35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Variations of Snoring Properties with Macro Sleep Stages in a Population of Obstructive Sleep Apnea Patients*

S. Akhter, U. R. Abeyratne, *Senior Member IEEE* and V. Swarnkar

Abstract— **Snoring is common in Obstructive Sleep Apnea (OSA) patients. Snoring originates from the vibration of soft tissues in the upper airways (UA). Frequent UA collapse in OSA patients leads to sleep disturbances and arousal. In a routine sleep diagnostic procedure, sleep is broadly divided into rapid eye movement (REM), non-REM (NREM) states. These Macro-Sleep States (MSS) are known to be involved with different neuromuscular activities. These differences should influence the UA mechanics in OSA patients as well as the snoring sound (SS). In this paper, we propose a logistic regression model to investigate whether the properties of SS from OSA patients can be separated into REM/NREM group. Analyzing mathematical features of more than 500 SS events from 7 OSA patients, the model achieved 76% (± 0.10) sensitivity and 75% (± 0.10) specificity in categorizing REM and NREM related snores. These results indicate that snoring is affected by REM/NREM states and proposed method has potential in differentiating MSS.**

I. Introduction

Obstructive sleep apnea (OSA) is a highly prevalent disease among adults [1] in which upper airway (UA) collapses during sleep. A complete UA collapse is termed as apnea while a partial collapse is known as hypopnea [2]. Frequent UA collapse and associated arousals can seriously disrupt the overall sleep architecture of a patient. Fatigues, day time sleepiness, lack of concentrations are most common diurnal symptoms while snoring, chocking, gasping are the common nocturnal symptoms of OSA [1]. OSA is a serious health concern as it increases risks of developing cardiovascular disease, diabetes, stroke and neuro-cognitive deficits [1].

Current reference technique for OSA diagnosis is Polysomnography (PSG). The test monitors sleep by recording a range of neuro-physiological and cardio-respiratory signals throughout the night. The main outcomes of PSG test are OSA severity measures such as the apnea-hypopnea Index (AHI) and the Arousal Index (AI).

PSG also provides information about the Macro-Sleep Architecture (i.e. the temporal course of Rapid-Eye-Movement (REM) sleep and non-REM (NREM) sleep [2]) of the patient and subsequent diagnostic measures in terms of REM and NREM. This could provide more detail about the quality of sleep that is unavailable via the overall indices.

* This work was partially supported by the Australian Research Council under grant DP120100141 to Dr Abeyratne.

Akhter S, Abeyratne U R and Swarnkar V are with the School of ITEE, The University of Queensland, 4072, Australia.
(e-mail: udantha@itee.uq.edu.au)

In routine sleep diagnostic procedures, MSS is done manually by applying complex and visual scoring rules simultaneously on multiple electrophysiological signals (e.g. EEG, EOG, EMG) [2, 3]. The process is subjective, time-consuming, tedious and costly. In general, due to the need for technical expertise and complexity, MSS scoring is not included in devices targeting for mass screening of OSA.

Snoring is one of the earliest common symptoms of OSA. Snoring originates from the vibration of the soft tissues (e.g. tongue, soft palate, pharyngeal wall) in the UAs [4]. UA muscle activations [5, 6] and cross-sectional area [5] are reported to be varied with REM and NREM sleep. Muscle tone variation with MSS states is one of the major reasons for the change of acoustical properties of UA (and hence snore characteristics).

Recent studies reveal that properties of snoring sound carry vital information related to the UA collapse which has potential in characterizing OSA/Non-OSA [7-9]. However a very few researchers have looked at the effect of MSS on snore. To the best of our knowledge only one group of researchers have attempted [10] to derive MSS specific information from snores of OSA patients. The group investigated the duration and peak spectral components (100-300 Hz) of snores from contact tracheal microphone on 7 OSA patients. The study reported in [10] was limited to a presentation of descriptive statistics of snores from known sleep states where the results were not validated on a prospective dataset. Moreover, snore sounds have bandwidth extending beyond 10 kHz while the bandwidth of a tracheal recording is far below this.

In this context, we hypothesise that variation in UA muscle activities due to MSS states are embedded in snore sounds and snore features can be used to identify these states.

