Severity Analysis of Upper Airway Obstructions: Oesophageal Pressure Versus Snoring Sounds

Mrunal Markandeya
School of Information Technology and Electrical Engineering
The University of Queensland
Brisbane, Australia
mrunal.markandeya@gmail.com

Udantha R. Abeyratne
School of Information Technology and Electrical Engineering
The University of Queensland
Brisbane, Australia
udantha@itee.uq.edu.au

Roneel V. Sharan
Australian Institute of Health Innovation
Macquarie University
Sydney, Australia
roneel.sharan@mq.edu.au

Craig Hukins
Department of Respiratory and Sleep Medicine
Princess Alexandra Hospital
Brisbane, Australia
craig.hukins@health.qld.gov.au

Brett Duce
Department of Respiratory and Sleep Medicine
Princess Alexandra Hospital
Brisbane, Australia
brett.duce@health.qld.gov.au

Karen McCloy
School of Information Technology and Electrical Engineering
The University of Queensland
Brisbane, Australia
k.mccloy@uq.net.au

***Abstract*—Obstructive sleep apnea (OSA) is a sleep related breathing disorder. Identifying severity of airway obstruction is important in OSA severity analysis as well as for treatment success. The apnea hypopnea index (AHI), defined as the total number of full and partial upper airway obstructions per hour, is widely used to diagnose and characterize the severity of OSA. However, recent research shows that AHI provides a crude summary of overnight dynamics of upper airway obstructions. Oesophageal pressure manometry (Pes) is the gold standard method for identifying the severity of individual airway obstruction but, due to the invasive nature, it is less commonly used in sleep laboratories. There is a need for simple automated technology to characterize the severity of airway obstruction. In this work, we propose a method to classify the severity of airway obstruction by analyzing snoring sounds collected through an iPhone 7 smartphone, which requires no physical contact with a subject. For the development of methods, we segmented more than 2000 snoring sound epochs of 5 seconds duration from 7 patients undergoing a polysomnography (PSG) along with Pes. Based on Pes data, we labelled snoring epochs as mild, moderate or severe airway obstruction. We extracted audio features from snoring epochs and used them to train a classifier for multiclass classification. Using 10-fold cross-validation, our methods achieved average accuracy greater than 80% in classifying the severity of airway obstructions. Our results indicate the feasibility of snoring sound in characterizing the severity of airway obstructions. Our non-contact, snoring sound-based technology has the potential to develop into an automatic individual airway obstruction severity analysis system.**

Keywords—snoring sounds, oesophageal pressure manometry, obstructive sleep apnea

I. Introduction

Obstructive sleep apnea (OSA) is a serious sleep related breathing disorder affecting 9% to 38% of the general population [15]. The repetitive upper airway obstruction is a characteristic of OSA. A full closure of the upper airway is known as apnea; while partial closure is known as hypopnea. The total number of both such events per hour is known as the apnea-hypopnea index (AHI). Currently, the diagnosis and classification of OSA are based on OSA-related clinical symptoms and AHI.

Recent research shows that AHI is a crude summary of the overnight dynamics of upper airway obstructions [13], [16]. The 'all or nothing approach' of the AHI fails to detect the severity of these individual upper airway obstructions. Upper airway obstructions which fall short of tight time based and threshold based criteria of apnea and hypopnea events are completely ignored by the AHI [13]. The impact of severity of an individual obstructive event on morbidity and mortality [12] related to OSA suggests that there is a need to identify the severity of an individual obstructive event.

Respiratory effort is the gold standard measure of airway obstruction severity and estimated using oesophageal pressure manometry [3]. It represents total inspiratory muscle activity. Thus, the inspiratory effort increases with an increase in severity of obstruction and the oesophageal pressure become more negative. Although oesophageal pressure manometry identifies and quantifies respiratory effort, it is less commonly used in routine polysomnography (PSG) due to its invasive nature [18].

Snoring is one of the commonest symptoms of OSA. Snoring sounds are generated due to the vibrations of collapsing soft tissues of the upper airway. During a complete collapse of the airway, a large pressure differential builds up around the site of the collapse. In the process of the airway reopening, this differential drives a rapid stream of air through a narrow opening, causing higher frequency snoring sounds carrying information about the event. Additionally, depending on the level of closure of the airway during obstructive events, it will generate snoring sounds carrying the signatures of those events. As such, snoring sound analysis should reveal significant information about the severity of airway obstruction.

Earlier studies, such as [1],[2],[4],[5], mostly focus on considering snoring sound as statistical entities and machine classifying patients into different AHI bands. In our previous work [10], based on snoring sounds we developed a method to characterize upper airway obstructions. However, we only validated our results in characterizing full and partial collapse of the upper airway using polysomnography.

