Automated Detection of Sleep Apnea and Hypopnea Events Based on Robust Airflow Envelope Tracking in the Presence of Breathing Artifacts

Marcin Ciołek, Maciej Niedźwiecki, *Senior Member, IEEE*, Stefan Sieklicki, Jacek Drozdowski, and Janusz Siebert

***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, based on a robust envelope detector, 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 proposed approach, evaluated on 30 overnight polysomnographic recordings, was assessed using diagnostic measures such as accuracy, sensitivity, specificity, and Cohen's coefficient of agreement; the achieved levels were equal to 95%, 90%, 96%, and 0.82, respectively. The results suggest that the algorithm may be implemented successfully in portable monitoring devices, as well as in software-packages used in sleep laboratories for automated evaluation of sleep apnea/hypopnea syndrome.**

***Index Terms*—Breathing artifacts, envelope detector, median filters, sleep apnea and hypopnea (SAHS).**

I. Introduction

SLEEP apnea/hypopnea syndrome (SAHS) is a sleep-breathing disorder characterized by repetitive episodes of complete obstruction (sleep apnea event) or partial obstruction (hypopnea event) of the upper airway, resulting in a blood oxygen desaturation or arousals leading to sleep fragmentation. The usual daytime manifestation is excessive sleepiness, fatigue, and poor concentration, which can escalate to traffic accidents, depression, and memory loss. The major risk factors for the disorder include obesity, male gender, and age [1]. Untreated SAHS may lead to cardiovascular dysfunction, stroke, and possibly (since this supposition is still not well documented) to the ischemic heart disease [2]. SAHS is a noticeable problem of social and health life, affecting 3% of children [3], 2% of female adults, and 4% of male adults worldwide [4]. In fact, still up to 80% of cases of moderate or severe SAHS have gone undiagnosed despite adequate access to health care [5], [6].

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. PSG study is considered as the "gold standard" method for SAHS diagnosis [7]. It involves recording and studying simultaneously many signals such as electrocardiogram (ECG), nasal airflow (NAF), and blood oxygen saturation (SaO_2).

To reach the final conclusion, the recorded signals are analyzed 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 $5 \leq \text{AHI} < 15$; moderate, when $15 \leq \text{AHI} < 30$; and severe, when $\text{AHI} \geq 30$ (events per hour of sleep) [7], [8]. Currently, the clinical routine is based on manual correction of the results obtained by automated analysis, which is an extremely tedious and time-consuming task [8]. Unfortunately, the drawbacks of the full PSG study, such as high cost, a long list of patients waiting to be tested, and a feeling of discomfort due to a large number of sensors placed on the patient's body during the overnight test, suggest the need for developing alternative methods of diagnosis, based on information from selected channels of PSG, which could be implemented in portable monitoring devices for evaluation of SAHS [1], [9], [10].

Numerous methods exist, based on the evaluation of various signals, for detection of SAH events. This includes methods based on analysis of airflow signals [11]–[15], ECG signals [16]–[19], pulse oximetry (SaO_2) signals [20], [21], tracheal sound signals [22], or some combinations of signals listed previously [23]. A number of portable devices for sleep apnea monitoring and diagnosis are available, such as ApneaLink, SleepStrip, and LifeShirt, to name only the most popular solutions (see [24] and references therein). All signals mentioned previously provide only supportive evidence of SAH events and do not allow one to localize precisely their beginning and end. It often happens that the primary evidence (significant reduction in the NAF signal) is not observed. In the case of an SaO_2 signal, the supportive evidence, in the form of a blood oxygen desaturation, is delayed in time in relation to the moments of occurrence of SAH events. Additionally, it is highly

Manuscript received December 5, 2013; revised April 29, 2014; accepted May 12, 2014. Date of publication May 23, 2014; date of current version March 2, 2015. This work was supported in part by the National Science Center under Agreement 2927/B/T02/2011/40.

M. Ciołek, M. Niedźwiecki, and S. Sieklicki are with the Department of Automatic Control, Faculty of Electronics, Telecommunications and Computer Science, Gdańsk University of Technology, 80-233 Gdańsk, Poland (e-mail: marcin.ciolek@pg.gda.pl; maciekn@eti.pg.gda.pl; sieklick@eti.pg.gda.pl).

J. Drozdowski is with the Department of Pneumology and Allergology, Medical University of Gdańsk, 80-210 Gdańsk, Poland (e-mail: jacekdr@amg.gda.pl).

J. Siebert is with the Department of Family Medicine, Medical University of Gdańsk, 80-210 Gdańsk, Poland, and also with the Gdańsk University of Technology, 80-233 Gdańsk, Poland (e-mail: jsiebert@gumed.edu.pl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JBHI.2014.2325997

dependent on many factors, such as calibration of the measuring device, physiological conditions of the patient, presence of artifacts, etc. [8]. In the recent years, several studies have focused on detection of SAH events based exclusively on the analysis of the airflow signal. Almost all of these studies are based on a "black box" approach and use such techniques as artificial neural networks [11], spectral analysis [12], feature selection, [13] or support vector machines [14], [15].

