AUTOMATED DETECTION OF SLEEP APNEA AND HYPOPNEA EVENTS BASED ON ROBUST AIRFLOW ENVELOPE TRACKING

Marcin Ciołek**, Maciej Niedźwiecki***, Stefan Sieklicki***, Jacek Drozdowski*† *and Janusz Siebert*‡

* Department of Automatic Control, Gdańsk University of Technology, Gdańsk, Poland
† Department of Pneumology and Allergology, Medical University of Gdańsk, Gdańsk, Poland
‡ Department of Family Medicine, Medical University of Gdańsk, Gdańsk, Poland

ABSTRACT

The paper presents a new approach to detection of apnea/hypopnea events, in the presence of artifacts and breathing irregularities, from a single-channel airflow record. The proposed algorithm identifies segments of signal affected by a high amplitude modulation corresponding to apnea/hypopnea events. It is shown that a robust airflow envelope—free of breathing artifacts—improves effectiveness of the diagnostic process and allows one to localize the beginning and the end of each episode more accurately. The performance of the approach, evaluated on 30 overnight polysomnographic recordings, was assessed using diagnostic measures such as accuracy, sensitivity, specificity, and Cohen's coefficient of agreement; achieving 95%, 90%, 96%, and 0.82, respectively.

Index Terms— Apnea, envelope detectors, median filters

1. INTRODUCTION

Sleep Apnea-Hypopnea Syndrome (SAHS) is a common sleep-breathing disorder characterized by repetitive episodes of complete obstruction (sleep apnea event) or partial obstruction (hypopnea event) of the upper airway, relating in a blood oxygen desaturation or arousals leading to sleep fragmentation. The usual daytime manifestation is excessive sleepiness and poor concentration, which can escalate to traffic accidents, depression, and memory loss [1].

Currently, in sleep laboratories, there are carried out overnight polysomnographic studies (PSG) aimed at early detection and assessment of the severity of SAHS in patients [1]. To reach the final conclusion, the recorded signals are analysed by a physician experienced in the field of sleep medicine. The final diagnosis is based on calculation of the apnea/hypopnea index (AHI) which reflects the number of sleep apnea/hypopnea (SAH) events per hour of sleep. It is assumed that accurate and reliable identification of SAH events is critical for case identification and for quantifying disease severity classified as: mild, when $5 \leq$ AHI < 15; moderate, when $15 \leq$AHI < 30; and severe, when AHI ≥ 30 (events per hour of sleep) [1].

The standard morphological criteria, given by the American Academy of Sleep Medicine (AASM) [1], describe SAH events as a significant reduction of airflow amplitude lasting at least 10 s. The reduction of the airflow amplitude is observed in relation to the level of breathing amplitude preceding and succeeding the respiratory event, further called the baseline value. Two thresholds of airflow reduction, 50% and 90%, were accepted to represent partial and complete obstruction of the upper airway, respectively. According to AASM the breathing baseline is defined as "the mean amplitude of stable breathing and oxygenation in the 2-min preceding onset of the event (in individuals who have a stable breathing pattern during sleep) or the mean amplitude of the three largest breaths in the 2-min preceding onset of the event (in individuals without a stable breathing pattern)". Although sufficiently clear for pulmonologists, the verbal description cited previously is not easy to turn into a reliable algorithm for automatic SAH scoring. In the presence of abnormally large peaks, further called breathing artifacts, automatic scoring becomes even more difficult.

Irregular breathing artifacts are usually associated either with the patient's motion during sleep [see Fig. 1(A)] (rapid body movements, changing of sleep position), or with a sudden opening of the upper airway succeeding a sleep apnea event [see Fig. 1(B)]. Such artifacts in airflow measurements can lead to incorrect identification of SAH events by automated sleep scoring methods, which in turn may result in incorrect diagnosis of the SAHS syndrome. For this reason, the physician localizes, based on visual inspection, signal segments corrupted by artifacts and manually marks them as the ones that should be ignored during automated sleep scoring. This is a very time-consuming process, and a subjective judgment is required to complete the job. Unsatisfactory quality of automated analysis based on the airflow signal is often caused by problems with adaptive tracking of the baseline value, with respect to which SAH events are identified. In the presence of artifacts, the correct morphological description of the baseline value is not a trivial task [2].

The main contribution of this paper is to demonstrate that SAH events can be effectively identified, in the presence of artifacts and breathing irregularities, based on analysis of a robust airflow envelope.

