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

Statistical Analysis of Tracheal Breath Sounds during Wakefulness for Screening Obstructive Sleep Apnea

Cameron A. MacGregor, *Student Member, IEEE*, Davood Karimi, Ali Azarbarzin, and Zahra Moussavi, *Senior Member, IEEE*

Abstract— **Obstructive sleep apnea (OSA) is a prevalent disorder. The accepted method of diagnosis in widespread clinical practice, polysomnography (PSG), is costly and very time consuming; therefore, quick screening methods, especially when there is a need for quick diagnosis, is of great interest. Diagnostic methods which exploit subtle differences in breath sounds recorded during wakefulness, such as our group's Awake-OSA technology, have shown their capability to diagnose OSA at the research stage. Simplifying the breath sound recording procedure employed in the Awake-OSA diagnostic method would increase its efficiency when used in a clinical setting. In this study, we adopted breath sound data collected during wakefulness in two positions (sitting upright and supine) and two breathing maneuvers (nose and mouth breathing) from our previous study [10], and ran hypothesis tests on a wide variety of sound features to select the most significant features correlated with OSA. The goal was to investigate which combinations of patient position and breathing maneuver contribute the least to the significant features amongst groups of people with differing OSA severity, thus permitting simplification of the recording protocol. The results show that all signals recorded by a combination of the two breathing maneuvers and two positions result in features significantly correlated with OSA severity; this makes it impossible to confidently recommend that a combination be omitted from the recording protocol. Nevertheless, the results show that the majority of significant features originated from recordings made in the supine position. Therefore, as a step toward simplification of the Awake-OSA diagnostic algorithm, we may use breath sound signals recorded only in the supine position and further investigate the accuracy of the algorithm in distinguishing amongst groups with differing OSA severity.**

I. Introduction

Repeated arousal from sleep can negatively affect health and wellbeing. Obstructive sleep apnea (OSA) is a prevalent disease [1] which significantly lowers sleep quality, thereby impacting the health of sufferers [2–3]. OSA is distinguished by recurring upper airway collapse during sleep due to the pharynx having greater inherent collapsibility [4]. Airway collapsibility depends on anatomical features which include the amount of fat present in the pharynx and the size of the mandible [5 cited in 4]. Because the pharyngeal dilator muscle is more active during wakefulness, OSA sufferers do not experience any upper airway obstruction during wakefulness [6]. In many patients, an episode of sleep apnea may be ended by an arousal, which increases dilator muscles' activity during sleep [6]. These arousals can occur without fully awakening the individual, leaving them with the effects of sleep deprivation but without knowledge of its cause. The accepted method of OSA diagnosis in widespread clinical practice is the overnight polysomnogram (PSG). The overnight PSG involves continuous recording of multiple physiological signals from the patient, which are then scored (interpreted) by a trained sleep lab technologist. Due to the high cost of overnight PSG on the health care system, its long waiting list, its inconvenience for patients, and also the fact that the majority of patients referred to overnight PSG may not have OSA in need of treatment, alternative quick screening methods have been of great interest [9].

Studies have shown that OSA can be diagnosed by analysis of breath sounds recorded during sleep [7]. Our group has been investigating the use of breath sound signals recorded during wakefulness as a potential diagnostic tool for screening OSA and its severity [8–10]; the resultant technology is called Awake-OSA, in which tracheal breath sounds are recorded in two positions, upright and supine, and two breathing maneuvers, nose and mouth, while the patient is awake.

The aim of this study has been to investigate whether we can shorten and simplify the recording protocol of Awake-OSA technology while retaining the representative sound features. Simplifying the breath sound recording procedure would improve the focus of research in this area, and increase the efficiency of this screening technique in a clinical setting. At the research stage, examining a subset of signals would result in fewer sound features per patient, which would in turn improve the ability of both the exhaustive and floating feature search algorithms used in [10] to find an optimum set of features for OSA diagnosis and severity estimation. For example, starting with a smaller set of features would allow the exhaustive search algorithm to find larger feature subsets within a reasonable amount of computational time, and would increase the likelihood that the floating search algorithm will encounter a combination of features that can classify (diagnose) patients more accurately. Furthermore, in a clinical setting, we envision that a more efficient version of our Awake-OSA diagnostic technology could be employed as a pre-screening tool either before surgery, or before referral to overnight PSG, thereby permitting more efficient use of overnight PSG resources. For example, even though the current recording procedure takes significantly less time (approximately 20 minutes) compared to overnight PSG, a shorter (e.g. 1-5 minutes) protocol may permit the use of this technology for rapid pre-screening of OSA in a doctor's office to decide whether to refer the patient to an overnight PSG for a detailed sleep assessment [10].

