability plot of the data in $x[n]$. However, the normal probability plot conducts a comparison between γ and probabilities of a would-be, or reference Gaussian dataset, if $x[n]$ was to represent a Gaussian data segment. This is thus calculated and referred to as the probability set g in (6).

The deviation of the probability plot of the analysed data (γ) to its reference Gaussian probability plot (g) was chosen as the measure of non-Gaussianity, or ψ. Linear regression was utilised to acquire a measure for this deviation and thus calculate ψ. It must be noted that $g[i]$ and $\gamma[i]$ represent the probabilities of the reference normal data, and the analysed data, respectively, with i ranging from the values 1 to N:

$$\psi = 1 - \left(\frac{\sum_{i=1}^{N} \left(g[i] - \overline{g} \right)^2}{\sum_{i=1}^{N} \left(\gamma[i] - \overline{\gamma} \right)^2} \right) \tag{6}$$

After calculating ψ, the obtained values were assigned to each analysed segment as an obtained score representing the frame of data. $x[n]$ was then windowed using rectangular windows of length L, and ψ values within the segments of length L of $x[n]$ were thresholded using threshold value β as displayed below:

$$x[n] \text{ for } (n = 1,2,3,\ldots N) = \begin{cases} 1 & \text{if}, \quad \psi > \beta \\ 0 & \text{if}, \quad \psi \le \beta \end{cases}$$

Finally, L-length segments of $x[n]$ that contained frames with ψ values above the threshold β were marked as apneaic snore episodes. A count of the apneaic snore episodes over a 5 hour period of the recorded SRS data was conducted. This count was normalised by time to achieve a new index indicating the count of apneaic episodes per hour for each subject. This index was titled the Apnea Count Index (ACI).

IV. Results

The SRS data, recorded during PSG, was utilised for training and testing of the developed method. To conduct the experiment, the recorded data from the 86 subjects was grouped into training and testing sets. This division can be seen in Table I. The evaluation of the developed method and the ACI was then carried out using four independent experiments. Each experiment utilised a different RDI value as the ground truth for the diagnosis. Hence, four RDI values were utilised to classify the training and test set subjects into OSA and non-OSA patients. The RDI values employed were equal to 5, 10, 15, and 30. However, as RDI of 10 is most commonly employed in OSA diagnosis, the results of this experiment were quoted throughout the paper.

Each of the four experiments was carried out for 10 independent L-length segmentations of the SRS signals. The L values were chosen to be: 1, 3, 5, 7, 9, 10, 15, 20, 25, and 30 seconds in length. This was done to mark apneaic episodes. Typically a segment length of 30 seconds is utilised in PSG studies to obtain the estimated apneaic snore episodes and thus obtain RDI. In this study, the 10 distinct L values were utilised to study the affects of varying the episode lengths.

To conduct each of the L-length experiments, initially, ψ values for frames of length (N=100ms) of the SRS signals were obtained. After that, the training set was utilised to obtain an optimum β value for the thresholding process. To do this, the training sets were thresholded using a range of β values. For each β value associated ROC curves were plotted through variation of the obtained ACI values. The area under the ROC curves was utilised as the detectability measure and the ROC curve with the largest detectability measure was marked. The β value was then recorded and utilised to test the testing set. Hence, the test set was also tested at various L values. However, a set β value was utilised. The ACI value utilised for diagnosis of the subjects was also chosen from the training experiments and utilised to diagnose the test subjects. ROC curves were then plotted, through variation of the ACI values, to obtain detectability measures for the testing experiments. Table II displays the optimum variable values and the final detectability accuracy achieved using these variables for each of the four experiments. It can be seen that the detectability of the method is reduced below 90% as the diagnostic RDI value is raised to 30.

TABLE I
Training and Testing Set Divisions of SRS Database

Set Type	Subjects	RDI Range	RDI < 10	RDI ≥ 10
Training	20	0.5 – 92.9	5	15
Testing	66	0.6 – 106.7	17	49
Total	86	0.5 – 106.7	22	64

Fig. 1 (a) and (b) display the ACI versus RDI graphs for the test subjects and two of the four experiments. Each of the graphs represent the diagnosis of the subjects based on an optimum L, β, and ACI values obtained during the training experiments. Each experiment utilised a set RDI for the reference diagnosis. The diagnostic RDI and ACI values are both marked on each graph using dashed lines. In addition, the False Positive (FP) and False Negative (FN) regions are accordingly marked. It can be seen that as the reference RDI increases for the diagnosis. The ACI diagnosis produces larger FP errors.