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.

II. Breathing Artifacts

The standard morphological criteria, given by the American Academy of Sleep Medicine (AASM) [1], describe SAH events as a significant reduction of the 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 the 1999 AASM "Chicago consensus paper" [1], 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)." This specification was upheld in two later AASM recommendations presented in 2007 by Iber *et al.* [9], and in 2012 by Berry *et al.* [10]. 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 NAF 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 [8], [25]. In [26], the authors propose to track the baseline value in adaptive way, based on an exponentially weighted average of past peaks. To eliminate the impact of large positive and negative peaks, only peaks that remain within 40% of the current baseline value are considered. This approach seems to work well until a sudden change of breathing rhythm appears, leading to problems with fast updating of the baseline value. The approach proposed in [27] is based on airflow signal modeling. The model is used to reconstruct fragments corrupted by artifacts.

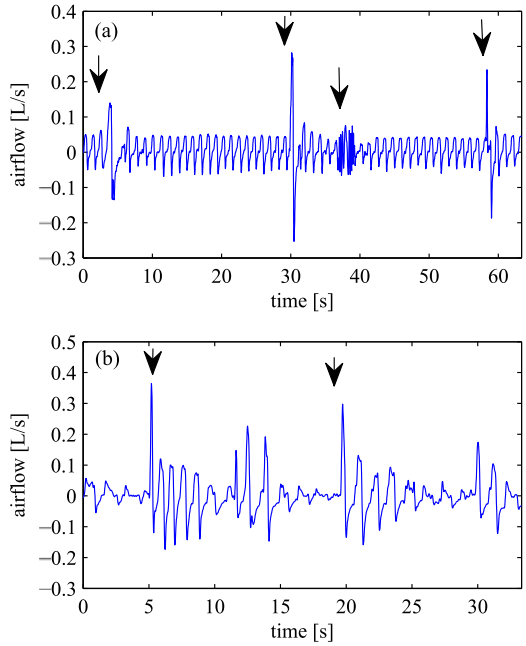

Fig. 1. Breathing artifacts: (a) rapid body movements and changing of the sleep position and (b) sudden opening of the upper airway succeeding a sleep apnea/hypopnea event. Artifacts are marked with arrows.

III. 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 the 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 [28], 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 $t = \ldots, -1, 0, 1, \ldots$ denotes normalized (dimensionless) discrete time, $m(t) \geq 0$ denotes the baseband message, ω_c denotes the 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 low-pass signal with bandwidth W much smaller than the carrier frequency ω_c, the amplitude

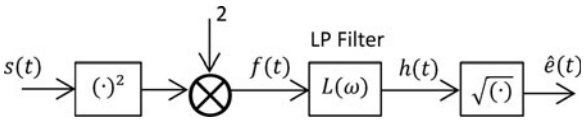

Fig. 2. Square-law envelope detector.

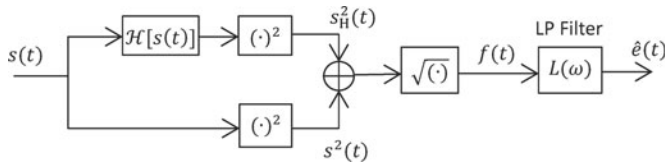

Fig. 3. Envelope detector based on the Hilbert transform.

envelope of the signal $s(t)$ is defined as

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

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 the 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.

A. Square-Law Envelope Detector

The flowchart of the square-law detector [28] 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

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

The first term on the right-hand side of (3) is a low-pass signal with a cutoff frequency $2W$, and the second one is a band-pass signal with spectrum confined to the frequency bands $(-2\omega_c - 2W, -2\omega_c + 2W)$ 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 low-pass component of $f(t)$ can be extracted using a low-pass FIR filter $L(\omega)$ with a cutoff frequency $2W$

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

where $l_i = l_{-i}, i = 0, \ldots, k$, denote impulse response coefficients of the filter. The estimated value of the envelope can be obtained from

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

further referred to as an ESL envelope.

