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

Sleep-Quality Assessment from Full Night Audio Recordings of Sleep Apnea Patients

E. Dafna, A. Tarasiuk, Y. Zigel, *member, IEEE*

Abstract: **In this work, a novel system (method) for sleep quality analysis is proposed. Its purpose is to assist an alternative non-contact method for detecting and diagnosing sleep related disorders based on acoustic signal processing. In this work, audio signals of 145 patients with obstructive sleep apnea were recorded (more than 1000 hours) in a sleep laboratory and analyzed. The method is based on the assumption that during sleep the respiratory efforts are more periodically patterned and consistent relative to a waking state; furthermore, the sound intensity of those efforts is higher, making the pattern more noticeable relative to the background noise level. The system was trained on 50 subjects and validated on 95 subjects. The system accuracy for detecting sleep/wake state is 82.1% (epoch by epoch), resulting in 3.9% error (difference) in detecting sleep latency, 11.4% error in estimating total sleep time, and 11.4% error in estimating sleep efficiency.**

Keywords: **Sleep Quality Estimation, Obstructive Sleep Apnea, Audio Signal Processing, Snore Detection.**

I. Introduction

Obstructive sleep apnea (OSA) is a prevalent sleep breathing disorder affecting 2% to 7% of adults that can lead to considerable morbidity. Partial or complete collapse of the upper airway during sleep has different effects on the human body, ranging from noisy breathing (snoring) to cardiovascular morbidity [1].

A patient with OSA usually snores during sleep. The snoring is then interrupted by a long silent period during which there is no breathing. This is followed by a loud snort and gasp, as the patient tries to breathe. This pattern usually repeats [2].

The most common approach to diagnose OSA is attended polysomnographic (PSG) study in the sleep laboratory. During PSG the subject is connected to numerous electrodes and sensors attached literally from head to toe, monitoring biological signals such as EEG, EMG, ECG, and respiratory efforts. PSG study is uncomfortable and expensive, [2] and sleep conditions are un-ideal. This led to seeking alternative methods of OSA diagnosis and sleep quality assessment as a whole.

In order to reliably estimate pathological respiratory events, as well as other sleep disorders during sleep, it is important to accurately estimate sleep quality parameters [3,4]. A typical sleep quality parameter set may include: 1) *total sleep time* (TST) – the overall duration of sleep stages when concatenated, 2) *sleep latency* (SL) – the elapsed time to falling asleep from lying in bed, 3) *sleep efficiency* (SE) – the ratio between TST and total time in bed, 4) *wake-time after sleep onset* (WASO) – the summation of all awakening episodes during sleep, and 5) *awakening index* (AI) – the average awakening per hour.

Recently, the use of ambulatory at-home respiratory measurement technologies has become more frequently used. In these technologies the assessment of sleep/wake pattern is determined by a sleep diary completed by the patient and sometimes by a single channel EEG [5] and/or built-in actigraphy technology [3,6], and cardiac tempo [7,8]. Following whole-night data acquisition, sleep technologists must review the data in order to determine pathological respiratory events and sleep quality parameters. The main disadvantage of the available sleep quality technologies is that these methods require wearing costly sensors and devices; these devices are sensitive to use, they require ongoing maintenance, and tend to break. To the best of our knowledge, there is no technology available to estimate sleep/wake pattern using non-contact inexpensive technology.

This paper describes a pioneering attempt to estimate sleep quality parameters using only audio signal. In this study we propose the use of a non-contact sensor – a microphone and a digital audio recorder that records the patient sounds from a one meter distance and enables a more natural sleep, and therefore more reliable sleep quality parameters.

II. Methods

One hundred forty-five patients (over 18 years old) scheduled for the sleep laboratory were recorded during a full-night with a digital audio recorder device (EDIROL R-4) using a directional microphone (RODE NTG-1) at a distance of 1 meter above the head level and stored along with the PSG signals; the acquired audio signals are digitized at a sampling frequency of 44.1 kHz, PCM, 16 bits per sample. The raw audio signal is processed using the proposed system, which is shown in Figure 1.