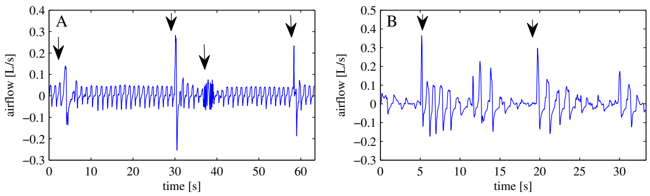

Fig. 1: Breathing artifacts : (A) rapid body movements, changing of sleep position and (B) sudden opening of the upper airway succeeding a sleep apnea/hypopnea event.

2. ENVELOPE DETECTION

During a routine analysis, a physician can easily track the "true" signal envelope using visual inspection, even in the presence of artifacts. Such an envelope corresponds to a smooth curve that matches, in a way that is robust to breathing artifacts, the main peaks of the waveform, and follows closely sudden variations in signal amplitude. If a partial or complete reduction of airflow takes place, then SAH events take the form of characteristic "valleys" visible in the signal envelope. A physician identifies and classifies these valleys as hypopnea or sleep apnea events, based on the standard morphological criteria and his/her own experience. In the proposed approach, we try to reproduce such a procedure.

Envelope detection has numerous applications in the field of signal processing and communications [3], one of which is demodulation of amplitude-modulated (AM) signals governed by

$$s(t) = A[1 + \beta m(t)] \cos \omega_c t \tag{1}$$

where $t = \dots, -1, 0, 1, \dots$ denotes normalized (dimensionless) discrete time, $m(t) \geq 0$ denotes the baseband message, ω_c denotes carrier (angular) frequency, A denotes the carrier amplitude, and $\beta > 0$ is the so-called amplitude sensitivity of the modulator. When $m(t)$ is a lowpass signal with bandwidth W much smaller than the carrier frequency ω_c, the amplitude envelope of the signal $s(t)$ is defined as

$$e(t) = A[1 + \beta m(t)] \geq 0. \tag{2}$$

Envelope can be "extracted" from the AM signal using devices known as envelope detectors. The two most frequently used envelope detectors that will be briefly described below are those based on the square-law (full rectification) principle, and on the Hilbert transform, respectively. Due to irregularities in the breathing rhythm, the airflow signal only approximately fits the AM model (1) which adversely affects performance of the classical envelope detectors. The situation becomes even more complicated in the presence of breathing artifacts. We will show that both problems indicated previously can be taken care of, if the classical detection schemes are suitably modified.

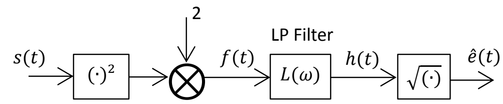

Fig. 2: The square-law envelope detector.

2.1. Square-law envelope detector

The flowchart of the square-law detector [3] is shown in Fig. 2. When $s(t)$ is an AM signal governed by (1), the scaled output of the squaring device can be written down as a sum of two components

$$f(t) = A^2[1 + \beta m(t)]^2 + A^2[1 + \beta m(t)]^2 \cos 2\omega_c t. \tag{3}$$

The first term on the right-hand side of (3) is a lowpass signal with a cutoff frequency $2W$, and the second one is a bandpass signal with spectrum confined to the frequency bands $(-2\omega_c - 2W, -2wc + 2W)$ and $(2\omega_c - 2W, 2\omega_c + 2W)$, centered around $\pm 2\omega_c$. Hence, when the condition $\omega_c > 2W$ is met, the lowpass component of $f(t)$ can be extracted using a lowpass FIR filter $L(\omega)$ with a cutoff frequency $2W$

$$h(t) = L[f(t)] = \sum_{i=-k}^{k} l_i f(t - i) \cong A^2[1 + \beta m(t)]^2 \tag{4}$$

where $l_i = l_{-i}, i = -k, \dots, k$, denote impulse response coefficients of the filter. The estimated value of the envelope can be obtained from

$$\widehat{e}(t) = \sqrt{h(t)} \tag{5}$$

further referred to as an ESL envelope.

