Breathing Rate Estimation during Sleep Using Audio Signal Analysis

E. Dafna, T. Rosenwein, A. Tarasiuk, Y. Zigel

***Abstract*—Sleep is associated with important changes in respiratory rate and ventilation. Currently, breathing rate (BR) is measured during sleep using an array of contact and wearable sensors, including airflow sensors and respiratory belts; there is need for a simplified and more comfortable approach to monitor respiration. Here, we present a new method for BR evaluation during sleep using a non-contact microphone. The basic idea behind this approach is that during sleep the upper airway becomes narrower due to muscle relaxation, which leads to louder breathing sounds that can be captured via ambient microphone. In this study we developed a signal processing algorithm that emphasizes breathing sounds, extracts breathing-related features, and estimates BR during sleep.**

A comparison between audio-based BR estimation and BR calculated using the traditional (gold-standard) respiratory belts during in-laboratory polysomnography (PSG) study was performed on 204 subjects. Pearson's correlation between subjects' averaged BR of the two approaches was *R*=0.97. Epoch-by-epoch (30 s) BR comparison revealed a mean relative error of 2.44% and Pearson's correlation of 0.68. This study shows reliable and promising results for non-contact BR estimation.

***Keywords*: Breathing sounds, respiratory analysis, breathing rate estimation, audio signal processing.**

I. Introduction

In recent years, several studies have explored new approaches to estimate breathing rate (BR) during sleep using minimal contact (non-disturbing) sensors. Bates et al. [1] explored the ability of a tri-axial accelerometer worn on the torso to record breathing activity and to calculate BR (wireless). They compared their BR estimation with nasal cannula sensor and showed a correlation of 0.870-0.928. Johnston and Mendelson [2] used a photo-plethysmographic signal that was recorded by a reflectance pulse oximeter sensor mounted on the subject's forehead. Their BR analysis was subsequently processed by a time domain filtering and frequency analysis; they showed, qualitatively, a good agreement between their approach and respiratory belts. Murthy et al. [3] estimated respiratory rate using a sensitive thermal infra-red (IR) camera (thermal sensitivity of 0.025°C); They measured airflow due to temperature difference resulting from respiratory activity. Accuracy of BR estimation between 83% and 98% depending on the tested subject (10 participants) was reported.

Breathing activity involves pressure waves that generate breathing sounds due to the flow of air in high resistance airways. As soon as a person falls asleep, there is decreased activity of the pharyngeal dilator muscles, resulting in air turbulence and more audible breathing sounds [4, 5]. Breathing sound during sleep may exhibit a wide range of sound intensity, ranging from very quiet to very loud (typically observed as snore) [6]. Recently, we have shown that it is possible to detect even quiet breathing sounds with a non-contact microphone, characterized by sound intensity of about 20 dB (lower than a quiet bedroom's noise) [7-9]. Using noise reduction and breathing emphasis algorithms [7-9], we found that most of breathing cycles are captured, making BR estimation from breathing sound feasible.

In the current study, we developed a non-contact audio-based BR estimation (during sleep) algorithm. By taking advantage of the nature of breathing activity, i.e., breathing is repetitive and statistically periodic, BR estimation can be reliable. To our knowledge, this is the first reported attempt to determine BR from breathing sounds using non-contact audio signals during sleep. This proposed algorithm produces reliable BR estimations regardless of gender, age, body-mass-index (BMI), apnea-hypopnea-index (AHI), and snoring intensity, and is suitable for other related audio-based (sleep) diagnoses.

II. Methods

A total of 204 patients were recorded at the Sleep-Wake Disorder Unit (Soroka University Medical Center) during routine PSG study. The Institutional Review Committee of Soroka University Medical Center approved this study protocol (protocol number 10141). All patients were recorded using a digital audio recording device (Edirol *R-4 Pro*) connected to a non-contact microphone (RØDE NTG-1) that was placed 1.0 m above the patient's head. Sixty of the patients were also recorded using a handy audio recorder (Olympus LS-5) that was positioned on the dresser behind the patient's head. BRs, estimated using the audio recordings, were compared to the gold standard respiratory activity monitoring, achieved using abdomen and chest effort belts, which are based on respiratory inductance plethysmography (RIP) (SomniPro 19 PSG, Deymed Diagnostic, Hronov, Czech Republic). The acquired audio signals were digitized and stored at 16 kHz, 16 bits. RIP

*This work was supported in part by the Israel Ministry of Economics, The Kamin Program, award no. 46168.