Cameron MacGregor (ummacgrc@cc.umanitoba.ca), Davood Karimi (d_karimi@umanitoba.ca), Ali Azarbarzin (umazarba@cc.umanitoba.ca), and Zahra Moussavi (mousavi@cc.umanitoba.ca) are with the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, Manitoba, Canada.

TABLE I. FEATURES DERIVED FROM THE ESTIMATED POWER SPECTRUM OF THE BREATH SOUND SIGNAL [10]

Feature name	Description	Equation(s)*	
Signal power[a]	Power contained within the specified frequency band	$P_s\|_{f_l \le f \le f_u} \triangleq \sum_{f_l \le f \le f_u} P(f)\Delta f$	(1)
Relative signal power[a]	Power contained within the specified frequency band relative to the power contained over the entire estimated power spectrum	$\frac{P_s\|_{f_l \le f \le f_u}}{P_s\|_{100 \le f \le 2500}}$	(2)
Spectral centroid[a,b]	Frequency around which most of the power is centered	$SC \triangleq \frac{\sum_{f_l \le f \le f_u} f P(f)\Delta f}{P_s\|_{f_l \le f \le f_u}}$	(3)
Spectral bandwidth[a,b]	Power spectral density-weighted average of the squared distance between the spectral centroid and the different frequency components	$\frac{\sum_{f_l \le f \le f_u} (f-SC)^2 P(f)\Delta f}{P_s\|_{f_l \le f \le f_u}}$	(4)
Spectral flatness[b]	A measure of how similar the breath sound is to a pure tone (i.e. its tonality)	$\frac{\left(\Pi_{f_l \le f \le f_u} P(f)\right)^{\frac{1}{f_u - f_l}}}{P_a}$, where $P_a \triangleq \frac{P_s\|_{f_l \le f \le f_u}}{(f_u - f_l)}$	(5) (6)
Crest factor[b]	A different measure of breath sound tonality	$\max_{f_l \le f \le f_u}\left(\frac{P(f)}{P_a}\right)$	(7)

* $P(f) \triangleq$ estimated power spectrum; $f_l \triangleq$ lower frequency; $f_u \triangleq$ upper frequency.

a. Feature was computed over six sub-bands of the power spectrum: $f_l = 100$, $f_u = 150$; 150, 450; 450, 600; 600, 1200; 1200, 1800; and 1800, 2500 Hz.

b. Feature was computed over the entire power spectrum $f_l = 100$, $f_u = 2500$ Hz.

II. METHOD

A. Data Collection

The data used in this study were adopted from our previous studies [8, 10], which were collected at Misericordia Hospital, Winnipeg, from awake patients before they proceeded to an overnight PSG assessment. The tracheal breath sounds were collected by a miniature microphone (Sony ECM-77B) over the suprasternal notch, and recorded by an amplifier which band-pass filtered the signal between 0.05 Hz and 5 kHz (Biopac DA100C) at a sampling rate of 10240 Hz. The apnea-hypopnea index (AHI) of each patient was collected from the PSG assessment by the sleep lab technologist, prospectively. Patients were instructed to breathe deeply with a constant flow rate following the hand movements of the recorder (to assist in keeping the same flow rate in breathing) in the upright and again in the supine position, while breathing through their nose and then their mouth. Thus, from every patient we recorded 4 signals: (I) nose breathing while upright, (II) mouth breathing while upright, (III) nose breathing while supine, and (IV) mouth breathing while supine. To ensure breath phase accuracy in all our recordings, the experimenter gave a vocal cue (audible on the recording) at the start of the patient's inspiratory phase.

B. Data Analysis

Inspiratory and expiratory phases were separated using the technique described in [10], which resulted in a series of individual breath sound signals. Each breath sound signal, consisting of discrete audio samples, was split into 50 ms (512 sample) contiguous segments, and the statistical variance of each segment computed. The stationary portion of each breath sound signal was automatically chosen by discarding all audio samples located outside of a 300 ms window centered around the 50 ms segment with the highest variance. These 300 ms windows were individually standardized to have a zero mean and standard deviation of 1 in order to minimize the influence of plausible differences in respiratory flow on the sound features. Next, the power spectrum over 100 – 2500 Hz and the bispectrum over 100 – 2600 Hz were estimated for the sounds in each of these 300 ms segments for every corresponding breath sound, as described in detail in [10].

