34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

A Comparison between Recording Sites of Snoring Sounds in Relation to Upper Airway Obstruction*

Ali Azarbarzin, *Student Member, IEEE*, and Zahra Moussavi, *Senior Member, IEEE*

Abstract— **This paper presents the results of our study on investigating the acoustical properties of snoring sounds (SS) recorded by two microphones (one over trachea and one hung in the air within 30-50 cm away from the subject) in relation to sleep apnea. Several features were extracted from SS segments of 50 snorers with different Apnea-Hypopnea Index (AHI). We used an optimal subset of the sound features to cluster the SS segments into two clusters (A and B). Then, the number of SS segments in cluster A was calculated and normalized by the total number of SS segments for each subject, resulting in 50×1 vector R. A correlation analysis was run between AHI and R. The results show a difference in acoustical properties of the tracheal and ambient snoring sounds and their ability to distinguish two types of snoring; the ambient snoring sounds are not as characteristic as tracheal snoring sounds.**

I. Introduction

Snoring is a highly prevalent disorder, which affects 20-40% of adult population and increases by age [1]. It is also a major sign of obstructive sleep apnea (OSA) syndrome [2]. However, not every snorer is apneic; those are usually referred as simple snorers. Studies have shown a positive correlation between snoring sounds (SS) intensity and Epworth sleepiness scale [3] and a significant negative correlation between Apnea-Hypopnea Index (AHI) and peak and mean frequencies of the SS power spectrum [4].

While SS analysis has often been used to diagnose OSA [5-7], it is of great interest to know how the SS characteristics change due to severity of OSA. Furthermore, it is of interest whether the snoring sounds recorded by a microphone hung in the vicinity of the subject (ambient microphone) would be as useful as those recorded by a microphone directly on the trachea (tracheal microphone).

There are many studies on SS analysis in relation to OSA. their applications are mostly limited due to small number of SS segments (<40 segments per subject) (as in [7]), inability to fully extract the SS characteristics probably due to non-Gaussian and nonlinear nature of the snoring sounds (as in [5, 6]), and also the use of only ambient microphone without paying attention to location of microphone and its effect on the recorded sound. Moreover, all the previously published studies used a supervised algorithm to investigate the properties of SS segments in relation to sleep apnea.

*Resrach supported by NSERC Canada.

Ali Azarbarzin is with the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB, Canada, R3T 5V6. (abarzin@ee.umanitoba.ca)

Zahra Moussavi is with the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB, Canada, R3T 5V6. She is also an adjunct scientist in TRLabs, and a research affiliate of Riverview Health Center, Winnipeg, MB, Canada. (moussavi@ee.umanitoba.ca).

This paper addresses the relationship of SS with OSA at different level of its severity as well as the location of the microphone by using a relatively large dataset (50 participants with an average of 408±268 SS segments per participant) and examining a wide range of acoustical markers of SS. We hypothesize that the acoustical properties of SS segments change in relation to OSA. It is known that the post-apneic SS segments show different acoustical properties than regular SS segments [8]. Therefore it is expected to see at least two different clusters of SS segments for an apneic patient (one cluster associated to regular SS segments and another associated to apneic SS segments). We also hypothesize that there is a correlation between the severity of OSA and the number of apneic SS segments. In this paper, we present the results of our study on investigating the above hypotheses as well as the effect of microphone's location on clustering results.

II. Method

A. Data Recording

Data for this study was adopted from our previous studies [9, 10]. Out of the 68 participants of that study, data of 50 individuals, who were snorers, were selected for this study. Data were recorded simultaneously with full-night Polysomnography (PSG). The respiratory sounds of the participants were collected by two miniature microphones (ECM-77B): one placed over the suprasternal notch of trachea (tracheal microphone) and the other one hung in the air 40 cm above participants' head (ambient microphone). The recorded sounds were digitized at 10240 Hz sampling rate. The participants' anthropometric information of this study is shown in Table I. The AHI value of each participant was determined by the PSG study scored by the sleep lab technicians.