E. Dafna, T. Rosenwein are with the Department of Biomedical Engineering, Ben-Gurion University of the Negev, Beer–Sheva, Israel (elirandafna@post.bgu.ac.il).

A. Tarasiuk is with the Sleep-Wake Disorders Unit, Soroka University Medical Center, and Department of Physiology, Faculty of Health Sciences, Ben-Gurion University of the Negev, Israel (tarasiuk@bgu.ac.il).

Y. Zigel is with the Department of Biomedical Engineering, Ben-Gurion University of the Negev, Beer–Sheva, Israel (corresponding author, phone: +972-8-642-8372; fax: +972-8-642-8371; e-mail: yaniv@bgu.ac.il).

channels were stored at 40 Hz, 16 bits. Each subject's BR was estimated continuously throughout the night in two different ways: 1) using analysis of RIP signals from the PSG, and 2) using analysis of the audio signal that was recorded using the digital audio recorder. Both estimations of BR were compared in order to evaluate the similarities between the approaches. The comparison includes a variety of statistical analysis tools (see section D).

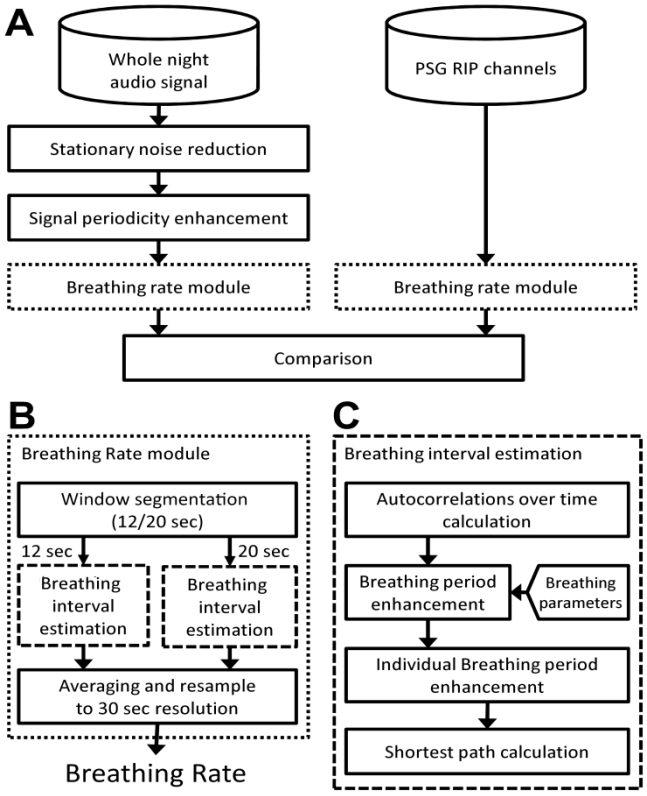

Figure 1. Block-diagram of the BR estimation algorithm.

Figure 1 presents the block diagram of the BR estimation algorithm using both of the approaches (RIP and audio signals). The RIP approach includes two channels that measure the abdomen and thorax activities. The RIP channels were processed in the BR module (see BR module section) to estimate the BR.

In the Audio signal approach, the audio signal undergoes an adaptive noise reduction technique to suppress any stationary background noise that disrupts and covers breathing events. Next, the enhanced signal follows a periodicity enhancement module that aims to suppress spectral components that did not follow any periodic pattern. The final stage in this process of BR estimation is the BR module (same as in the RIP approach but for different signals).

A. Audio and PSG-RIP signals

Our proposed algorithm of BR estimation can be applied on either audio signals or on PSG RIP signals with the appropriate preprocessing modules.

Audio signal – In order to reliably capture breathing cycles, a breathing signal enhancement procedure is required. The first step is to remove adaptively any stationary background noise such as air-conditioner or fan noises. We used our previously designed noise reduction algorithm [7], which results in an average of +6 dB signal-to-noise-ratio enhancement. The next step is to remove short transient noises (<200ms [7]) that are unlikely to be related to breathing activity. In order to do so, we remove these noises using 400 ms median filter (temporal filtering) on the signal's spectrum components.