Research supported by the National Health and Medical Research Council (NHMRC), Australia under grant #: APP1126600.

978-1-5090-0617-5/19/$31.00 ©2019 IEEE

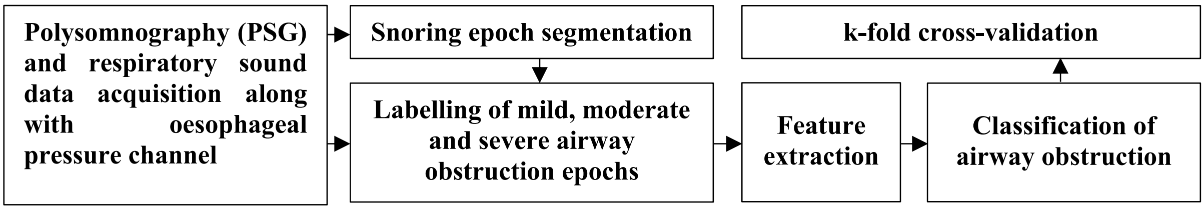

Fig. 1. Block diagram of the method under study. Our method consists of three main stages: (1) identification and labelling of snoring sound epochs as mild, moderate or severe airway obstruction with reference to oesophageal pressure channel, (2) the extraction of features and classifier design, and (3) performance validation of the model to classify individual snoring epochs based on airway obstruction severity.

This study aims to analyze snoring sound from OSA patients and explore their efficacy in classifying the severity of airway obstruction. Our approach is to acquire simultaneous oesophageal pressure manometry and snoring sound data from OSA patients and develop a classifier to predict the severity of an individual obstructive event. To the best of our knowledge, this is the first attempt to analyze snoring sound to automatically characterize severity of airway obstruction with reference to the gold standard oesophageal pressure manometry.

II. Methodology

The overall approach we propose in this paper is summarized in Fig. 1 and the details provided in the following subsections.

A. Data acqusition protocol

The data acquisition environment for this research was the Sleep Disorder Centre of the Princess Alexandra Hospital in Brisbane, Australia. For PSG data, the sleep data acquisition system (Model Siesta, Compumedics®, Sydney, Australia) was used. Both oral and written consent was taken from the subjects as approved by the Human Ethics Committees of Princess Alexandra Hospital and The University of Queensland.

Overnight respiratory sounds were recorded using a smartphone (Model - iPhone 7, Apple Inc.). The smartphone was placed at a distance of about 40–70 cm from the subject's mouth. The sampling frequency was kept at 44.1 kHz with 16-bit resolution.

Oesophageal pressure monitoring is performed as part of an overnight PSG recording for diagnostic studies. An air-filled balloon catheter was inserted after one nostril was anaesthetized with 3-5ml of viscous lignocaine. The balloon catheter was passed by a technician into a selected nostril. The distance at which the catheter was positioned in the oesophagus was calculated based on the patient's height.

For airway obstruction severity analysis, delta oesophageal pressure (ΔPes) is determined as the peak to trough difference in the waveform [17]. Fluctuations of ΔPes of less than 10 cmH_2O are considered normal [9]. As there is no agreement in the literature, cut-off values for oesophageal pressure are largely based on clinical experience [9]. For the present study, we are considering ΔPes between 10-20 cmH_2O as mild airway obstruction, ΔPes between 20-30 cmH_2O as moderate airway obstruction, and ΔPes > 30 cmH_2O as severe airway obstruction.

B. Snoring sound identification and labelling

For characterizing the severity of upper airway obstruction, we selected snoring epochs from overnight respiratory sound data. Our group previously defined snoring objectively, based on the presence of detectable pitch in the audio signal. The overnight respiratory sound data was divided into non-overlapping, five-second respiratory epochs. We manually extracted snoring epochs as respiratory epochs which contained the partial or full presence of observable periodicity. Snoring epochs were extracted by listening to all respiratory sound data and simultaneously visually observing their time-domain waveforms on Adobe Audition CS6 software.

C. Feature extraction and classifier design

Step 1: In this step, our aim was to extract audio features from the snoring epochs to be used in the classification algorithm for determining the severity of upper airway obstruction. Drawing inspiration from pathophysiologic changes during upper airway obstruction as well as from the success of past studies, we computed the following 30 features from each snoring epoch: 3 × *R-index*, 2 × formants, 1 × non-Gaussianity index, 12 × Mel-frequency cepstral coefficients (MFCCs), and 12 × derivatives of MFCCs. Details on the extracted features are given below.