2.2. Envelope detector based on the Hilbert transform

The second classical method of envelope detection, based on the Hilbert transform [3], is depicted in Fig. 3. The Hilbert transform of an analog real-valued finite energy signal $s(t_c)$, $-\infty < t_c < \infty$, is defined as [3]

$$s_H(t_c) = \mathcal{H}[s(t_c)] = \int_{-\infty}^{+\infty} \frac{s(\tau) d\tau}{\pi(t_c - \tau)} \tag{6}$$

where t_c denotes the continuous-time variable. Envelope detection involves creation of the complex-valued analytic signal, defined as

$$y(t_c) = s(t_c) + js_H(t_c). \tag{7}$$

The analytic counterpart of a discrete-time signal $s(t)$ can be evaluated either directly - using the FFT-based frequency-domain approach, or indirectly - by computing an "analytic" signal $s_H(t)$ (using the discrete-time FIR approximation of the Hilbert transform) and combining it with $s(t)$

$$y(t) = s(t) + js_H(t). \tag{8}$$

Since Hilbert transform shifts the phase of all sinusoidal components by $-\pi/2$, for the "ideal" AM signal (1), one obtains

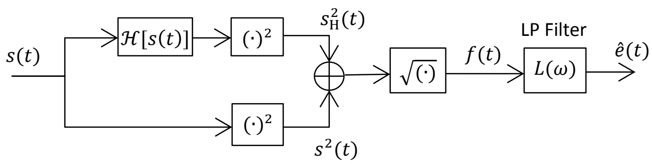

Fig. 3: Envelope detector using the Hilbert transform.

$s_{\mathrm{H}}(t) \cong A[1+\beta m(t)] \sin \omega_{\mathrm{c}} t$ which means that the envelope of $s(t)$ can be obtained by evaluating the magnitude of the analytic signal

$$f(t) = |y(t)| = \sqrt{s^2(t) + s_{\mathrm{H}}^2(t)} \cong A[1 + \beta m(t)]. \qquad (9)$$

For AM-like signals, such as speech signals or biomedical signals, the envelope extracted in this way suffers from high-frequency fluctuations, called ripples. Ripples can be removed by passing the signal $f(t)$ through a lowpass filter, leading to

$$\widehat{e}(t) = L[f(t)] \qquad (10)$$

further referred to as an EHT envelope.

2.3. Modification in the envelope detection procedures

The estimation of the airflow envelope obtained using the approach based on the square-law or on the Hilbert transform suffers from envelope distortions caused by artifacts which are only partially eliminated at the stage of lowpass filtering. The envelope distortions of short duration and of relatively high amplitude may seriously affect the estimated baseline values by setting them at too high levels. This may lead to a large number of false-positive decisions, some of which cause erroneous distinction between hypopnea and sleep apnea events. The second problem with analysis based on the classical envelope detection results is related to the filter-induced time-shift effects occurring at the beginning and at the end of each apnea episode. Since the sleep apnea event should last at least 10 s, wrong localization of its endpoints may result in an erroneous event classification.

To eliminate both drawbacks mentioned previously, a cascade made up of a standard median (SM) filter and a recursive median (RM) filter is used instead of the linear lowpass FIR filter $L(\omega)$ in the two envelope detection methods depicted in Figs. 2 and 3. Median filtering is a popular method of noise removal in applications involving signal and image processing. This non-linear technique has proven to be a good alternative to linear filtering as it can effectively suppress impulsive noise while preserving the edge information [4].

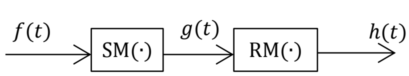

Fig. 4: A cascade of a standard and a recursive median filters.

The output $g(t)$ of the SM filter is the median of the input data inside the window centered at the point t, and is given by

$$g(t) = \mathrm{med}\{f(t-m), \ldots, f(t), \ldots, f(t+m)\} \qquad (11)$$

where $M = 2m + 1$ denotes the window size and $\mathrm{med}\{\cdot\}$ denotes the central value of the ordered sequence of samples. To effectively reduce the influence of artifacts on the airflow envelope, while preserving the sharp envelope "edges" at the beginning and at the end of each apnea episode, the SM window size should be properly selected. If the window size is too small, not all artifacts are suppressed. If the window size is too large, the blurring effect can be observed, similarly as in the case of image processing applications. Therefore the window size should be at least two times larger than the length of the segments affected by artifacts, but smaller than the distance between two neighboring SAH events. Based on the observation that artifacts are usually confined to one breathing cycle, whereas the breathing frequency changes from 0.18 Hz to 0.4 Hz, we suggest that the SM window should cover 15 s of the airflow signal, i.e., that M should be set to 301 under 20 Hz sampling.