In this paper we developed a novel method based on logistic regression (LRA) model to explore this hypothesis. We extracted characteristics features of snore sounds recorded from non-contact microphones and applied the LRA model to learn and validate the properties of REM and NREM related snores. The overall methodology followed in this paper is described in the next Section.

II. Method

A. Data Acquisition

Data acquisition environment for the work of this paper was the Sleep Diagnostic Laboratory of The Princess

Alexandra Hospital, Brisbane, Australia. Our subject population includes patients referred to the hospital for routine PSG test and an informed consent was made.

Clinical PSG equipment (Siesta, Compumedics®, Sydney, Australia) was used for the standard set-up protocols in [2]. Overnight (6-10 hours) audio recordings were carried on simultaneously with PSG through a separate high fidelity computerized data acquisition system. The sound recording system consists of a matched pair of low-noise free-field microphones having a hypercardoid beam pattern (Model NT3, RODE, Sydney, Australia). A professional quality pre-amplifier and A/D converter unit, (Model Mobile-Pre USB, M-Audio, California, USA) was used for data acquisition with a sampling rate of 44.1 KHz and 16 bit/sample resolution. The nominal distance from the microphones to the mouth of the patient was 50cm.

Standard sleep scoring rules [3] and AASM guidelines [2] were followed for the diagnosis. We denoted full night audio data as snore related sound (SRS) recordings. For each patient, we collected SRS along with the complete PSG dataset and standard diagnostic reports from the hospital.

B. SRS Data Pre-processing

Our focus is the effects of MSS states on snore properties. In addition to sleep stages, body position can also affect the UA during sleep, due to the effects of gravity on the UA mechanics. In this paper, to control any position related effects, we limited our study to snore sounds associated with supine sleep.

A typical snore sound consists with inspiratory and expiratory phases. During the inspiratory phase of breathing there is a stronger tendency for airway narrowing/collapses due to the negative pressures developed within the UA. This should provide an ideal opportunity to observe the balance between negative pressures and muscle activities that keep the airway open. In this paper, we considered only the inspiratory part of snore to capture this dynamic state of the UA during sleep.

Overnight SRS recording in the hospital may contain snore sounds, breathing, speech sounds from the patient and also background and electrical noises. To avoid speech and other sounds, we selected snore samples manually from the SRS during NREM and REM sleep in supine position. Then the samples were passed through a high pass filter (5^{th} order Butterworth filter with cut off frequency at 20 Hz) to remove the low frequency noises from the recording environment.

We used arithmetic mean of the signals recorded from the two microphones for further analysis. Let s(n) represents the mean inspiratory part of n^{th} snore sample, where n = 1,2,3….N. N represents the total number of collected snores from P patients. We segmented each s(n) into k number of equal sized blocks where $s_k(n)$ indicates the k^{th} sub-block.

C. Feature Extraction

1) Pitch Period

In speech processing [11], pitch is the fundamental frequency of speech signal. Considering the similarities between snore sound and speech [8], sleep stage related muscle tone reduction in the UA with increased inspiratory effort should have modification on the fundamental frequency of snore. Pitch Period (PP) in OSA patients is found to carry vital information about the UA collapse [7]. We computed the PP of $s_k(n)$ using autocorrelation with centre clipping algorithm as described in [12].

2) Formant Frequencies

Formants are the resonance frequencies of the vocal tract [11]. Different formant are considered to represent the functionality of different sections of vocal tract (i.e. F1-F3 corresponds to the pharyngeal constriction, tongue advancement and lip-rounding respectively [11]). In this paper, we considered F1 and F2 of snores in our feature set and used Linear predictive coding (LPC) scheme with the Yule–Walker autoregressive parameter estimation method in [13] to estimate the formants of $s_k(n)$.