The FP errors increase as the reference RDI increases. It can be said that, using a larger RDI value, such as 30, could cause under-diagnosis of the test subjects. Hence, the ACI could be more accurate than the RDI reference value utilised for the diagnosis. It can be said that the errors produced may be due to the high accuracy of the system. Table II provides a summary of the results obtained for the four experiments.

V. Conclusion

This paper proposed a novel algorithm for conducting non-contact OSA diagnosis. A new feature, called the Non-Gaussianity Index (NGI), was developed to perform a measure of non-Gaussianity. This feature was employed to mark possible apneaic episodes in analysed snore sound recordings of 86 subjects. A count of the apneaic episodes was then conducted to obtain the Apnea Count Index (ACI) used for the diagnosis. It was shown that the ACI could be optimised and adjusted to various diagnostic RDI values. Finally, it was seen that the ACI value corresponding to the diagnostic RDI value of 10 provided a detectability of 97.34% in conducting diagnosis of the 86 subjects.

TABLE II
Results Achieved and Variables Used Per Experiment

RDI	L	B Value	Detectability	ACI
5	25	0.54	92.00	7.61
10	**25**	**0.36**	**97.34**	**15.44**
15	30	0.21	96.67	32.98
30	30	0.15	86.30	58.00

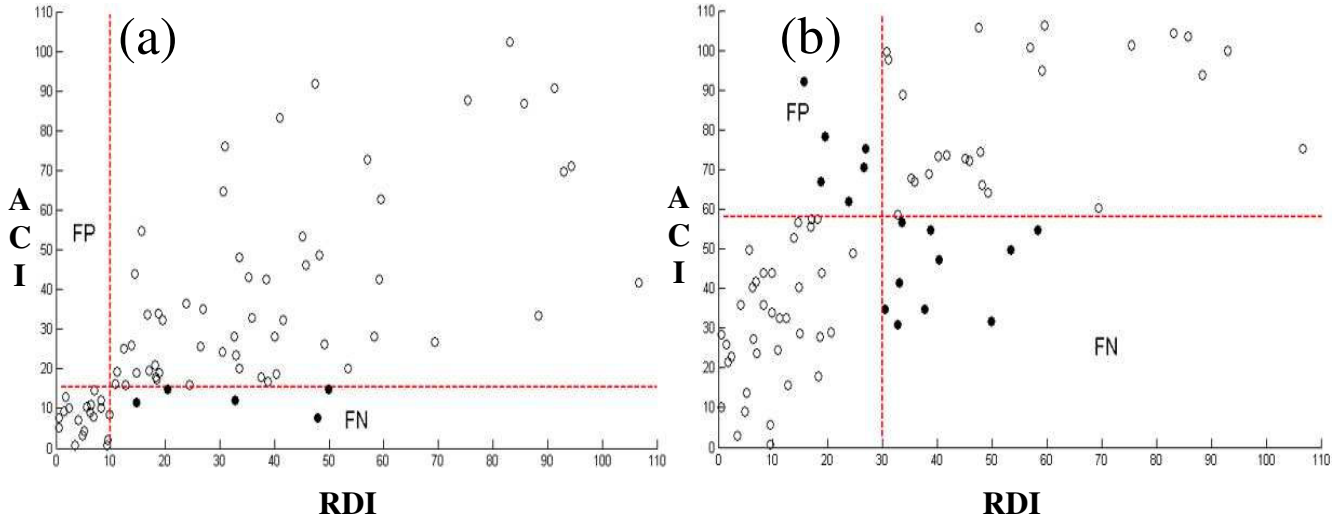

Fig. 1 (a) ACI versus RDI graph with horizontal dashed line marking (RDI=10) and vertical dashed line marking (ACI=15.44). (b) ACI versus RDI graph with horizontal dashed line marking (RDI=30) and vertical dashed line marking (ACI=58.00). The horizontal dashed lines mark the border region for the ground truth diagnosis using RDI. The vertical dashed lines mark the border region for the conducted diagnosis based on the proposed method using ACI. FP and FN represent the False Positive and the False Negative error regions respectively.