- *R-index* – feature developed in our previous work [10] based on the premise that the energies of snoring sounds shift to the higher frequencies due to narrowing of the upper airway. For this purpose, we divided the snoring epoch into 'n' equal parts. For each part, we calculated the total energy (P_t), the average energy of the frequency band (P_b), and the average percentage of the total energy (R) by using power spectral density. The *R-index* for a given frequency band was obtained as the difference between the ratio of the maximum to the minimum average energy and the ratio of the percentage of the total energy at the locations of the average energy. The values of *R-index* focus on the maximum energy content of a given frequency band. In this paper, we calculated 3 - *R-index* features, from 5000-10000Hz, 10000-15000Hz, and 15000- 22000Hz bands.

- Formants – formants are the resonance frequencies of the vocal tract. The size and shape of acoustic spaces in the vocal tract and their coupling indicate corresponding formants. The first formant is known to represent pharyngeal constriction, whereas the second formant is related to the degree of advancement of tongue [19]. For each snoring epoch we computed 2 formants. We used linear predictive coding (LPC) scheme based on the Yule–Walker autoregressive parameter estimation [11] for this purpose.

- Non-Gaussianity index (NGI) – represents the deviation of probability plot of analyzed data to its reference Gaussian probability plot [7]. Turbulent airflow through collapsing upper airways causes deviation in detectable periodicity within snoring epochs. We anticipated that variation in the severity of upper airway obstruction cause change in the periodic pattern of snoring sounds and hence

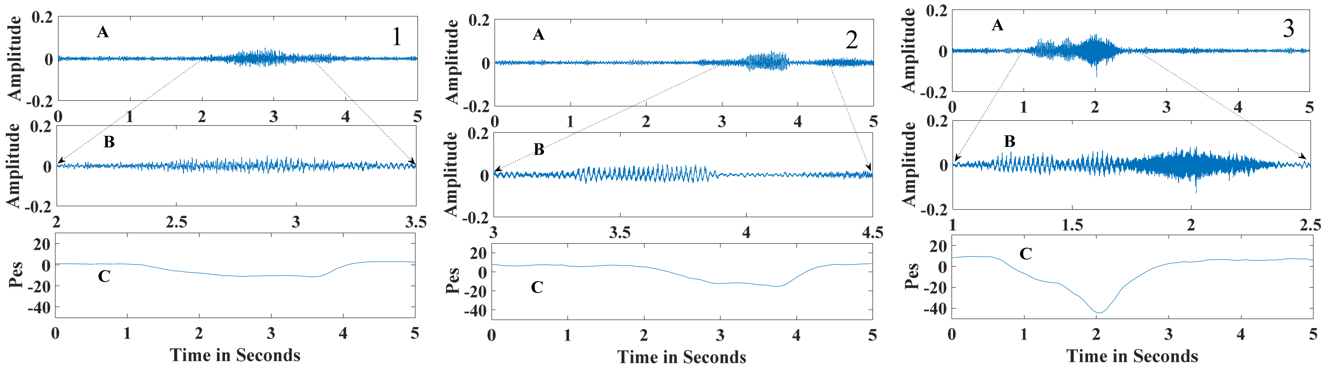

Fig. 2. Illustration of snoring sound epochs along with the corresponding oesophageal pressure signal. As airway obstruction severity increases, esophageal pressure becomes more negative. Column 1 – mild airway obstruction, column 2 – moderate airway obstruction, and column 3 – severe airway obstruction. A – 5-second snoring epoch, B – an expanded version of the snoring epoch, and C – oesophageal pressure signal.

Gaussianity deviates accordingly. To obtain this index, let $x[n]$ represent the discrete-time snoring epoch signal. We first computed the inverse of the normal cumulative distribution function (CDF) for the data using (1) as follows:

$$\gamma = F^{-1}(p|\mu,\alpha) = \{\gamma : F(\gamma|\mu,\alpha) = p\} \tag{1}$$

where

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

μ = mean of $x[n]$ and σ = standard deviation of $x[n]$ in (2). We then assigned probabilities p in (1) to obtain the values of γ to plot the normal probability plot of $x[n]$. The deviation of the probability plot of the analyzed data (γ) to its reference Gaussian probability plot (g) was referred as NGI. NGI was calculated using linear regression using (3) as:

$$NGI = 1 - \left(\frac{\sum_{i=1}^{N}(g[i] - \bar{g})^2}{\sum_{i=1}^{N}(\gamma[i] - \bar{\gamma})^2}\right) \tag{3}$$

- Mel-frequency cepstral coefficients (MFCCs) – this feature set is known to effectively represent vocal tract configuration. The human ear resolves sound frequencies non-linearly across the audio spectrum. Considering the similarities between snore sound and speech generation, several snoring sound analysis studies [1], [8] show a high level of performance accuracy using MFCC feature set. In computing MFCC, the snoring epoch was split into 30ms frames. For each frame we calculated 12 MFCCs and 12 differentials MFCCs, using a triangular Mel-filter bank. We then calculated mean of each coefficient across the epoch, thus getting 24 mean MFCC based features for each snoring epoch.