A. Preprocessing & noise reduction

The full-night audio signal is down-sampled to 16 kHz. A noise reduction (spectral subtraction) algorithm is applied to the signal based on the Wiener-filter, which is based on tracking a priori SNR using the decision-directed method proposed by Scalart et al. [9]; this step is very important since it emphasizes the non-stationary events that were recorded during sleep such as snoring and low breaths.

*This work was supported in part by the Israeli Ministry of Industry and Trade, The Kamin Program, award no. 46168.

E. Dafna is with the Department of Biomedical Engineering, Ben-Gurion University of the Negev, Beer–Sheva, Israel (elirand@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: +972-8-6428372; fax: +972-8-6428371; e-mail: yaniv@bgu.ac.il).

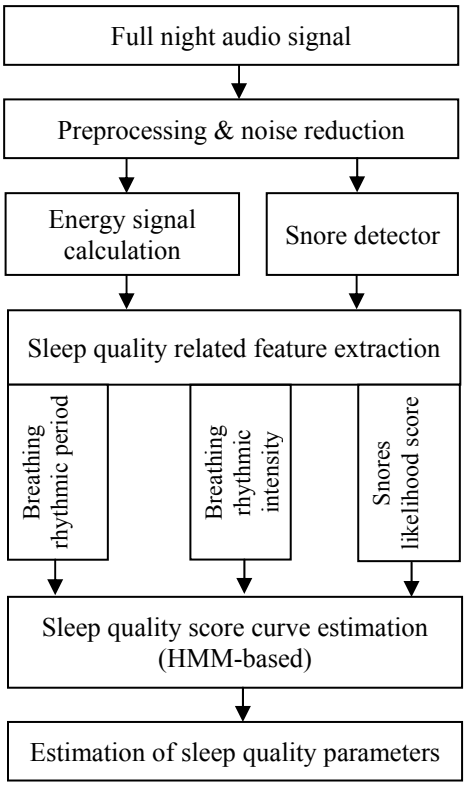

Figure 1. Block diagram of the system.

B. *Energy signal calculation*

Once the audio signal is enhanced using the preprocessing stage, an energy signal, $e(n)$, is calculated using a 60 ms Gaussian window with 75% overlap. The energy values are stored in dB scale in order to further emphasize the low events.

C. *Snore detection algorithm*

The snore detection algorithm consists of two main steps [10,11]: 1) *audio event detection* – audio events are automatically detected and segmented using an adaptive energy threshold and 2) *snore detection* – based on Gaussian mixture model (GMM). Each audio event is scored by a likelihood ratio score that represents the probability that this event is a snore event.

D. *Sleep quality related feature extraction*

Using the energy signal, $e(n)$, two conjugated features are extracted: *Breathing rhythmic period* and *Breathing rhythmic intensity*. The purpose of those features is to evaluate the rhythmic pattern of respiratory action; the more periodic the respiratory action the better the sleep quality; any movement or body posture changes will be translated as temporary poorer sleep quality. In order to estimate the rhythmic features, the energy signal was segmented into 24-second segments with 19-second overlap (i.e., 5 second resolution). From each segment the autocorrelation vector $R(\tau)$ was calculated and the first peak within the range of 1 sec to 10 sec was picked; this peak holds information about the basic rhythm period (location of the peak) and its intensity (its amplitude relative to its neighbors). The intensity feature, R_I, is calculated as the product of the first peak amplitude value, $R(\tau_{peak})$, and the initial correlation area, *Area*

$$R_I = R(\tau_{\text{peak}}) \cdot Area \quad (1)$$

where *Area* is actually the normalized square area between the $R(\tau)$ curve and the $a\tau+1$ linear line from $R(0)$ to $R(\tau_{\text{peak}})$, calculated as:

$$Area = \frac{1}{\tau_{peak}} \sum_{\tau=0}^{\tau_{peak}} (a\tau + 1 - R(\tau))^2 \quad (2)$$

where a is the estimated linear slope. The more harmonic the energy pattern the greater is $R(\tau_{peak})$ and the greater the *Area*. A demonstration is shown in Figure 2.