3) Non-Gaussianity Score

A typical snore sound might have voiced, silence and unvoiced components. Voiced component of snore is the pseudo periodic sequences with detectable pitch [14] while unvoiced component is the aperiodic sequences without detectable pitch. Change in the UA muscle activity during MSS states should vary the nature of vibration. Non-Gaussianity Score (NGS) provides a measure of deviation from the Gaussian distribution of data. We employed normal probability plot γ of snore segment $s_k(n)$ to compute the deviation from the reference Gaussian probability δ using (1). In (1), ψ_k is the NGS of $s_k(n)$ and L is the length of $s_k(n)$. The details of NGS can be found in [15].

$$\psi_k = 1 - \left(\frac{\sum_{i=1}^{L}(\delta[i]-\bar{\delta})^2}{\sum_{i=1}^{L}(\gamma[i]-\bar{\gamma})^2}\right) \quad (1)$$

4) Log Energy

Voiced or unvoiced components of snore sound is considered to differ in terms of energy content [14]. Accordingly, energy of snore segment will differ in REM and NREM sleep. We used (2) to compute the Log Energy (LogE) of $s_k(n)$. In (2), L is the length of $s_k(n)$ and ε is an arbitrary constant to avoid any computation of log 0.

$$LogE_k = 10\log_{10}\left(\frac{1}{L}\sum_{i=1}^{L}(s_k^i(n))^2 + \varepsilon\right) \quad (2)$$

At this stage, if P patients have N^{NR} snore samples from NREM sleep and N^{R} snore samples from REM sleep, then the size of the feature matrix became $N^{NR} \times k \times f$ for NREM and $N^{R} \times k \times f$ for REM. Where 'k' is the snore sub-blocks and 'f' represents five features computed in this section. These two feature matrices were then used for classification of REM and NREM related snores with the help of a LRA model. LRA is a generalized linear model, which uses several independent predictors (features) to estimate the probability (Y) of a categorical event (dependent variable). In our work, the dependent variable Y is assumed to be equal to 'zero' ($Y = 0$) for REM snore samples and 'one' ($Y = 1$)

for NREM snore samples. The model used regression function to estimate the probability of Y from the independent variables (i.e. features) as below:

$$Pr\,ob(Y = 1|_{f_1, f_2, \ldots.. f_F}) = \frac{e^z}{e^z + 1} \quad (3)$$

$$z = \beta_0 + \beta_1 f_1 + \beta_2 f_2 + \ldots\ldots\ldots + \beta_F f_F \quad (4)$$

In (3) and (4) f_1, f_2 …f_F are the elements of feature vector (independent variables), β_0 is called the intercept and β_1, β_2 and so on are called the regression coefficient of independent variables obtained from the Training set. We used the Receiver-Operating Curve (ROC) analysis on the Training model to select the optimal decision threshold λ from Y (that the snore is from NREM sleep if Y is above λ otherwise REM) to maximize the performance. We calculated the performance of the LRA model in terms of area under the ROC (AUC), sensitivity and specificity of the differentiation of REM and NREM snores from the ROC.

III. Results

A. Clinical Database

For the work of this paper we analyzed N = 546 snore samples from P = 7 patients SRS data recordings. Out of N = 546 events N^{NR} = 391 events were from NREM sleep and N^R = 155 were from REM sleep. Table 1 presents the demographic details of each patients and the record of collected REM/NREM snore samples. We varied number of sub-blocks k from 1 to 10.

B. Classification and validation of REM/NREM Snores

We developed a LRA model to classify snores into subsequent REM and NREM stages and used Receiver-Operating Curve (ROC) analysis to calculate the sensitivity and specificity of classification performance. We used the snore features computed in Section 2 for two cases of classification: (i) within-patient and (ii) across-patient.

1) Case 1: Within-Patient

REM and NREM related snore samples from each individual patient were considered for mutually exclusive Training and Testing set.

TABLE I. Patient Details and Collected Snore Samples (P1-P7 = Patient No., BMI = Body Mass Index, RDI = Respiratory Disturbance Index, NR = NREM, R = REM, S^{NR} = NREM Sleep, S^R = REM Sleep, N^{NR} = NREM Samples, N^R = REM Samples)

Details	P1	P2	P3	P4	P5	P6	P7
Age	51	51	50	61	46	29	27
BMI	30.8	31.8	42.2	35.6	40.9	36.8	45.5
RDI	11.4	13.8	19.5	30.8	33.0	48.2	94.4
R AHI	56.5	28.0	40.9	54.1	61.9	18.8	124.0
NR AHI	10.3	9.3	15.9	26.9	29.4	56.1	87.4
$\%S^{NR}$	97.5	76.1	85.4	84	87	78.8	80.9
$\%S^R$	2.5	23.9	14.6	14.0	11.3	21.1	19.2
N^{NR}	68	69	31	39	50	26	108
N^R	32	27	15	12	7	20	42

Training set consists of 50% of N^R samples and 50% of N^{NR} samples from a patient and the Testing set included the remaining 50% of N^R and N^{NR} samples of that patient. Later we applied the feature matrix of the Training and Testing set to the LRA model respectively for learning and cross-validation.