The algorithm proposed in [9] was run on each individual's respiratory sounds to extract all SS segments. The algorithm resulted in start and end of each SS segment. However, to ensure 100% accuracy, all the detected SS segments were validated by visual and auditory means in the time-frequency domain and the occasional misclassified cases were removed from the database. This resulted in 20401 SS segments in total.

TABLE I. Anthropometric Information of Participating Individuals

Group	Number of subjects	Age	Body mass index	AHI
OSA	37 (6 females)	52.6± 12.1	33.1± 5.6	35.5±33.5
Non-OSA, Simple Snorers	13 (4 females)	51.3±9.6	30.1± 3.9	2.5±1.2

B. Feature Extraction

All SS segments ($s(n)$) were first band-pass filtered in the frequency range of 150-5000 Hz (to remove the effect of heart sounds and high frequency noises). The following features were calculated for each segment:

Power (P): The signal's power was computed as:

$$P = \frac{1}{N}\sum_{n=1}^{N} |s(n)|^2, \qquad (1)$$

where N is the length of each SS segment.

Duration (D)

Zero crossing rate (ZCR): It was defined as

$$ZCR = \frac{1}{2N}\sum_{n=1}^{N-1} |sgn[s(n+1)] - sgn[s(n)]|, \qquad (2)$$

where $sgn[.]$ represents the sign function.

Skewness and kurtosis: we used skewness ($\gamma_1 = \frac{c_3(0,0)}{\sigma_s^3}$) and kurtosis ($\gamma_2 = \frac{c_4(0,0,0)}{\sigma_s^4}$) as two higher order statistical (HOS) features where σ_s is the standard deviation of $s(n)$ and $c_3(0,0)$ and $c_4(0,0,0)$ are its zero-lag 3rd and 4th order cumulants respectively [11]. It should be noted that skewness is a measure of asymmetry of the probability distribution function (pdf) while kurtosis is a measure of peakedness of the pdf.

Formant Frequencies ($F_1 - F_3$): The first three formant frequencies [12] were calculated from each SS segment using linear predictive coding (LPC). To meet stationarity assumption, $s(n)$ was divided into 50 ms overlapping frames (50% overlap and Hanning window). In each frame, the autoregressive (AR) model (with order 14) of the signal was estimated and the roots of AR model were calculated. Then, the first three formant frequencies (F_1, F_2, and F_3) were estimated by taking median over all frames.

Spectral entropy (SE): SE measures the flatness of the spectrum. Larger values of SE correspond to the broader spectral contents [13]. It is computed as:

$$SE = -\sum_{f} P_n(f).\ln(P_n(f)), \qquad (3)$$

where $P_n(f)$ refers to normalized power spectrum density (PSD) at discrete frequency f. The PSD was estimated using Welch method [14] using Hanning windows of size 50 ms and 50% overlap between the successive windows.

Central Tendency Measure (CTM): CTM is a variability measure from second order difference plots [15]. CTM is calculated as below:

$$CTM = \frac{1}{N-2}\sum_{n=1}^{N-2} \delta(n), \qquad (4)$$

where $\delta(n)$ is defined as:

$$\delta(n) = \begin{cases} 1, & \sqrt{\left(s(n+2)-s(n+1)\right)^2 + \left(s(n+1)-s(n)\right)^2} < \rho \\ 0, & otherwise \end{cases}$$

Parameter ρ defines the radius of a circular region around the origin in second order difference plots. We used $\rho = 1$ for this study.

A. Blind Clustering and counting

The above mentioned features were calculated for all SS segments of all subjects. The distributions of the extracted features were different both within and between the participants. We investigated all plausible subset of features which could result in distinct clusters. Theoretically, if all the features are in the same range, we will not observe distinct clusters implying that only one type of SS exists.