Luckily, breathing activity is a repetitive process that usually involves periodic intervals. We took advantage of this phenomenon to further enhance the periodicity of the audio signal in a segmented window (see section B) by eliminating 75% of the least periodic spectral components of the audio signal (keeping 64 out of 256 spectral coefficients of STFT). For more information see a detailed explanation of this process in [8]. The outcome of the two steps listed above is a 2D image (matrix) with 64 spectral components (64 rows in the matrix) calculated from 50ms frames at 40 Hz frame-rate (each column represents a time frame).

PSG RIP signals – PSG RIP signals are designed to capture any changes in abdomen and thorax perimeters as the subject dynamically changes his lung volume in order to breathe. These changes are periodic and follow the subject's breathing cycle. The whole night RIP signals are then processed on the BR module as a 2D image (matrix) with two signals (two rows).

B. BR module

The BR module represents the core of the BR estimation process. The module input is either the RIP signals (2D image, 2 rows) or a presentation of the spectral contents of the audio signal thought the night (2D spectrogram, see section C). Inputs are segmented into either 12s or 20s windows in 5s time increments. Segment lengths are arbitrarily selected in order to capture the several breathing cycles that are needed for BR estimation and yet not too long window, which will result in a smeared BR; a reasonable value of 15 BR/min will contain three and five respiratory cycles in 12s and 20s windows, respectively). Each segment (12s or 20s window, separately) undergoes a *breathing interval estimation* stage. The breathing interval (BI) estimations from both segment lengths will later be averaged, and resampled to the standard 30s epoch rate. The outcome of the *BR module* will be calculated to form a BR per minute using the transformation function BR=60/BI.

C. Breathing interval estimation

The BI estimation block calculates from the segmented window the most likely breathing interval sequence during the night using a periodicity measurement. Periodicity measurement is calculated for every signal *i* (64 audio-related signals or two RIP signals) in a window segment, and it resembles the autocorrelation function as described in (1)

$$P_{i,l} = \frac{1}{T \times FR - l} \sum_{n=0}^{T \times FR} \left(X_{i,n} - \bar{X}_i \right) \times \left(X_{i,n+l} - \bar{X}_i \right), \quad (1)$$

where l represents the time-lag in samples, T represents the window's length in seconds, FR represents the frame-rate for the input signals in Hz. We calculate time-lags up to 75% of window's length to remove non-reliable estimations (insufficient number of samples). Figure 2A demonstrates an example of such P function computed from 20s windows during the night (2D image). In this example, the P function calculated on the averaged 25% (i.e., 64) most periodic frequency components [8] in each segmented window.

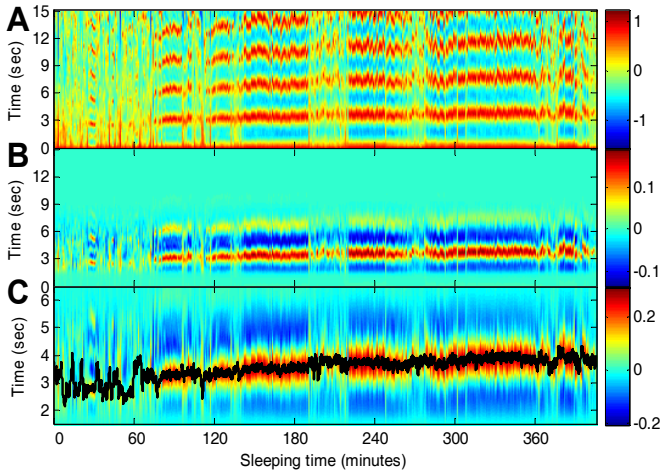

Figure 2: Example of the breathing interval estimation. A) The autocorrelation function P according to Eq.(1). B) The enhanced breathing intervals after emphasis of the general $BIPF_{4,2}$. C) The enhanced breathing intervals after emphasis of the optimal BIPF. The black line represents the shortest path acquired using Dijkstra's algorithm. Y axes are present time-lag in seconds rather than samples for convenient.

Once an enhanced 2D image of autocorrelation along time is achieved, another breathing period emphasis process is applied to the 2D image. The emphasis process consists of two consecutive steps in order to find the individual's *BI* probability function (*BIPF*); see equation (2):

$$BIPF_{\mu,\sigma} = N(\mu,\sigma)\times\left[\frac{\arctan\left(4(t-2)\right)}{\pi}+0.5\right], \qquad (2)$$

where N is the normal distribution with mean of μ seconds, and standard deviation of σ; the *Sigmoid* function (inside the square brackets in (2)) aims to attenuate impossibly low BI; t is the time index (seconds). We formulated this function to achieve realistic BI probabilities based on empiric trials. Figure 3 shows illustrations of three *BIPF* functions.