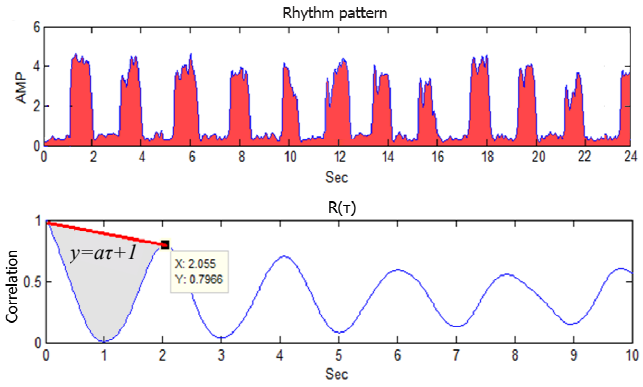

Figure 2. Rhythm pattern and correlation. Upper panel represents the energetic pattern of snores and breathing episodes. Bottom panel is the autocorrelation $R(\tau)$ of the segment. Note the rhythmic period is located at about 2 seconds and its intensity is calculated over the captured area between $R(\tau)$ and the line connected between $[R(0), R(\tau_{peak})]$.

The snores likelihood score (SLS) feature is calculated as the maximum of event scores $s(x)$ within a one minute segment (segment rate of 5 seconds). Calculated as:

$$SLS = \max_i \left(s(\mathbf{x}_i)\right) = \max_i (\log p(\mathbf{x}_i|\lambda_s) - \log p(\mathbf{x}_i|\lambda_n)) \quad (3)$$

where λ_s and λ_n are the snore model and noise model, respectively, and $p(\mathbf{x}_i|\lambda)$ is the probability of event feature vector $\mathbf{x}_i$ given the model λ.

E. *Sleep quality score curve estimation*

Using the three sleep quality features, a sleep quality analysis (SQA) score curve is estimated for each patient. Since the sleep quality parameters are directly connected with two sleep phases, sleep and awake, a two-state hidden Markov model (HMM) is used; one state represents the sleep phase and the other state represents the awake phase (Figure 3).

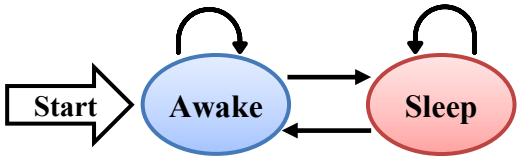

Figure 3. Two-states HMM.

The HMM transition matrix was estimated using the labeled (30 sec) PSG epochs from the design database. Each state probability density function (pdf) is estimated using the labeled feature vectors. In the validation phase, integrating the transition probability with the pdf score of each feature vector, an SQA score is calculated – one score for every feature vector (5 second resolution). This SQA score represents a continuous scale between "Awake" (poor sleep quality) and "Sleep" (good sleep quality). A demonstration of the SQA score curve is shown in Figure 4 and Figure 5.