While snoring sounds may indeed be of several types, as the first attempt in blind clustering, in this study we limited the number of clusters (i.e. types of snoring sounds) to 2, and used k-means clustering method [16] for all SS segments of each subject. Since this algorithm is an unsupervised clustering, the label of each SS segment was not given; the results of the clustering had to be validated because not all the feature subsets necessarily produce two distinct and compact clusters. Thus, we used Davies-Bouldin-Index (DBI [17]) to evaluate the clustering performance in terms of between-cluster separation and intra-cluster compactness. The lower values of DBI are equivalent to more separated and compact clusters.

For every plausible feature subset, the SS segments of each participant were clustered, and the corresponding DBI was calculated. The summation of DBI (DBI_S) was also calculated. After an exhaustive search over all feature subsets, the one that minimized the DBI_S was selected. The feature subset with minimum DBIs (FS_{opt}) was considered as the optimal feature subset in terms of between-cluster separation and intra-cluster compactness.

Once FS_{opt} was determined, we investigated the accumulative AHI differences between the two clusters being formed by the FS_{opt}. The following measures, calculated within the two clusters formed by FS_{opt}, give prior information on the acoustical differences between apneic snoring (*type A*) and non-apneic snoring (*type B*):

$$AHI_k^{Acc} = \sum_{j=1}^{50} R_k^{P_j} AHI^{P_j}, \qquad k = 1,2 \quad (5)$$

$$R_k^{P_j} = \frac{\sum_{i=1}^{N_j} \boldsymbol{I}(SS_i^{P_j} \in c_k)}{N_j}, \qquad k = 1,2 \quad (6)$$

where c_k is the cluster obtained from k-means algorithm (here we have two clusters c_1 and c_2), AHI_k^{Acc} represents the accumulative AHI for cluster c_k, N_j is the total number of SS segments for subject P_j, AHI^{P_j} is the AHI value for subject P_j, $R_k^{P_j}$ is the ratio of SS segment falling into cluster c_k over N_j for subject P_j, $\boldsymbol{I}(.)$ is the indicator function (the output is 1 if the condition inside argument true and 0 otherwise), and

$SS_i^{P_j}$ is the i-th SS segment of subject P_j. If $AHI_1^{Acc} > AHI_2^{Acc}$ then c_1 is denoted as cluster A and c_2 as cluster B and vice-versa otherwise. It is worth noting that for the two-cluster case: $R_2^{P_j} = 1 - R_1^{P_j}$. Therefore, we need to calculate $R_k^{P_j}$ only for one cluster (i.e. cluster A); this measure is called R^{P_j} hereafter. To investigate how the severity of OSA (measured by AHI) is correlated with the occurrence of snoring type A or B, the Kendall's Tau-b two-tailed test [18] was employed. Testing the association of AHI and the occurrence frequency of SS segment type A (or type B) is equivalent with computing the correlation between AHI and vector $R = \{R^{P_j} | j = 1, \ldots, 50\}$.

III. RESULTS

The optimal feature subset was found to be a 3-D feature vector including power (P), central tendency measure (CTM), and skewness (γ_1). This feature subset resulted in a minimum value of DBI equaled to 13.6 for tracheal microphone and a minimum value of 32.5 for ambient microphone. As shown in Table II, the difference between the accumulative AHI ($AHI_1^{Acc} - AHI_2^{Acc}$) of two clusters of tracheal recordings is larger than that of ambient recording. The characteristics of snoring type A and B for both ambient and tracheal data set are shown in Table II.

Furthermore, the minimum DBI value for clustering validation is lower for tracheal microphone than ambient microphone. This indicates that tracheal recordings resulted in more separated and compact clusters than ambient recordings. This is also shown in Fig. 1, in which the clustering results were shown for a typical snorer using tracheal and ambient microphone. For tracheal recordings, snoring type A was characterized with lower values of CTM (0.78±0.13), lower values of F_1 (261±38 Hz), and lower values of γ_1 (-0.65±0.38). Snoring type B, on the other hand, was characterized by higher values of CTM (0.95±0.05), higher values of F_1 (336±32 Hz), and higher values of γ_1 (-0.1±0.35). In addition, the value of vector R was positively correlated with AHI (p<0.01) for tracheal microphone while the correlation between vector R and AHI was not significant (p>0.05).