The RM filter, used to process the median-prefiltred signal samples $g(t)$, is given by

$$h(t) = \mathrm{med}\{h(t-n), \ldots, h(t-1), g(t), \ldots, g(t+n)\} \quad (12)$$

where $N = 2n + 1$ denotes the window size. The RM filter is more sensitive to window size than the SM filter. If the window is too wide, it can excessively smooth out the signal leading to deformation of the envelope. It is proposed to set N approximately to $M/2$. According to our experiments, the proposed cascade of two median filters yields better results than those obtained when only one of the filters is applied. The proposed modification in the airflow envelope detection procedures allows one to obtain robust envelopes based on the square-law or on the Hilbert transform, further denoted as RESL and REHT, respectively.

Figs. 5 A - 5 B illustrate robustness of the proposed modified envelope detectors in two practically important cases. Fig. 5 A demonstrate insensitivity of RESL and REHT envelopes to breathing artifacts—unlike the classical ESL/EHT envelopes, which are affected by airflow outliers, the RESL/REHT envelopes are robust to short-lived breathing artifacts. This allows one to keep the baseline (which is set to the local envelope maximum) at a level corresponding to regular, i.e., undisturbed breathing. Fig. 5 B demonstrate another advantage of nonlinear filtering—preservation of sharp envelope "edges". When linear lowpass filter is used in lieu of the proposed cascade of nonlinear filters, the corresponding envelopes slowly decay/rise at the beginning/end of each reduced-airflow episode. Since the length of such an episode is an important diagnostic factor, its understatement can result in overlooking or misclasification of SAH events. Median filters do not introduce time shifts mentioned previously. Additionally, the RM filter is very efficient in smoothing out (without blurring envelope edges) some local signal fluctuations that can be observed at the output of an SM filter [4]. As a result, the envelope "valleys" corresponding to SAH events usually have only one local minimum.

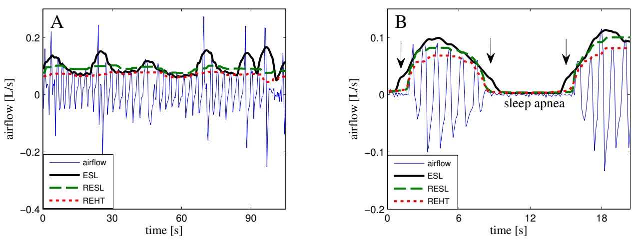

Fig. 5: A - envelope distortions caused by artifacts, B - time-shift effects in the envelope introduced by lowpass filtering.

3. IDENTIFICATION OF SAH EVENTS

Once the airflow envelope has been obtained, identification of SAH events is easy: they can be localized (if present) between the succeeding local maxima of the robust envelope. Each time a new local maximum appears, the baseline value is updated and set to the value of this local maximum. Consider a sequence of samples $\{\widehat{e}(t_1), \ldots, \widehat{e}(t_2)\}$ corresponding to a segment of a robust envelope, where t_1 and t_2 denote two consecutive leftmost local maximum points. The baseline value is set to $\widehat{e}(t_1)$ and two thresholds are computed corresponding to 50% and 10% of the baseline value for hypopnea and apnea events, respectively. The following decisions are made at each time instant t, $t_1 < t < t_2$

$$d_{\mathrm{h}}(t) = \begin{cases} 1 & \text{if} \quad \widehat{e}(t) \le \frac{1}{2}\widehat{e}(t_1) \\ 0 & \quad\text{otherwise} \end{cases}$$

$$d_{\mathrm{a}}(t) = \begin{cases} 1 & \text{if} \quad \widehat{e}(t) \le \frac{1}{10}\widehat{e}(t_1) \\ 0 & \quad\text{otherwise} \end{cases}$$

where $d_{\mathrm{h}}(t)$ and $d_{\mathrm{a}}(t)$ denote sequences of binary values indicating which samples in the analysed segment can be classified as hypopneic/apneic activity. If the segment includes less than 10 s of continuous hypopneic/apneic activity, it is classified as normal breathing. Otherwise hypopnea or apnea is detected. When both types overlap, only the sleep apnea episode is scored (see Fig. 6).