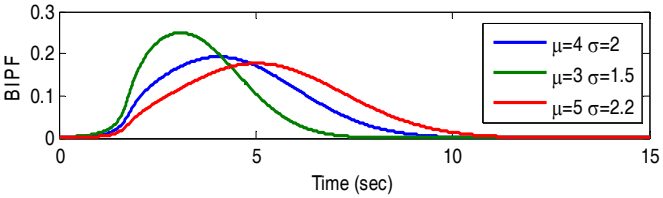

Figure 3: Breathing interval probability function (*BIPF*). This function was formulated in order to emphasize the most probable breathing intervals.

In order to find an individual's BIPF two parameters, the first step includes searching the subject's most recurrent BI according to the autocorrelation peaks histogram extracted from the 2D image (initial values of μ=4s and σ=2, providing a reasonable BR of 15 cycles per minute). See Figure 2B.

Once the most recurrent BI is calculated, the optimal μ and σ are estimated and will be used for breathing emphasis. For further simplicity, we denote the multiplication between the 2D image of autocorrelation and the optimal *BIPF* as **W**. See Figure 2C.

Based on a priori knowledge of the subject's *BI*s, the final stage in the *breathing interval estimation* module is to find the most likely *BI* sequence in **W** throughout the night. In order to do so, we used the Dijkstra shortest-path algorithm [10]. The advantage of this algorithm is its ability to find the optimal path (BI sequence) while coping with "unsuccessful" periodicity measurements during this process. We constructed a nodes path according to the 2D image from left to right (beginning to ending), where each node is connected to nine adjacent nodes located on the right, with a "distance" determined by the corresponding **W** values. For example, the $Node_{l,t}$ is connected to $Node_{l+j,t+1}$ where j is an integer between -4 and +4. This node path connectivity will allow a reasonable dynamic change rate of *BI* (a maximum *BI* change of ±450ms in 30s). A demonstration of the process is presented in Figure 2C. The starting and ending nodes are forced to be between 2s and 8s breathing interval zones (8-30 BR/min).

D. Comparison and Analysis

Based on the two estimations of BR (audio and PSG), we evaluated within- and between-subjects comparisons. Within-subjects comparison for each individual's BR estimations (epoch by epoch) included: mean absolute error (*MAE*), mean square error (*MSE*), mean relative error (*MRE*), and Pearson's correlation (R_{ind}). Between-subjects comparison included a global score of correlation (R_{glob}) to the subject's mean BR using both approaches. The influences of anthropometric parameters on BR estimation results were also examined.

III. Results and Discussion

Two major types of measurements were executed in order to assess the similarities between *BR* estimations according to the two approaches (audio and PSG). Figure 4 presents global estimation of a subject's mean BR using both approaches (left panel), and histograms of within-subjects', epoch-by-epoch, similarity scores (right panels).

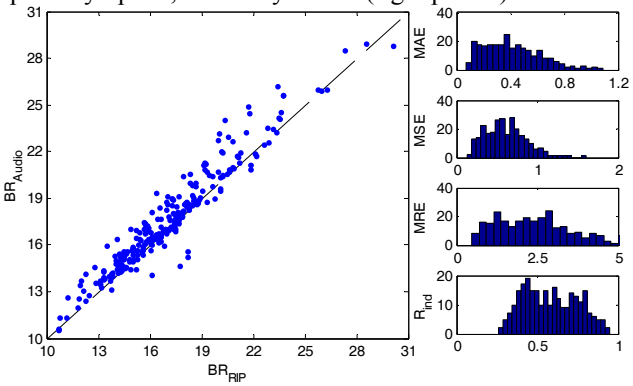

Figure 4: Comparison of BR estimations. On the left panel, mean BR according to audio and PSG approaches – each dot represents an individual subject. The dashed line represents the identity y=x. To the right, histograms of several comparison (epoch-by-epoch) measurements of *BR* for each individual, including mean absolute error (MAE), mean squared error (MSE), mean relative error (MRE) in %, and Pearson's correlation (R).