B. Envelope Detector Based on the Hilbert Transform

The second classical method of envelope detection, based on the Hilbert transform [28], 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 [28]

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

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

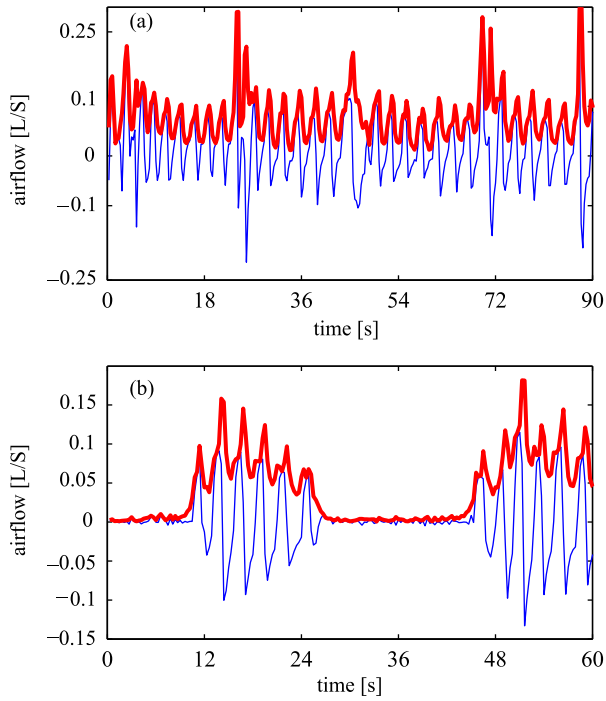

Fig. 4. Output of the Hilbert-transform-based envelope detector *prior* to low-pass filtering (thick line): (a) normal breathing in the presence of incidental artifacts and (b) sleep apnea patterns. Note the presence of high-frequency fluctuations called ripples. Thin lines show the airflow (input) signal.

defined as

$$y(t_c) = s(t_c) + j s_H(t_c). \quad (7)$$

The analytic counterpart of a discrete-time signal $s(t)$ can be evaluated either directly—using the FFT-based frequency-domain approach [29], or indirectly—by computing an "analytic" signal $s_H(t)$ (using the discrete-time FIR approximation of the Hilbert transform) and combining it with $s(t)$

$$y(t) = s(t) + j s_H(t). \quad (8)$$

Since the Hilbert transform shifts the phase of all sinusoidal components by $-\pi/2$, for the "ideal" AM signal (1), one obtains $s_H(t) \cong A[1 + \beta m(t)] \sin \omega_c t$ which means that the envelope of $s(t)$ can be obtained by evaluating the magnitude of the analytic signal

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

For AM-like signals, such as speech signals or biomedical signals, the envelope extracted in this way suffers from high-frequency fluctuations, called ripples [30] (see Fig. 4). Ripples

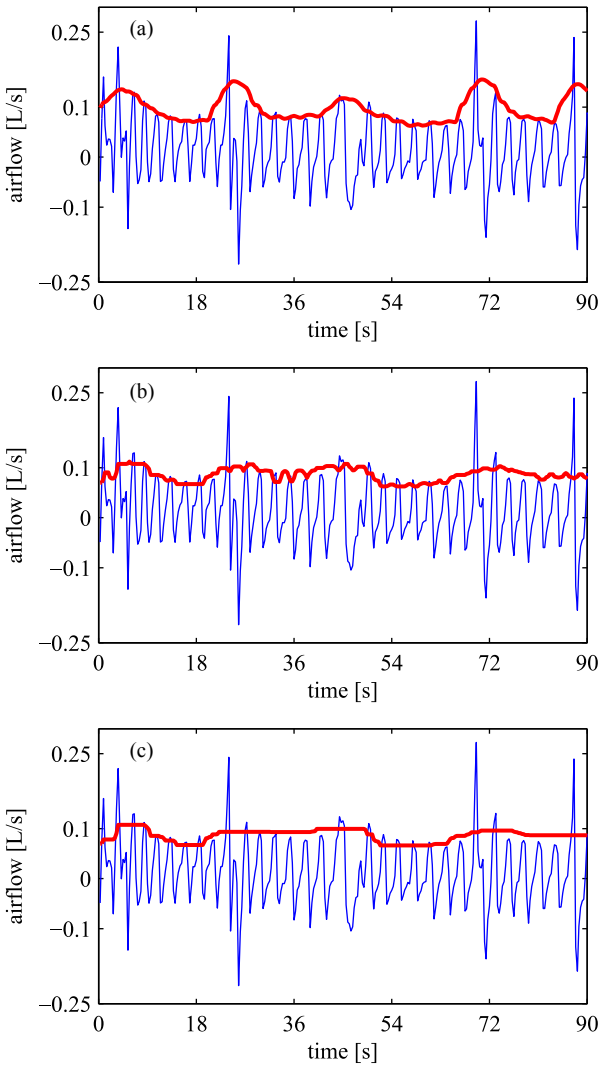

Fig. 5. Performance of (a) the classical square-law envelope detector and (c) its robust version in the presence of breathing artifacts. (b) Middle figure shows the intermediate detection results observed at the output of the SM filter. Thick line—envelope, thin line—airflow signal.