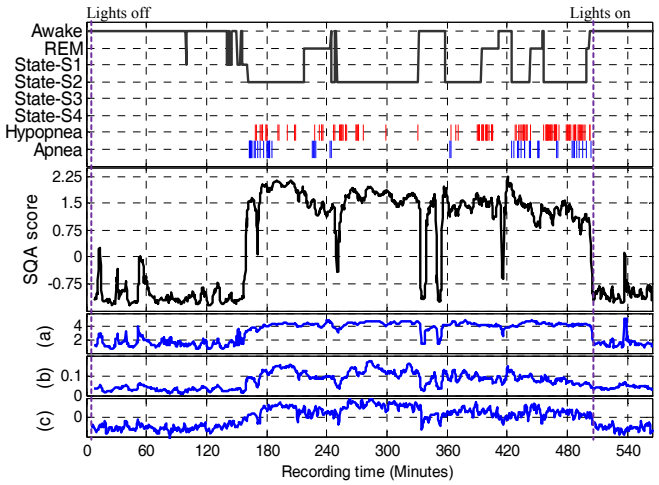

Figure 4. A sleep pattern example of whole night recording from an OSA patient (female, age 66, BMI 41, AHI 32). Upper figure: the sleep state along with the manual annotations of OSA events monitored with PSG test. Lower figure: the SQA score and the three extracted features. (a) *Breathing rhythmic period*, (b) *Breathing rhythmic intensity*, (c) *Snores likelihood scores*.

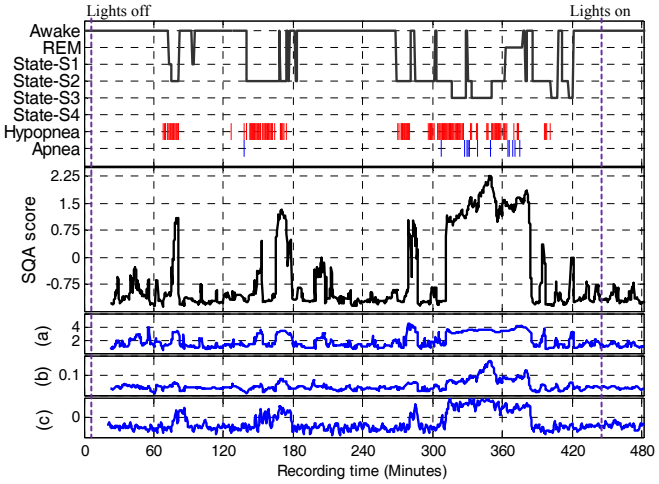

Figure 5. Another example of a sleep pattern of a patient with OSA and insomnia (male, age 80, BMI 39, AHI 51). See Figure 4 for details

F. Estimation of sleep quality parameters

Looking at Figure 4 and Figure 5 one can see that the awake/sleep phase can be easily determined from the SQA score curve by a discriminating threshold (e.g., at SQA score = 0). Affected by the demand for the application, one can arbitrarily formulate a measure from the full night curve (e.g., amount of continuous sleep stage, the variance within sleep stage, and so on).

We chose to estimate the five most common sleep quality parameters: SL, SE, TST, WASO, and AI (see Introduction).

III. RESULTS

The experiment (study) was conducted using the database that is shown in Table I. In order to train (design) and validate our system, a common sample rate (segment rate) for the PSG and the SQA is necessary; therefore, we low-pass-filtered and resampled every feature vector to a 30 second rate so it can be compared with the accepted 30 sec PSG epoch.

TABLE I. PATIENT CHARACTERISTICS.

	Design	**Validation**
Number of Patients	50	95
Male/Female	31/19	59/36
AGE (years)	52.8 ± 12.6	56.6 ± 16.0
BMI (kg/m^2)	31.7 ± 5.0	32.6 ± 6.4
AHI (events/hr)	20.4 ± 18.5	21.6 ± 18.5
Recorded length (min)	443.1 ± 63.2	451.9 ± 66.1

BMI – Body mass index, AHI – Apnea and hypopnea index.
Values are mean ± SD

Table II shows the mean and SD values of the estimated sleep quality parameters for the design and validation study during PSG and the proposed SQA.

TABLE II. SLEEP PARAMETERS FOR THE DESIGN AND VALIDATION STUDY.