Classification results for this case are presented in Table 2. Number of sub-blocks k varied from 1 to 2. It can be observed in Table 2 that the Training performance can achieve 100% sensitivity, specificity and AUC for most of the patients. This performance remains steady as k increased to 2. For the Testing Network in Table 2, performance varied from 70-100%. Low performance with increasing k in the Testing Network might be from the increase in size of the feature matrix with k.

2) Case 2: Across-Patient

In this case, we followed the *Leave-One-Out* cross validation (LOOCV) technique to validate the LRA model. LOOCV included all features of N^R and N^{NR} samples from 6 out of 7 patients for Training and the features from the 7^{th} one for Testing. This was repeated for 7 times so that each of the 7 patients was used for validation at least once.

TABLE II. Summary of Performance Matrices of Training and Testing Network for each Patients with Variable Segment Length (k = number of sub-blocks, AUC = Area under the ROC curve, Sen = Sensitivity and spec = specificity of ROC Curve)

Data Set	k = 1							k = 2						
	Patient No.							Patient No.						
	1	*2*	*3*	*4*	*5*	*6*	*7*	*1*	*2*	*3*	*4*	*5*	*6*	*7*
Training Set Performance														
AUC	85	100	98	100	100	100	95	99	100	100	100	100	100	96
Sensitivity	81	100	100	100	100	100	90	100	100	100	100	100	100	95
Specificity	76	100	87	100	100	100	89	97	100	100	100	100	100	94
Testing Set Performance														
AUC	72	98	98	90	87	92	92	84	98	91	87	76	74	84
Sensitivity	69	100	100	83	75	90	86	81	93	100	83	75	80	86
Specificity	71	94	94	85	76	85	87	79	91	88	70	64	85	81

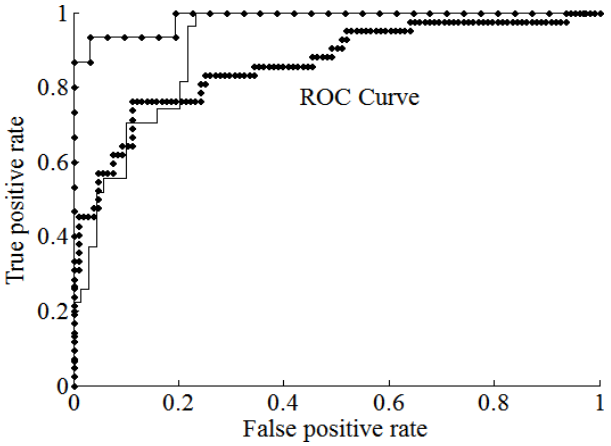

Figure 1. ROC Curves of Testing Dataset in Case 2 from LRA model. This figure represents the ROC of LOOCV for three out of 7 different set of combinations of the patients.

Fig. 1 shows the classification results for k=1. Table 3 represents the mean and standard deviation of AUC, sensitivity and specificity as calculated from the ROC curves. We varied k from 1 to 10. It is to be noted in Table 3 that maximum performance in the Training Network is at k=10 while for Testing Network it is at k=1. It indicates that more details about the characteristics of snore might be obtained by increasing number of sub-blocks k. However, this could affect the performance of the Testing Network as the size of the feature matrix increased k times compared to the number of samples being used. Overall the performance of the classifier reached to 76%/75% sensitivity/specificity and 83% AUC for k = 1.

IV. Conclusion

In this paper, we developed an LRA model and showed the possibility that the effect of REM/NREM sleep states embedded into the properties of snore sound can be utilized to classify MSS states. The results indicate that the model can achieve up to 85%-90% (sensitivity and specificity) for each individual patient. Both the sensitivity and specificity drops to 75% when computed across a group of patients.

The proposed approach illustrates the potential for extracting macro-sleep staging information from OSA patients based on snoring. It may find valuable use in snore-sound based non-contact population screening/monitoring devices for OSA. The method requires further validation on larger clinical datasets and across different sleep positions.