Using the power spectrum and bispectrum estimated for each breath, all 28 features listed in Table 1 and all 112 features listed in Table 2 were computed separately for each breath sound. Given that there are 4 signals recorded per patient and that the inspiratory and expiratory phases are analyzed separately, there are 8 signals per patient. Each signal consists of a series of discrete breaths, with 140 (112 + 28) feature values associated with each breath. Mean feature values are computed for all 140 features over all breaths originating from each signal. Thus, there are 1120 feature values available per patient (140 features × 8 signals). We also looked at the differences of the mean feature values derived from nose versus mouth breathing, and upright versus supine positions. Therefore, in total, we have 2240[1] feature values per patient.

Two statistical t-tests and a one-way unbalanced Analysis of Variance (ANOVA) test were used to determine which features had a significant ($P \le 0.01$) correlation with AHI, and thus have the potential to be used for classification of patients with OSA. The first t-test was used to find the statistically significant features between the patients with AHI ≤ 5 and those with AHI ≥ 30. The second t-test was the same as the first, except that the patients were grouped according to AHI ≤ 10 and AHI ≥ 20. The one-way unbalanced ANOVA was used to find the statistically significant features amongst three groups of patients: those with AHI ≤ 5, 10 ≤ AHI ≤ 25, and AHI ≥ 30. The sound features that passed all three tests were then grouped according to the breathing maneuver(s), position(s), and breath phase they originated from. These AHI intervals were chosen in order to be consistent with our previous studies that were based on the clinical importance of these AHI intervals [7–10].

III. RESULTS

The results of all three tests are summarized in Table 3. Of all the 2240 available features, a total of 41 features were commonly found to be significantly ($P \le 0.01$) different

[1] (4 power spectrum (PS) features × 6 sub-bands + 4 PS features × 1 band + 46 bispectrum (BSP) invariant features + 55 BSP average magnitude features + 9 other BSP features + 2 BSP median bifrequency features) × (4 position and maneuver combinations + 2 nose minus mouth differences + 2 supine minus upright differences) × (2 phases)

TABLE II. FEATURES DERIVED FROM THE ESTIMATED BISPECTRUM OF THE BREATH SOUND SIGNAL BETWEEN 100 AND 2600 HZ [10]

Feature name	Portion of bispectrum	Equation(s)†	
Bispectral invariant parameter $P(a)$	Estimated over radial lines with slopes starting at 1º and ending at 45 º, in 1º increments	$P(a) = \text{atan}\left(\frac{\Im\mathfrak{m}\{I(a)\}}{\Re\mathfrak{e}\{I(a)\}}\right)$, where $I(a) \cong \sum_{k=1}^{\left\lfloor\left(\frac{N}{2}-1\right)/(1+a)\right\rfloor}\left(p\, C_3(k,\lfloor ak\rfloor) + (1-p)C_3(k,\lceil ak\rceil)\right)$	(8) (9)
Average magnitude	Ω = each square portion of the non-redundant region of the bispectrum formed when each of the f_1 and f_2 axes are divided into 10 equally sized sub-bands between 100 and 2600 Hz (55 subregions in total)	$\frac{1}{N}\sum_{\Omega}\lvert C_3(f_1,f_2)\rvert$	(10)
Average magnitude	Ω = the non-redundant region of the bispectrum	$\frac{1}{N}\sum_{\Omega}\lvert C_3(f_1,f_2)\rvert$	(11)
Average power		$\frac{1}{N}\sum_{\Omega}\lvert C_3(f_1,f_2)\rvert^2$	(12)
Normalized entropy		$-\sum_n p_n \log\frac{\lvert C_3(f_1,f_2)\rvert}{\sum_{\Omega}\lvert C_3(f_1,f_2)\rvert}$	(13)
Normalized squared entropy		$-\sum_n q_n \log\frac{\lvert C_3(f_1,f_2)\rvert^2}{\sum_{\Omega}\lvert C_3(f_1,f_2)\rvert^2}$	(14)
Sum of logarithmic amplitudes		$\sum_{\Omega}\log\lvert C_3(f_1,f_2)\rvert$	(15)
Sum of logarithmic amplitudes	Diagonal elements of the bispectrum	$\sum\log\lvert C_3(f,f)\rvert$	(16)
First-order moment of the logarithmic amplitudes		$\sum f.\log\lvert C_3(f,f)\rvert$	(17)
Second-order moment of the logarithmic amplitudes		$\sum(f-H_3)^2.\log\lvert C_3(f,f)\rvert$	(18)
Phase entropy	Ω = the non-redundant region of the bispectrum	$\sum_n p(\psi_n)\log\left(p(\psi_n)\right)$, where $p(\psi_n) = \frac{1}{N}\sum_{\Omega} 1\left(\phi(C_3(f_1,f_2) \in \psi_n)\right)$, and $\psi_n = \left\{\phi \mid -\pi + \frac{2\pi n}{M} \le \phi \le -\pi + \frac{2\pi(n+1)}{M}\right\}$	(19) (20) (21)
Median bifrequency	The frequency at which the area under the bispectrum is equal on both sides; estimated for both bispectral frequencies according to the algorithm in [10].		