	Design (N=50)		**Validation(N=95)**	
	PSG	**SQA**	**PSG**	**SQA**
SL (min)	59.2 ± 50.0	59.7 ± 50.8	64.3 ± 69.0	54.8 ± 59.2
SE (%)	65 ± 14	69 ± 17	65 ± 13	69 ± 16
TST (min)	286 ± 62	307 ± 84	290 ± 58	309 ± 68
WASO (min)	47 ± 36	48 ± 58	43 ± 31	52 ± 54
AI (e/hr)	5.1 ± 4	5.5 ± 6.6	4.7 ± 3.3	5.3 ± 5.1

SL – sleep latency, SE – sleep efficiency, TST – total sleep time, SE – sleep efficiency, WASO – wake-time after sleep onset, AI – awakening index.
Values are mean ± SD

The accuracy of the proposed method was tested by analyzing the SQA score of each epoch, using the PSG labeled (sleep/wake) epochs as a gold standard. System performances were evaluated by ROC curve (Figure 6), where true positive (TP) is the correct detection rate of sleep phases, and false positive (FP) is the incorrect detection of awake phases as sleep.

The overall accuracy in detecting sleep phases (epoch by epoch) was found to be 82.1%.

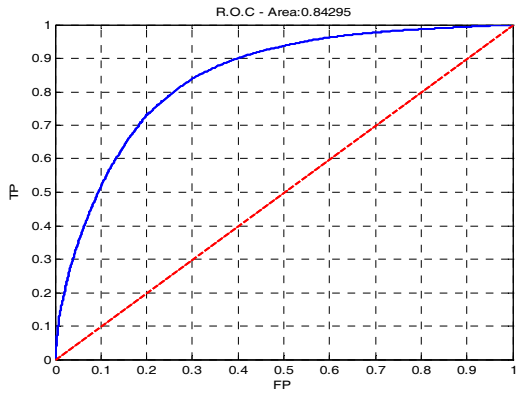

Figure 6. Performance evaluation, true positive (TP) detection of sleep phases vs. false positive (FP).

Bland-Altman plot [12] is a method of data plotting used, particularly, in analyzing the agreement between two different medical measures. Examining the Bland and Altman plots and comparing the proposed SQA vs. the gold-standard PSG sleep quality parameters (Figure 7) showed no major consistent bias. The plots also show (for SL, WASO, and AI) that the SQA values are closely matched to PSG in the left part of the plot, and therefore present less difference when the parameter values are relatively small; this implies reliability of the estimated parameters.

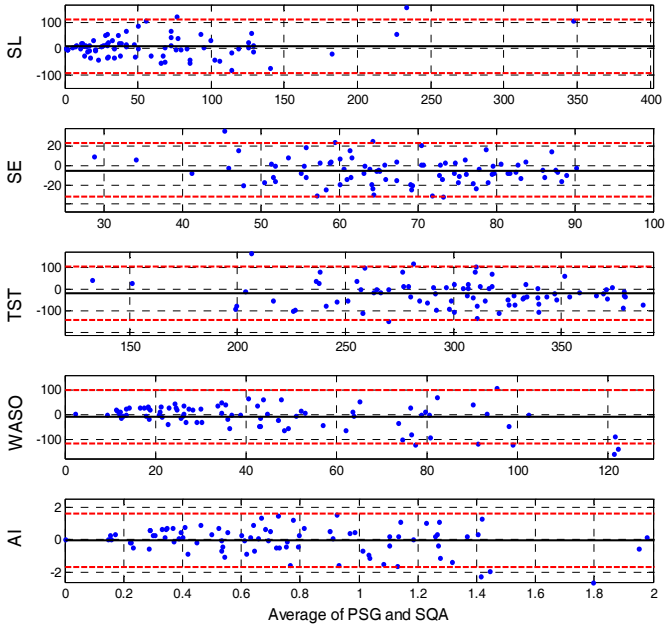

Figure 7 – Blend-Altman plot of the PSG & SQA. X-axis represents the mean sleep parameter value between the PSG & SQA in the relevant parameter units (SL, TST, and WASO in minutes, SE in %, and AI in e/hr). The Y-axes represent the difference between the PSG and SQA parameter. The solid line represents the mean difference and dashed lines represent the 95% confidence interval.

When calculating the absolute difference between PSG & SQA normalized by the record (signal) duration, the SL disagreement is 3.9 ± 5.4% (mean ± SD), TST and SE is 11.4 ± 9.1%, and WASO is 8.2 ± 8.3%. The absolute difference between the AI parameters is 0.64 ± 0.54 e/hr.