References

[1] T. L. Lee-Chiong and C. A. Polnitsky, "Sleep Breathing Disorders," in *Review of Sleep Medicine (Second Edition)*, pp. 43-65, 2007.

[2] C. Iber, S. A. Israel, A. L. Chesson, and S. F. Quan, Eds., *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. 2007.

[3] A. Rechtschaffen and A. Kales, Eds., *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects*. 1968.

[4] F. Dalmasso and R. Prota, "Snoring: analysis, measurement, clinical implications and applications," *European Respiratory Journal,* vol. 9, pp. 146-159, 1996.

[5] A. I. Pack, *Sleep Apnea: Pathogenesis, Diagnosis, and Treatment*: Marcel Dekker, 2002.

[6] A. S. Jordan, A. Malhotra, D. P. White, Y. L. Lo, A. Wellman, D. Eckert, S. Yim-Yeh, M. Eikermann, S. Smith, and K. Stevenson, "Airway dilator muscle activity and lung volume during stable breathing in obstructive sleep apnea," *SLEEP,* vol. 32, pp. 361-368, 2009.

[7] S. d. Silva, U. R. Abeyratne, and C. Hukins, "A method to screen obstructive sleep apnea using multi-variable non-intrusive measurements," *Physiological Measurement,* vol. 32, p. 445, 2011.

[8] A. S. Karunajeewa, U. R. Abeyratne, and C. Hukins, "Multi-feature snore sound analysis in obstructive sleep apnea-hypopnea syndrome," *Physiol Meas,* vol. 32, pp. 83-97, 2011.

[9] U. R. Abeyratne, S. d. Silva, C. Hukins, and B. Duce, "Obstructive sleep apnea screening by integrating snore feature classes," *Physiological Measurement,* vol. 34, pp. 99, 2013.

[10] H. Nanako, T. Ikeda, M. Hayashi, E. Ohshima, and A. Onizuka, "Effects of Body Position on Snoring in Apneic and Nonapneic Snorers," *SLEEP,* vol. 26, pp. 169-172, 2003.

[11] W. R. Zemlin, *Speech and Hearing Science: Anatomy and Physiology*, 3rd ed.: Prentice Hall, 1988.

[12] M. Sondhi, "New methods of pitch extraction," *IEEE Transactions on Audio and Electroacoustics,* vol. 16, pp. 262-266, 1968.

[13] J. Markel, "Digital inverse filtering-a new tool for formant trajectory estimation," *Audio and Electroacoustics, IEEE Transactions on,* vol. 20, pp. 129-137, 1972.

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

[15] H. Ghaemmaghami, U. R. Abeyratne, and C. Hukins, "Normal probability testing of snore signals for diagnosis of obstructive sleep apnea," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5551-5554, 2009.

TABLE III. Summary of Performance Matrices of Training and Testing Network for the OSA Patient Group

Data Set	k = 1	k = 2	k = 3	k = 4	k = 5	k = 6	k = 7	k = 8	k = 9	k = 10
Training Set Performance										
AUC Mean (Std)	82 (2)	84 (2)	87 (2)	86 (2)	89 (1)	88 (1)	89 (1)	90 (1)	90 (1)	90 (1)
Sensitivity Mean (Std)	76 (3)	77 (2)	81 (2)	80 (1)	83 (2)	81 (2)	82 (2)	83 (2)	83 (2)	84 (1)
Specificity Mean (Std)	75 (3)	77 (2)	81 (1)	80 (1)	83 (2)	81 (1)	82 (2)	83 (1)	83 (2)	83 (1)
Testing Set Performance										
AUC Mean (Std)	83 (12)	73 (9)	75 (10)	69 (12)	72 (10)	70 (10)	70 (11)	69 (11)	69 (11)	69 (13)
Sensitivity Mean (Std)	76 (10)	69 (5)	68 (9)	66 (9)	70 (9)	66 (8)	67 (9)	66 (9)	68 (8)	66 (7)
Specificity Mean (Std)	75 (10)	67 (6)	68 (8)	65 (9)	67 (8)	66 (8)	64 (8)	66 (9)	65 (8)	65 (7)