TABLE II. THE ACOUSTICAL CHARACTERISTICS OF SNORING SOUND SEGMENTS, ACCUMULATIVE AHI FOR TWO CLUSTERS, DBI, AND THEIR DIFFERENCES BETWEEN AMBIENT AND TRACHEAL MICROPHONES.

Feature	Snoring type	Tracheal Microphone	Ambient Microphone
$F_1(Hz)$	A	261±38	272±82
	B	336±32	307±104
CTM	A	0.78±0.13	0.98±0.02
	B	0.95±0.05	0.99±0.006
γ_1	A	-0.65±0.38	-0.02±0.2
	B	-0.1±0.35	-0.01±0.25
AHI_1^{Acc}		634	555
AHI_2^{Acc}		256	335
minimum DBI		13.6	32.5

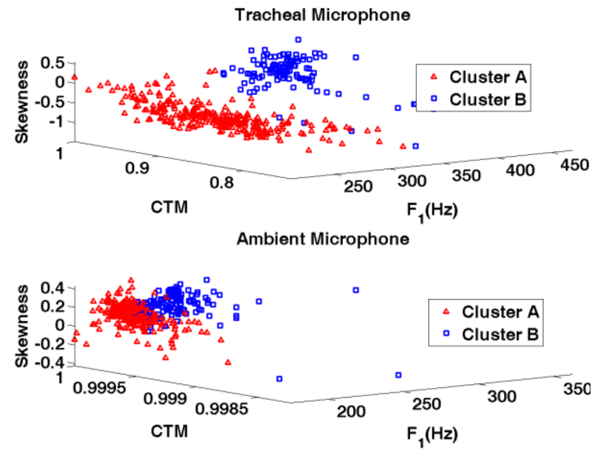

Figure 1. K-means clustering on a 3-D subset of features of an individual (AHI=23.3) with 396 SS segments in total.

IV. DISCUSSION

In this study, we characterized the snoring sounds collected by two tracheal and ambient microphones from both apneic and non-apneic individuals, and investigated whether they form different distinct clusters. Since people snore due to different reasons, we expected that the source of snore generation would impact on the acoustical properties of the snoring sounds. Therefore, assuming that the cause of OSA also changes the acoustical characteristics of snoring, we hypothesized that there exist at least two distinct clusters of snoring sounds. Furthermore, we hypothesized that the frequency of the occurrence of each type of snoring is correlated with AHI. As a pilot study, we limited the number of clusters to 2, and investigated whether the type of snoring can give any indication of the apnea's severity.

We found that the SS collected by the ambient microphone was unable to distinguish between snoring sounds of two different sources; this could be due to the path that sound wave travels from its source to the recording location. The signal to noise ratio and power of signal also plays an important role in the separation ability of SS signals. This is very important as several studies (e.g. [5, 6]) used ambient recordings for acoustical analysis of SS. Therefore, we recommend recording the SS over trachea.

Our observations for tracheal recordings showed that participants with higher AHI snored more frequently of SS segments characterized by the lower central tendency measure, lower skewness, and lower first formant frequency. Therefore, one may conclude that OSA and snoring sounds are related to each other in terms of the above-mentioned effects. This result is supported partially by the results reported in [8] on a small subset of data. However these results were not seen in ambient microphone as snoring type A and B fully overlapped for these features implying that only one type of snoring exist in ambient microphone and therefore some crucial information was lost by changing the recording location. Therefore, this is the reason that no correlation was observed between occurrence of snoring type A and severity of sleep apnea.

ACKNOWLEDGMENT

This study was supported in part by *Natural Sciences and Engineering Research Council* of Canada (NSERC), and Telecommunication Research Lab (TRLabs) Manitoba.