† $C_3(f_1,f_2) \triangleq$ estimated bispectrum at frequency f_1 and f_2; $\Omega \triangleq$ an arbitrary two-dimensional region of the bispectrum over the f_1 and f_2 axes; $N \triangleq$ number of discrete points in region Ω.

among all the sets of AHI groups tested. Of these 41 features, some are derived from one of the four signals recorded for each patient, and some are difference features, i.e. nose versus mouth, upright versus supine, which are calculated by simple subtraction of the above features. Thus, difference features are derived from two of the four signals recorded for each patient. Table 4 lists the number and proportion of these 41 features when grouped according to the breath sound signal(s) and phase from which they were derived.

Overall, the greatest proportion of features was derived from mouth breathing recorded in the supine position (39%). When the features are categorized according to which position (supine, upright) or position combination (supine minus upright) they originate from, the majority of features (76%) were derived exclusively from recordings of the patient while supine. When the features are categorized according to which breathing maneuver (nose, mouth) or maneuver combination (nose minus mouth) they originate from, the majority of features (54%) were derived exclusively from recordings of mouth breathing. (These percentages do not add up to 100% because each refers to a different grouping of the 41 features, as shown in Table 4.)

TABLE III. THE NUMBER OF FEATURES TO PASS EACH HYPOTHESIS TEST AND NUMBER OF FEATURES COMMONLY PASSED BY THE TESTS

First t-test between AHI ≤ 5, AHI ≥ 30	64	53 features pass both t-tests	41 features pass all three tests
Second t-test between AHI ≤ 10, AHI ≥ 20	94		
ANOVA between AHI ≤ 5, 10 ≤ AHI ≤ 25, AHI ≥ 30	98		

Overall, the number of features derived from the inspiratory phase (20) and the number of features derived from the expiratory phase (21) are approximately equal. When grouped according position and maneuver combination or difference thereof (each thick-bordered cell in Table 4), the distribution of features amongst breath phase is skewed towards inspiration almost as much as it is skewed towards expiration (with the exception of features derived from the combination of nose minus mouth breathing in the supine position, which are equal). Some features derived from specific breath phases, and one difference feature, did not pass all three hypothesis tests.

IV. DISCUSSION

The number of statistically significant features derived from breathing sounds recorded in the upright position, both exclusively and in combination with the supine position (8 and 2, respectively) were not comparable with the number of features derived exclusively from breathing sounds recorded in the supine position (31). Since approximately three-quarters (76%) of the features were derived exclusively from breath sound recordings of the patient in the supine position, the recording procedure could probably be shortened by recording only nose and mouth breathing in the supine position.

In the supine position, the number of features derived exclusively from mouth breathing (16) versus the number of

TABLE IV. THE NUMBER OF FEATURES TO PASS ALL THREE HYPOTHESIS TESTS GROUPED ACCORDING TO THE COMBINATION OF PATIENT BREATHING MANEUVER(S), PATIENT POSITION(S), AND BREATH PHASE THEY ORIGINATE FROM

Feature Sources			Patient breathing maneuver				*Nose minus mouth (both maneuvers)*		*Totals for each position*	
			Nose		*Mouth*					
			Inspiration	*Expiration*	*Inspiration*	*Expiration*	*Inspiration*	*Expiration*	*Inspiration*	*Expiration*
Patient position	*Supine*		5 (13%)		16 (39%)		10 (24%)		31 (76%)	
	Inspiration	*Expiration*	2	3	11	5	5	5	18	13
	Upright		1 (2%)		6 (15%)		1 (2%)		8 (19%)	
	Inspiration	*Expiration*	1	0	1	5	0	1	2	6
Supine minus upright (both positions)			2 (5%)		0 (0%)		–		2 (5%)	
Inspiration	*Expiration*		0	2	0	0			0	2
Totals for each maneuver			8 (20%)		22 (54%)		11 (26%)		41 (100%)	
Inspiration	*Expiration*		3	5	12	10	5	6		

features derived from nose breathing, including its difference with mouth breathing (5 plus 10, respectively) is comparable; this makes it hard to reject recordings of nose breathing in favor of mouth breathing in the supine position. Overall, even though the greatest proportion of selected features originated from mouth breathing in the supine position (39%), the number of features derived from nose breathing in both patient positions, including its difference with mouth breathing (8 plus 11, respectively) is likewise comparable; this suggests that the diagnostic information provided by nose breathing, both on its own and when in tandem with mouth breathing, is valuable enough for it to remain a part of the recording procedure.