A detailed analysis is presented in Table I. The comparison also evaluates the similarities under anthropometric parameters' break-down including: gender, age, body mass index (BMI), obstructive sleep apnea severity of apnea hypopnea index (AHI), and snores' intensity (OSI [7]).

TABLE I. BREATHING RATE ESTIMATIONS AND COMPARISON

Group	*N*	*Between subjects*			*Within subjects*			
		Audio	RIP	R_{glob}	*MAE*	*MSE*	*MRE*	R_{ind}
All+PSG	264	17.6 (3.3)	17.1 (3.2)	0.972	0.42 (0.21)	0.61 (0.25)	2.44 (1.20)	0.68 (0.16)
Edirol+PSG	204	17.6 (3.4)	17.0 (3.3)	0.974	0.41 (0.22)	0.61 (0.26)	2.43 (1.24)	0.68 (0.16)
LS-5 +PSG	60	17.5 (3.1)	16.9 (3.0)	0.966	0.42 (0.19)	0.58 (0.21)	2.47 (1.07)	0.68 (0.17)
Males	173	17.4 (3.3)	16.9 (3.2)	0.977	0.39 (0.20)	0.58 (0.24)	2.34 (1.15)	0.70 (0.17)
Females	91	18.0 (3.2)	17.4 (3.0)	0.960	0.47 (0.22)	0.67 (0.27)	2.71 (1.27)	0.64 (0.15)
Age≤50	104	17.2 (3.5)	16.8 (3.3)	0.989	0.35 (0.19)	0.54 (0.24)	2.09 (1.01)	0.73 (0.16)
Age>50	160	17.7 (3.2)	17.1 (3.2)	0.963	0.45 (0.22)	0.64 (0.25)	2.64 (1.27)	0.66 (0.16)
BMI≤30	119	16.7 (3.0)	16.3 (3.0)	0.995	0.37 (0.18)	0.57 (0.22)	2.28 (1.00)	0.68 (0.16)
BMI>30	145	18.3 (3.4)	17.6 (3.3)	0.963	0.45 (0.24)	0.63 (0.27)	2.56 (1.34)	0.68 (0.17)
AHI≤15	136	17.4 (3.2)	16.9 (3.1)	0.970	0.41 (0.21)	0.61 (0.21)	2.35 (1.21)	0.69 (0.16)
AHI>15	128	17.7 (3.4)	17.2 (3.4)	0.975	0.41 (0.21)	0.60 (0.20)	2.40 (1.20)	0.69 (0.17)
OSI≤40	82	17.2 (3.1)	16.4 (2.7)	0.972	0.47 (0.23)	0.69 (0.28)	2.88 (1.28)	0.65 (0.15)
OSI:40-55	127	17.8 (3.3)	17.2 (3.2)	0.975	0.43 (0.21)	0.63 (0.24)	2.48 (1.17)	0.67 (0.16)
OSI>55	55	17.6 (3.7)	17.1 (3.6)	0.987	0.36 (0.22)	0.51 (0.22)	2.11 (1.13)	0.72 (0.18)

Values are mean (std). Parameters are: mean absolute error (MAE), mean squared error (MSE), mean relative error (MRE) in %, and Pearson's correlation (R).

No significant differences were observed in the anthropometric parameters including gender, age, BMI, AHI, and OSI. This finding indicates the robustness of the proposed BR estimation approach using non-contact microphone.

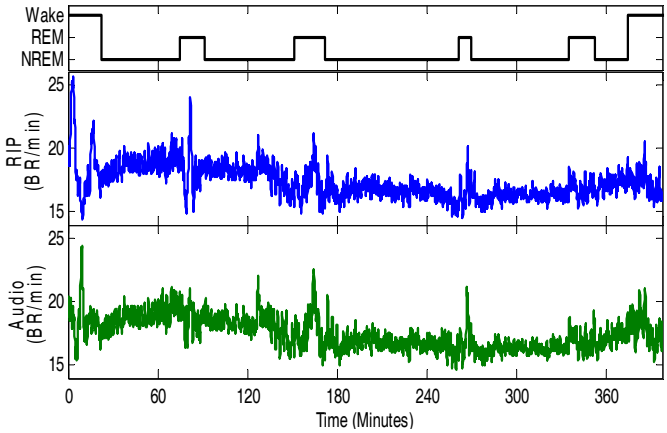

Figure 5: Typical example of BR estimation using RIP and audio-based methods (Male, age=54, BMI=27, AHI=9). Upper panel, macro sleep stages: wake, rapid-eye-movement (REM) sleep, and non-REM sleep according to the gold-standard PGS. In this example, similarity tests revealed: *MAE*=0.72, *MSE*=0.81, *MRE*=2.01% and *R*=0.63. Note that during REM sleep, *BR* pattern has high variability, suggesting high cognitive activity.