4. EXPERIMENTAL RESULTS

Overall, the polysomnograms of 30 sleep apnea patients were used to validate the proposed method. The detailed results are presented for a group of 15 representative patients, 9 male and 6 female [age: 53 ± 7 years (mean$\pm$standard deviation), duration of each study: 449 minutes] with different apnea/hypopnea indexes, ranging from 4.5 (patient #1) to 42.0 (patient #15); because of space limitations, only summary/average statistics are shown for all patients. The analyzed sleep studies were drawn from the database of the Medical University of Gdańsk, Gdańsk, Poland (only recordings containing breathing artifacts were selected). In all studies the airflow signal, bandpass filtered and sampled at the rate of 20 Hz, was measured using a nasal cannula pressure transducer. Respiratory events were detected based

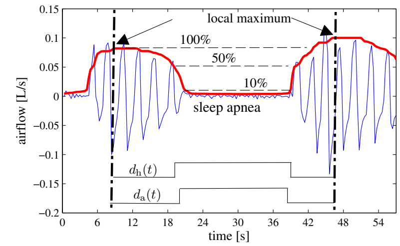

Fig. 6: Identification of SAH events based on the RESL.

on analysis of the airflow signal and scored, using the criteria proposed by AASM [1], by an expert—a pulmonologist with 15 years clinical experience (all recordings were scored by the same person). In agreement with [1], hypopnea was defined as an over 50% reduction in airflow from the baseline value, lasting for more than 10 s, and associated with at least 3% desaturation or an arousal. Sleep apnea was defined as the absence of airflow for more than 10 s. The clinical routine was based on manual correction of the results obtained by automated analysis performed by a commercial PSG software (RemLogic, default settings).

The common mistakes of the automated analysis are: overlooked episodes, false detections, and misclassification between hypopnea and sleep apnea events. In all studies 4559 SAH events were detected: 2193 apneas and 2366 hypopneas. The total number of artifacts present in airflow recording—the result of subjective scoring of abnormally large peaks—was 2353, ranging from 9 to 230 per recording.

The 127-tap FIR filter, approximating the Hilbert transform, was designed using the Parks-McClellan algorithm. The analytic signal was computed by adding the appropriately time-shifted real signal to its imaginary counterpart generated by the Hilbert filter. To reduce computational complexity, prior to median filtering the signal $f(t)$ was passed through a lowpass FIR filter with a cutoff frequency of 1 Hz, and then downsampled by a factor of $d = 6$. After decimation, the SM window size was set to $M = 51$ and the RM window size was set to $N = 21$.

Table I compares the results of automated detection of SAH events, obtained using Approach I based on the square-law, Approach II based on the Hilbert transform, and RemLogic, with decisions made by an expert. Patients were ordered according to their apnea/hypopne index (AHI), which reflects the number of sleep apnea/hypopnea events per one hour of sleep. The scores correspond to the number of detected apnea/hypopnea events (SAH), only apnea events (A), and only hypopnea events (H). It is easy to notice that, unlike the proposed approaches, RemLogic shows tendency to understate the AHI score.

Table II shows results of the event-by-event and of the epoch-by-epoch analysis. In this experiment no differentiation between apnea and hypopnea event was made. Four performance measures (accuracy, sensitivity, specificity, and Cohen's coefficient of agreement) were evaluated based on analysis of 30-s airflow epochs classified as SAH events or

Table 1: Quantitative comparison of scores provided by an expert with the results yielded by RemLogic and by two variants of the proposed method: Approach I based on the square-law, and Approach II based on the Hilbert transform.

		Expert				Approach I				Approach II				RemLogic			
Patient	Art	SAH	A	H	AHI	SAH	A	H	AHI	SAH	A	H	AHI	SAH	A	H	AHI
1	20	33	5	28	4.5	36	3	33	4.9	34	2	32	**4.6**	17	1	16	2.3
2	70	36	2	34	4.8	45	2	43	**6.0**	49	2	47	6.5	24	1	23	3.2
3	13	46	13	33	6.1	45	9	36	**6.0**	53	11	42	7.1	27	8	19	3.6
4	70	50	5	45	6.7	53	1	52	**7.1**	58	1	57	7.7	38	1	37	5.1
5	78	106	17	89	14.1	127	8	119	**16.9**	130	8	122	17.3	32	3	29	4.3
6	17	111	28	83	14.8	107	17	90	14.3	113	16	97	**15.1**	97	33	64	12.9
7	59	126	75	51	16.8	136	67	69	**18.1**	142	70	72	18.9	89	67	22	11.9
8	36	167	126	41	22.3	147	100	47	**19.6***	147	100	47	**19.6***	145	132	13	19.3
9	76	169	19	150	22.5	168	13	155	**22.4**	178	15	163	23.7	109	9	100	14.5
10	64	216	133	83	28.8	222	23	199	**29.6**	201	20	181	26.8	192	79	113	25.6
11	55	217	170	47	28.9	229	167	62	**30.5**	231	165	66	30.8	188	163	25	25.1
12	210	241	74	167	32.1	247	22	225	32.9	241	27	214	**32.1**	172	37	135	22.9
13	69	290	196	94	38.7	295	179	116	**39.3***	295	184	111	**39.3***	231	162	69	30.8
14	206	311	247	64	41.5	312	197	115	**41.6**	330	203	127	44.0	204	104	100	27.2
15	91	315	177	138	42.0	314	154	160	**41.9**	301	154	147	40.1	247	144	103	32.9
Σ_{15}	1134	2434	1287	1147		2483	962	1521		2503	978	1525		1812	944	868	
Σ_{30}	2353	4559	2193	2366		4753	1841	2912		4807	1873	2934		3261	1711	1550	