Step 2: For classification model development, we used a support vector machine (SVM). SVM is a binary classifier but has been shown to be useful in various multi-class classification problems using various strategies. The classifier was trained to learn the characteristics of mild, moderate, and severe airway obstruction classes. The validation of the classifiers was performed using 10-fold cross-validation.

For classifying the severity of upper airway obstructions in three classes, we compared the one-vs-all and one-vs-one multi-class classification techniques. In one-vs-all approach, the classifier with the highest probability assigns the class label whereas in one-vs-one approach the class which has been predicted most assigns the class label.

SVM algorithm separates the training data in the feature space by a hyperplane defined by the type of kernel function used. In this paper, we used a Gaussian Radial Basis Function (RBF) kernel. For RBF kernel, we need to determine two parameters to train the SVM model [14]. These are the penalty parameter 'c' and the kernel parameter 'γ'. The parameter 'c' determines the tradeoff between the minimization of the fitting error and the maximization of between-class margin. The parameter 'γ' determines the bandwidth of RBF. For optimizing this pair ['c', 'γ'], we used grid search using cross-validation method, as it performs exhaustive parameter search over the parameter space.

III. Results

Table I summarizes the demographic details of the subject population. Our database comprises 2104 snoring epochs segmented from 7 subjects (3-male, 4-female) referred to the sleep diagnostic center.

Fig. 2 shows the snoring epochs along with their expanded versions and oesophageal pressure signals. It can be observed that the duration and frequency of periodic peaks in snoring signals are increasing with an increase in the severity of airway obstruction. This supports our hypothesis that snoring sound carries information on the severity of upper airway obstruction.

To characterize the severity of upper airway obstruction, 2104 epochs were labelled into mild, moderate, and severe classes as described in section II-A resulting in 645, 467, and 992 epochs, respectively. As the minimum average duration per breath for an adult during rest is 5 seconds [6], we selected this as the epoch length in this work.

The classification accuracy values for each class using one-vs-all and one-vs-one multiclass-class classification techniques are shown in Table II, in the form of a confusion matrix. The airway obstruction severity prediction is

TABLE I. SUBJECT DEMOGRAPHIC AND CLINICAL DATA

	Age (years)	BMI (kg/m^2)	RDI
Male (3)	46.3±12.07	29.1±5.96	39.3±33.19
Female (4)	53.3±8.84	36.0±9.80	28.3±27.15

TABLE II. CLASS CONFUSION MATRIX IN 10-FOLD CROSS-VALIDATION USING SVM

One-vs-all – Confusion matrix values given in % as mean (standard deviation)			
	Predicted – Mild	**Predicted – Moderate**	**Predicted - Severe**
Labelled – Mild	88.22 (3.86)	10.07 (3.74)	1.70 (1.35)
Labelled – Moderate	22.90 (4.67)	58.69 (4.68)	18.40 (5.78)
Labelled – Severe	0.90 (1.11)	5.03 (3.39)	94.05 (3.63)
One-vs-one – Confusion matrix values given in % as mean (standard deviation)			
	Predicted – Mild	**Predicted – Moderate**	**Predicted - Severe**
Labelled – Mild	96.27 (1.97)	3.26 (2.13)	0.46 (0.75)
Labelled – Moderate	14.79 (4.53)	85.21 (4.53)	0 (0)
Labelled – Severe	0.71 (1.07)	0 (0)	99.29 (1.07)

compared with oesophageal pressure manometry. The results are reported in percentage (%) as the mean and standard deviation in 10-fold cross-validation. An average classification accuracy of 80% is achieved using one-vs-all multi-class classification technique and 93% using ove-vs-one multi-class classification technique.

The results presented in the paper were obtained with a cross-validation approach on small dataset (2104 epochs from 7 subjects). Results require further validation on larger, independent data set.

IV. CONCLUSION

Our results suggest the feasibility of snoring sounds to characterize the severity of upper airway obstruction when compared against the gold standard oesophageal pressure manometry. Our non-contact, smartphone-based method is more convenient than the existing methods and has the potential to develop into an automated technology for characterizing the severity of upper airway obstruction.