Figure 5 shows a typical example of BR estimations using both approaches. In addition, this example shows that BR measurement possesses some information that can indicate subject's sleep stages and even provide a bio-marker for sleep apnea presence (high BR variability).

IV. CONCLUSIONS AND FUTURE WORK

In this study we proposed a novel and robust method for estimation of BR using non-contact audio analysis. This method does not interfere with subjects' sleep, thus preserving natural sleep. The proposed algorithm can be applied to a variety of systems that require on-going BR monitoring. Moreover, BR estimation may be useful for assisting sleep/breathing diagnostic systems that rely on audio signals [7-9, 11-13]. Encouraged by our findings (Figure 5) regarding the relationship between BR dynamics and macro sleep stages, in the future we plan to explore the ability to differentiate between REM-sleep and non-REM sleep based on BR properties.

ACKNOWLEDGMENT

We would like to thank Mrs. Bruria Freidman from the Sleep Wake Disorder Unit of Soroka University Medical Center, for her support and collaboration.

REFERENCES

[1] A. Bates, M. J. Ling, J. Mann, and D. Arvind, "Respiratory rate and flow waveform estimation from tri-axial accelerometer data," in *Proc. Int. Conf. Body Sensor Networks*, 2010, pp. 144-150.

[2] W. Johnston and Y. Mendelson, "Extracting breathing rate information from a wearable reflectance pulse oximeter sensor," in *Conf Proc IEEE Eng Med Biol Soc.*, 2004, pp. 5388-5391.

[3] R. Murthy, I. Pavlidis, and P. Tsiamyrtzis, "Touchless monitoring of breathing function," in *Conf Proc IEEE Eng Med Biol Soc.*, 2004, pp. 1196-1199.

[4] C. Worsnop, A. Kay, Y. Kim, J. Trinder, and R. Pierce, "Effect of age on sleep onset-related changes in respiratory pump and upper airway muscle function," *Journal of Applied Physiology,* vol. 88, pp. 1831-1839, 2000.

[5] B. A. Edwards and D. P. White, "Control of the pharyngeal musculature during wakefulness and sleep: implications in normal controls and sleep apnea," *Head & neck,* vol. 33, pp. S37-S45, 2011.

[6] V. Hoffstein, S. Mateika, and D. Anderson, "Snoring: is it in the ear of the beholder?," *Sleep,* vol. 17, pp. 522-6, Sep 1994.

[7] E. Dafna, A. Tarasiuk, and Y. Zigel, "Automatic Detection of Whole Night Snoring Events Using Non-Contact Microphone," *PLoS One,* vol. 8, p. e84139, 2013.

[8] E. Dafna, A. Tarasiuk, and Y. Zigel, "Sleep-Wake evaluation from whole-night non-contact audio recordings of breathing sounds," *PLoS One,* vol. 10, p. e0117382, 2015.

[9] T. Rosenwein, E. Dafna, A. Tarasiuk, and Y. Zigel, "Detection of breathing sounds during sleep using non-contact audio recordings," *Conf Proc IEEE Eng Med Biol Soc.,* 2014.

[10] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische mathematik,* vol. 1, pp. 269-271, 1959.

[11] N. Ben-Israel, A. Tarasiuk, and Y. Zigel, "Obstructive apnea hypopnea index estimation by analysis of nocturnal snoring signals in adults," *Sleep,* vol. 35, p. 1299, 2012.

[12] E. Dafna, A. Tarasiuk, and Y. Zigel, "Sleep-quality assessment from full night audio recordings of sleep apnea patients," in *Conf Proc IEEE Eng Med Biol Soc.*, 2012, pp. 3660-3663.

[13] E. Dafna, A. Tarasiuk, and Y. Zigel, "OSA severity assessment based on sleep breathing analysis using ambient microphone," in *Conf Proc IEEE Eng Med Biol Soc.*, 2013, pp. 2044-2047.