The scores correspond to the number of detected apnea/hypopnea events (SAH), only apnea events (A), and only hypopnea events (H). The AHI index reflects the number of apnea/hypopnea events per one hour of sleep. The AHI scores that are closer to expert scores than those yielded by the competing approaches are shown in boldface. Equal scores are marked with asterisks. The number of artifacts present in airflow recordings (Art), based on subjective scoring of abnormally large peaks, ranged from 13 to 210 per recording. Σ_{15}: sum for a representative group of 15 patients, Σ_{30}: sum for all patients.

normal breathing. The epoch was classified as SAH event if at least 5 s of the epoch was affected by apneic/hypopneic activity. The total number of epochs for 30 patients was 26970. The obtained results in both experiments clearly indicate superiority of the square-law-based approach.

The proposed approach can be easily implemented in "screener" devices such as the ones described in [5, 6]. In order to increase the reliability of the proposed algorithm and decrease the number of false detection rate, a combined analysis of airflow signal and saturation of peripheral oxygen signal is proposed. Some further work is also needed to develop tools capable of distinguishing both types of events in a more reliable way.

5. CONCLUSIONS

The widely used classical envelope detection methods are not robust to artifacts present in airflow measurements. The proposed approach is a simple modification of the existing schemes, obtained by replacing the linear lowpass output filter with a cascade of two nonlinear filters—a standard median filter and a recursive median filter. The proposed approach gives better results than those yielded by RemLogic even though, unlike RemLogic, our method analyzes only the airflow signal.

Table 2: Comparison of the scores obtained for the commercial system RemLogic and for two variants of the proposed method—Approach I based on the square-law, and Approach II based on the Hilbert transform—against the "golden standard" expert scores.

	Event-by-Event			Epoch-by-Epoch			
Approach	TP_{SAH} (%)	FP (%)	FN (%)	Acc (%)	Sens (%)	Spec (%)	κ
RemLogic	67.4	**8.0**	32.6	90.3	59.0	**97.0**	0.61
Approach I	**94.0**	10.6	**6.0**	**95.1**	**90.3**	95.8	**0.82**
Approach II	91.2	14.6	8.8	93.7	87.6	94.6	0.77

The best scores are shown in boldface. TP_{SAH}: true positive apnea/hypopnea, FP: false positive, FN: false negative, Acc: accuracy, Sens: sensitivity, Spec: specificity, and agreement κ ($\kappa = 1$ indicates complete agreement).

REFERENCES

[1] R.B. Berry, et al., AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications, www.aasmnet.org, 2012.

[2] A. Otero, et al., "Algorithms for the analysis of polysomnographic recordings with customizable criteria," *Expert Systems with Applications*, vol. 38, pp. 10133–10146, 2011.

[3] S. Haykin, *Communication Systems*, Wiley 2009.

[4] I. Pitas et al., *Nonlinear Digital Filters: Principles and Applications.*, Springer-Verlag US, 1990.

[5] T. Shochat, et al., "An apnoea screener for the early detection of sleep apnoea syndrome, Eur. Respir. J., vol. 19, pp. 121-126, 2003.

[6] M. K. Erman,et al., "Validation of the ApneaLinkT M for the screening of sleep apnea: A novel and simple single-channel recording device, J. Clin. Sleep Med., vol. 3, no. 4, pp. 387392, 2007.