ACKNOWLEDGMENT

The authors would like to acknowledge Mr. Alexander Ryan from Princess Alexandra Hospital, Brisbane, Australia for his help with clinical data acquisition.

REFERENCES

[1] Akhter, S., U. R. Abeyratne, V. Swarnkar and C. Hukins (2018). "Snore Sound Analysis Can Detect the Presence of Obstructive Sleep Apnea Specific to NREM or REM Sleep." J Clin Sleep Med 14(6): 991-1003.

[2] ankışhan, H. and D. Yılmaz (2013). Comparison of SVM and ANFIS for Snore Related Sounds Classification by Using the Largest Lyapunov Exponent and Entropy.

[3] Berry, R., R. Brooks, C. Gamaldo, S. Harding, C. Marcus and B. Vaughn (2012). Darien, IL: American Academy of Sleep Medicine; 2012. The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications, version 2.0.

[4] Collop, N. A., S. L. Tracy, V. Kapur, R. Mehra, D. Kuhlmann, S. A. Fleishman and J. M. Ojile (2011). "Obstructive sleep apnea devices for out-of-center (OOC) testing: technology evaluation." J Clin Sleep Med 7(5): 531-548.

[5] Emoto, T., U. Abeyratne, T. Shono, R. Nonaka, O. Jinnnouchi, I. Kawata, M. Akutagawa and Y. Kinouchi (2016). Auditory Image Model for the Characterisation of Obstructive Sleep Apnoea. Proceedings of the American Thoracic Society, 5(2), 144-153.

[6] Flenady, T., Dwyer, T., & Applegarth, J. (2017). Accurate respiratory rates count: So should you! Australasian Emergency Nursing Journal, 20(1), 45-47.

[7] Ghaemmaghami, H., U. R. Abeyratne and C. Hukins (2009). "Normal probability testing of snore signals for diagnosis of obstructive sleep apnea." Conf Proc IEEE Eng Med Biol Soc 2009: 5551-5554.

[8] Herath, D. L., U. R. Abeyratne and C. Hukins (2015). "Hidden Markov modelling of intra-snore episode behavior of acoustic characteristics of obstructive sleep apnea patients." Physiol Meas 36(12): 2379-2404.

[9] Kushida, C. A., A. Giacomini, M. K. Lee, C. Guilleminault and W. C. Dement (2002). "Technical protocol for the use of esophageal manometry in the diagnosis of sleep-related breathing disorders." Sleep Medicine 3(2): 163-173.

[10] Markandeya, M. N., U. R. Abeyratne and C. Hukins (2018). "Characterisation of upper airway obstructions using wide-band snoring sounds." Biomedical Signal Processing and Control 46: 201-211.

[11] Markel, J. (1972). "Digital inverse filtering-a new tool for formant trajectory estimation." IEEE Transactions on Audio and Electroacoustics 20(2): 129-137.

[12] Muraja-Murro, A., A. Kulkas, M. Hiltunen, S. Kupari, T. Hukkanen, P. Tiihonen, E. Mervaala and J. Töyräs (2013). "The severity of individual obstruction events is related to increased mortality rate in severe obstructive sleep apnea." Journal of Sleep Research 22(6): 663-669.

[13] Punjabi, N. M. (2016). "COUNTERPOINT: Is the Apnea-Hypopnea Index the Best Way to Quantify the Severity of Sleep-Disordered Breathing? No." Chest 149(1): 16-19.

[14] Scholkopf, B., and A. Smola.(2002).Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond, Adaptive Computation and Machine Learning. Cambridge, MA: The MIT Press.

[15] Senaratna, C. V., J. L. Perret, C. J. Lodge, A. J. Lowe, B. E. Campbell, M. C. Matheson, G. S. Hamilton and S. C. Dharmage (2017). "Prevalence of obstructive sleep apnea in the general population: A systematic review." Sleep Med Rev 34: 70-81.

[16] Shahar, E. (2014). "Apnea-hypopnea index: time to wake up." Nature and science of sleep 6: 51-56.

[17] Simmons, J., A. Giacomini and C. Guilleminault (1993). "Routine use of a water-filled catheter for measuring respiration during NPSG studies: an overview of the procedure and clinical utility." J Sleep Res 22: 387.

[18] Vandenbussche, N. L., S. Overeem, J. P. van Dijk, P. J. Simons and D. A. Pevernagie (2015). "Assessment of respiratory effort during sleep: Esophageal pressure versus noninvasive monitoring techniques." Sleep Med Rev 24: 28-36.

[19] Zemlin, W. (1997). Speech and Hearing Science: Anatomy and Physiology. Lindon,UK, Pearson.