IV. Discussion and Conclusion

The experiment was conducted utilizing more than 1000 hours of audio recordings taken from 145 patients (50 for design, 95 for validation) during PSG study for OSA diagnosis as shown on Table 1; both groups were similar in demographic and PSG characteristics. According to the SQA score curve in Figures 4 and 5 one can see that there is a good correlation between the sleep pattern (sleep/wake states from the gold-standard PSG) and the proposed SQA curves. Since the SQA score curve is calculated using three sleep quality features, among them the snore likelihood score feature, a robust and low-false alarmed rate snore detector is necessary. We used our high-performance snore detector (97% accuracy with area of 0.9927 under ROC) [10]. Comparing our estimated sleep quality parameters with the wrist actigraphy method (Hedner et al. [3]), we found similar results.

This paper proposed a novel non-contact method for estimation of sleep quality parameters from patients undergoing OSA diagnosis based on analyzing full night audio signals. The performances of the system are very encouraging and could serve as a basis to estimate sleep-wake patterns using a simple single-channel non-contact audio technology.

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] H.K. Yaggi, J. Concato, W.N. Kernan, J.H. Lichtman, L.M. Brass, V. Mohsenin, "Obstructive sleep apnea as a risk factor for stroke and death," *N Engl J Med*, vol. 353, pp. 2034-2041, 2005.

[2] W. Flemons, "Obstructive sleep apnea," *N Engl J Med*, vol. 347, pp. 498-504, 2002.

[3] J. Hedner, G. Pillar, S.D. Pittman, D. Zou, L. Grote, D.P. White, "A novel adaptive wrist actigraphy algorithm for sleep-wake assessment in sleep apnea patients," *Sleep*, vol. 27, N. 8, pp. 1560-1566, 2004.

[4] T. Blackwell, S. Ancoli-Israel, S. Redline, K.L. Stone, "Factors that may influence the classification of sleep-wake by wrist actigraphy," *J Clin Sleep Med*, vol. 15, pp. 357-367, 2011.

[5] S. Cho, J. Lee, H. Park, K. Lee, "Detection of arousals in patients with respiratory sleep disorders using a single channel EEG," *Conf Proc IEEE Eng Med Biol Soc.*, vol. 3, pp. 2733-2735, 2005.

[6] N.L. Johnson, H.L. Kirchner, C.L. Rosen et al. "Sleep estimation using wrist actigraphy in adolescents with and without sleep disordered breathing: a comparison of three data modes," *Sleep,* vol. 30, pp. 899–905, 2007.

[7] B.H. Choi, G.S. Chung, J.S. Lee, D.U. Jeong, K.S. Park, "Slow-wave sleep estimation on a load-cell-installed bed: a non-constrained method," *Physiol. Meas*., vol. 30, pp. 1163-1170, 2009.

[8] T. Fukuda, Y. Wakuda, Y. Hasegawa, F. Arai, M. Kawaguchi, A. Noda, "Sleep Quality Estimation based on Chaos Analysis for Heart Rate Variability," *IEEE Transactions on Electronics, Information and Systems*. vol. 125, pp. 43-49, 2005.

[9] P. Scalart, J.V. Filho, "Speech enhancement based on a priori signal to noise estimation," *Conf Proc IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, pp. 629-632, 1996.

[10] E. Dafna, A. Tarasiuk, Y. Zigel, "Automatic detection of snoring events using Gaussian mixture models," *Proc 7th international workshop of models and analysis of vocal emissions for biomedical applications, MAVEBA*, 2011.

[11] N. Ben-Israel, A. Tarasiuk, Y. Zigel, "Obstructive apnea hypopnea index estimation by analysis of nocturnal snoring signals in adults," *Sleep*, 2012. (Epub ahead of print)

[12] J.M. Bland, D.G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *Lancet*, vol. 1, pp. 307-310, 1986.