REFERENCES

[1] V. Hoffstein, "Apnea and snoring: state of the art and future directions," *Acta Otorhinolaryngol Belg,* vol. 56, pp. 205-36, 2002.

[2] D. L. Bliwise, J. C. Nekich and W. C. Dement, "Relative validity of self-reported snoring as a symptom of sleep apnea in a sleep clinic population," *Chest,* vol. 99, pp. 600-608, 03/01, 1991.

[3] H. Nakano, T. Furukawa and S. Nishima, "Relationship between snoring sound intensity and sleepiness in patients with obstructive sleep apnea." *J Clin Sleep Med,* vol. 4, pp. 551-556, 2008.

[4] J. Fiz, J. Abad, R. Jane, M. Riera, M. Mananas, P. Caminal, D. Rodenstein and J. Morera, "Acoustic analysis of snoring in patients with simple snoring and obstructive sleep apnea," *Eur. Respir. J.,* vol. 9, pp. 2365-70, 1996.

[5] U. Abeyratne, A. Wakwella and C. Hukins, "Pitch jump probability measures for the analysis of snoring sounds in apnea," *Physiol. Meas.,* vol. 26, pp. 779-98, 2005.

[6] J. Sola-Soler, R. Jane, J. Fiz and J. Morera, "Pitch analysis in snoring signals from simple snorers and patients with obstructive sleep apnea," in *EMBS /BMES,* 2002, .

[7] A. Ng, T. Koh, U. Abeyratne and K. Puvanendran, "Investigation of Obstructive Sleep Apnea Using Nonlinear Mode Interactions in Nonstationary Snore Signals," *Ann. Biomed. Eng.,* vol. 37, pp. 1796-1806, 2009.

[8] J. Perez-Padilla, E. Slawinski, L. Difrancesco, R. Feige, J. Remmers and W. Whitelaw, "Characteristics of the snoring noise in patients with and without occlusive sleep apnea," *Am. Rev. Respir. Dis.,* vol. 147(3), pp. 635-644, 1993.

[9] A. Azarbarzin and Z. Moussavi, "Automatic and Unsupervised Snore Sound Extraction from Respiratory Sound Signals," *IEEE Trans Biomed Eng.,* vol. 58, pp. 1156 - 1162, 2011.

[10] A. Yadollahi, E. Giannouli and Z. Moussavi, "Sleep apnea monitoring and diagnosis based on pulse oximetery and tracheal sound signals," *Medical and Biological Engineering and Computing,* vol. 48, pp. 1087-1097, 2010.

[11] A. Swami, J. M. Mendel and C. L. Nikias, "Higher-order spectral analysis toolbox User's guide," The MathWorks, Tech. Rep. version2, 1998.

[12] X. Huang, A. Acero and H. Hon, *Spoken Language Processing: A Guide to Theory, Algorithm and System Development.* Prentice Hall PTR, 2001.

[13] T. Inouye, K. Shinosaki, H. Sakamoto, S. Toi, S. Ukai, A. Iyama, Y. Katsuda and M. Hirano, "Quantification of EEG irregularity by use of the entropy of the power spectrum," *Electroencephalogr. Clin. Neurophysiol.,* vol. 79, pp. 204-210, /9, 1991.

[14] J. G. Proakis and D. K. Manolakis, *Digital Signal Processing.* Prentice Hall, 2006.

[15] M. E. Cohen, D. L. Hudson and P. C. Deedwania, "Applying continuous chaotic modeling to cardiac signal analysis," *Engineering in Medicine and Biology Magazine, IEEE,* vol. 15, pp. 97-102, 1996.

[16] R. Duda, P. Hart and D. Stork, *Pattern Classification.* John Wiley and Sons, Inc, 2001.

[17] D. Davies and D. Bouldin, "A Cluster Separation Measure," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. PAMI-1, pp. 224-227, 04/27, 1979.

[18] M. G. Kendall, "A New Measure of Rank Correlation," *Biometrika,* vol. 30, pp. pp. 81-93, Jun., 1938.