REFERENCES

[1] [T. Young, M. Palta, J. Dempsey, et al., "The occurrence of SDB among middle-aged adults," *N. Engl. J. Med.*, vol. 328, pp. 1230-1235, 1993.

[2] J. Durán, S. Esnaola, R. Rubio, A. Iztueta, "Obstructive Sleep Apneahypopnea and related clinical features in a population-based sample of subjects aged 30 to 70 yr," *Am J Respir Crit Care Med*, vol. 63, pp. 685-689, 2001.

[3] K. Banno, M. H. Kryger, "Sleep Apnea: Clinical investigations in humans, " *Sleep Med*, vol. 8, pp. 400-426, 2007.

[4] T. Young, M. Palta, J. Dempsey, J. Skatrud, *S.* Weber and *S.* Badr, 'The Occurrence Of Sleep- Disordered Breathing Among Middle-Aged Adults", *New England Journal of Medicine,* 328(17), 1993, pp. 1230-1235.

[5] American Academy of Sleep Medicine Task Force, "Sleep related breathing disorder in adults: Recommendations for syndrome definition and measurement techniques in clinical research", *Sleep* 1999; 22: 667-689.

[6] H. Nakano, M. Hayashi, E. Ohshima, N. Nishikata, T. Shinohara, "Validation of a new system of tracheal sound analysis for the diagnosis of sleep apnea-hypopnea syndrome,", *Sleep*, vol. 27, pp. 951-957, 2004.

[7] C. Guilleminault, R. Stoohs, T. Shiomi, C. Kushida, I. Schnittger, "Upper airway resistance syndrome, nocturnal blood pressure monitoring, and borderline hypertension," *Chest*, vol. 109, pp. 901-908, 1996.

[8] W. W. Flemons, M. R. Littner, J. A. Rowley, P. Gay, W. M. Anderson, D. W. Hudgel, R. D. McEvoy, D. I. Loube, "Home diagnosis of sleep apnea: a systematic review of the literature," *Chest*, vol. 124, pp. 1543-1579, 2003.

[9] R.D. Chervin, "Sleepiness, fatigue, tired, and lack of energy in obstructive sleep apnea," *Chest*, vol. 118, pp. 372-379, 2000.

[10] M. Cavusoglu, M. Kamasak, O.Erogul, T. Ciloglu, Y. Serinagaoglu, T. Akcam, "An efficient method for snore/non-snore classification of sleep sounds," *Physiological Measurement,* vol. 28, pp. 841-853, 2007.

[11] Duckitt W D, Tuomi S K and Niesler T R 2006 Automatic detection, segmentation and assessment of snoring from ambient acoustic data *Physiol. Meas.* 27 1047–56.

[12] J. Sola-Soler, R. Jane, J. Morera, "Spectral envelope analysis in snoring signals from simple snorers and patients with obstructive sleep apnea," in *Proc. IEEE 25th Annu. Intl. Conf. Eng. Med. Bio. Soc.*, vol. 3, pp. 2527-2530, 2003.

[13] A. S. Karunajeewa, U.R. Abeyratne and C. Hukins, "Silence-breathing-snore classification from snore-related sounds", *Physiological Measurement*, Institute of Physics, UK, vol. 29, pp.227-243, 2008.

[14] A.K. Ng, T.S. Koh, E. Baey, K. Puvanendran, "Diagnosis of obstructive sleep apnea using formant features of snore signals," in *Proc. IFBME World Congress Med. Phy. Biomed. Eng.*, vol. 14, pp. 861-864, 2006.

[15] Abeyratne U R, Patabandi C K K and Puvanendran K 2001 Pitch-jitter analysis snoring signals in the diagnosis of obstructive sleep apnea presented at the *23rd Ann. Int. Conf. IEEE Engineering in Medicine and Biology Soc. (IEEE EMBC2001) (Istanbul, Turkey)*

[16] Ng, A.K.; Wong, K.Y.; Tan, C.H.; Koh, T.S., "Bispectral Analysis of Snore Signals for Obstructive Sleep Apnea Detection," *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE* , vol., no., pp.6195-6198, 22-26 Aug. 2007.

[17] Gray, A., Jr.; Markel, J., "A spectral-flatness measure for studying the autocorrelation method of linear prediction of speech analysis," *Acoustics, Speech and Signal Processing, IEEE Transactions on* , vol.22, no.3, pp. 207-217, Jun 1974