None of the four combinations of recording maneuvers and patient positions (upright nose, upright mouth, supine nose, supine mouth) were without features that passed all three hypothesis tests, indicating that none of these recording combinations can be outright rejected. On the other hand, all difference features derived from certain breath phases of certain recording combinations, and all features derived from the expiration phase of nose breathing in the upright position did not pass all three hypothesis tests, indicating that these phase and recording combinations can be rejected; these cases are shaded in Table 4. In the former case, for example, no difference feature between supine and upright positions during mouth breathing passed all three hypothesis tests. Thus, the sound features derived from these specific combinations of respiratory phase, position, and breathing maneuver may be ignored in the Awake-OSA diagnostic technique [8–9].

Omitting the features derived from the above sources would enable the feature search algorithms employed in [10] to exhaustively search for larger feature subsets and would increase the likelihood that the floating search would encounter a combination of features with better overall classification accuracy. Furthermore, the fact that the majority of features originate from supine recordings suggests that breath sounds recorded while the patient is supine may contain the most relevant information for diagnosing OSA; a future study may perform a feature search and classification using breaths recorded exclusively in the supine position. In order to have a final conclusion on optimizing the recording protocol, however, one has to run classification on the subsets of the commonly selected features in this study and verify the above results by the classification accuracies for the desired groups of AHI.

ACKNOWLEDGMENT

The authors would like to thank Mr. Ehsan Shams for assisting with data collection. The assistance of the staff of the Sleep Disorders Clinic, Misericordia Hospital, Winnipeg, MB, Canada in helping to facilitate the collection of data used in this study is greatly appreciated.

REFERENCES

[1] 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 J. of Medicine*, vol. 328, no. 17, pp. 1230-1235, April 1993.

[2] P. E. Peppard, T. Young, M. Palta, and J. Skatrud, "Prospective Study of the Association between Sleep-Disordered Breathing and Hypertension," *New England J. of Medicine*, vol. 342, no. 19, pp. 1378–1384, May 2000.

[3] E. Shahar, C. W. Whitney, S. Redline, E. T. Lee, A. B. Newman, F. Javier Nieto, G. T. O'Connor, L. L. Boland, J. E. Schwartz, and J. M. Samet, "Sleep-disordered breathing and cardiovascular disease: cross-sectional results of the Sleep Heart Health Study," *Amer. J. of Respiratory and Critical Care Medicine*, vol. 163, no. 1, pp. 19-25, Jan. 2001.

[4] A. Malhotra and D. P. White, "Obstructive sleep apnea," *The Lancet*, vol. 360, no. 9328, pp. 237-245, Jul. 20, 2002.

[5] J. E. Remmers, W. J. deGroot, E. K. Sauerland, and A. M. Anch, "Pathogenesis of upper airway occlusion during sleep," *J. of Appl. Physiology*, vol. 44, no. 6, pp. 931-938, Jun. 1, 1978.

[6] M. Younes, "Role of respiratory control mechanisms in the pathogenesis of obstructive sleep disorders," *J. of Appl. Physiology*, vol. 105, no. 5, pp. 1389-1405, Nov. 2008.

[7] A. Yadollahi, E. Giannouli, and Z. Moussavi, "Sleep apnea monitoring and diagnosis based on pulse oximetery and tracheal sound signals," *Medical & Biological Eng. & Computing*, vol. 48, no. 11, pp. 1087-1097, Nov. 2010.

[8] A. Montazeri, E. Giannouli, and Z. Moussavi, "Assessment of Obstructive Sleep Apnea by Respiratory Sound Analysis during Wakefulness," *J. Annals on Biomed. Eng.,* vol. 4, no. 4, pp. 916-924, 2012.

[9] D. Karimi and Z. Moussavi, "Diagnosis of obstructive sleep apnea based on spectral features of tracheal breath sounds during wakefulness" in *26th Ann. Meeting of the Associated Professional Sleep Societies, LLC*, Boston, MA, USA, 2012, pp. A136-A137.

[10] D. Karimi, "Spectral and Bispectral Analysis of Awake Breathing Sounds for Obstructive Sleep Apnea Diagnosis," M.Sc. thesis, Dept. of Elec. and Comp. Eng., Univ. of Manitoba, Winnipeg, MB, 2012, URI: http://hdl.handle.net/1993/18327 [Accessed Apr. 9, 2013].