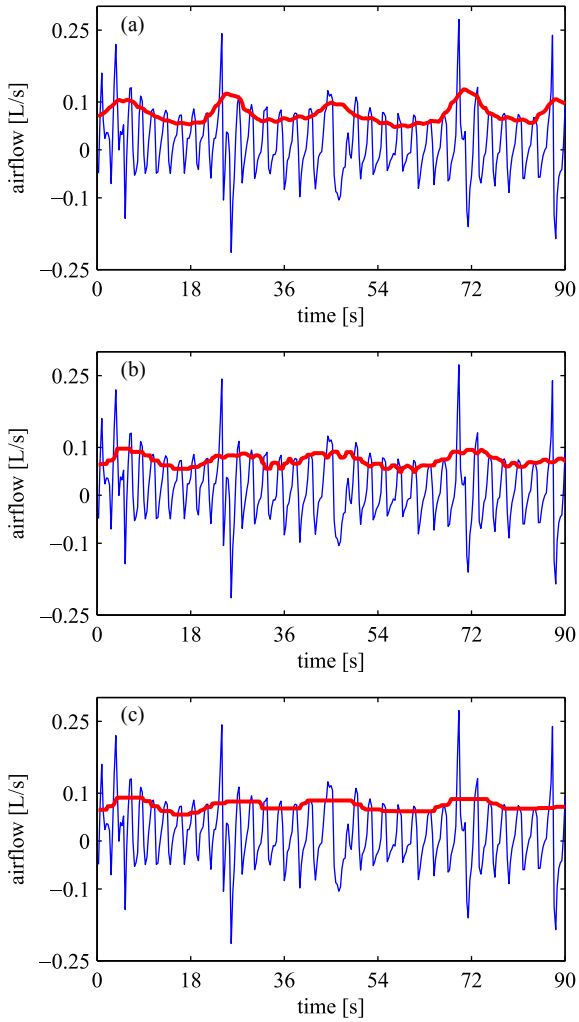

Fig. 6. Performance of (a) the classical Hilbert transform envelope detector and (c) its robust version in the presence of breathing artifacts. (c) Middle figure shows the intermediate detection results observed at the output of the SM filter. Thick line—envelope, thin line—airflow signal.

can be removed by passing the signal $f(t)$ through a low-pass filter, leading to

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

further referred to as an EHT envelope.

C. 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 low-pass 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 low-pass 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 nonlinear technique has proven to be a good alternative to linear filtering as it can effectively suppress impulsive noise while preserving the edge information [31], [32].

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

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

where $M = 2m + 1$ denotes the window size and $\text{med}\{\cdot\}$ denotes the central value of the ordered sequence of samples. To

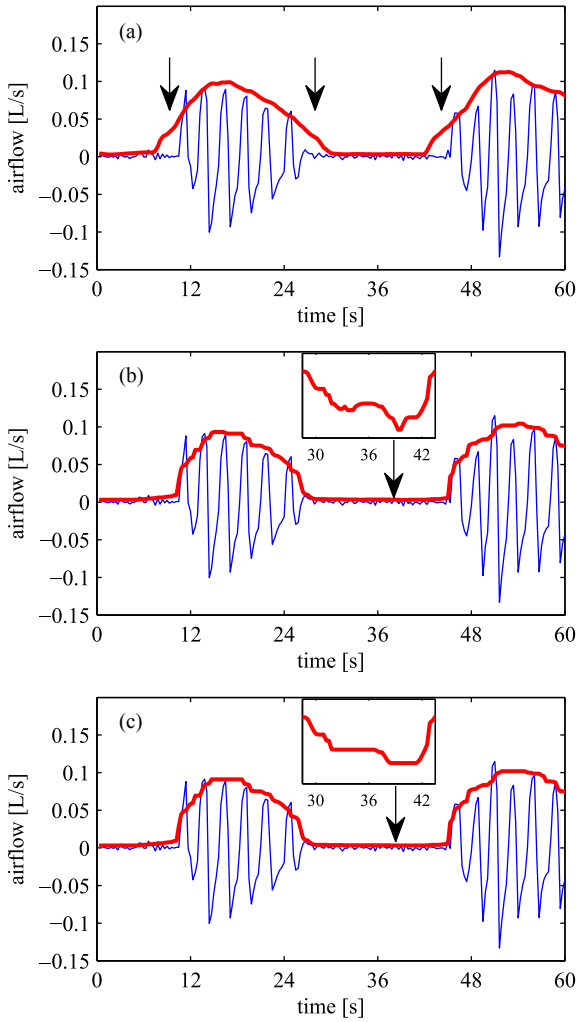

Fig. 7. Performance of (a) the classical square-law envelope detector and (c) its robust version in the presence of sleep apnea/hypopnea events. (b) Middle figure shows the intermediate detection results observed at the output of the SM filter. Thick line—envelope, thin line—airflow signal.

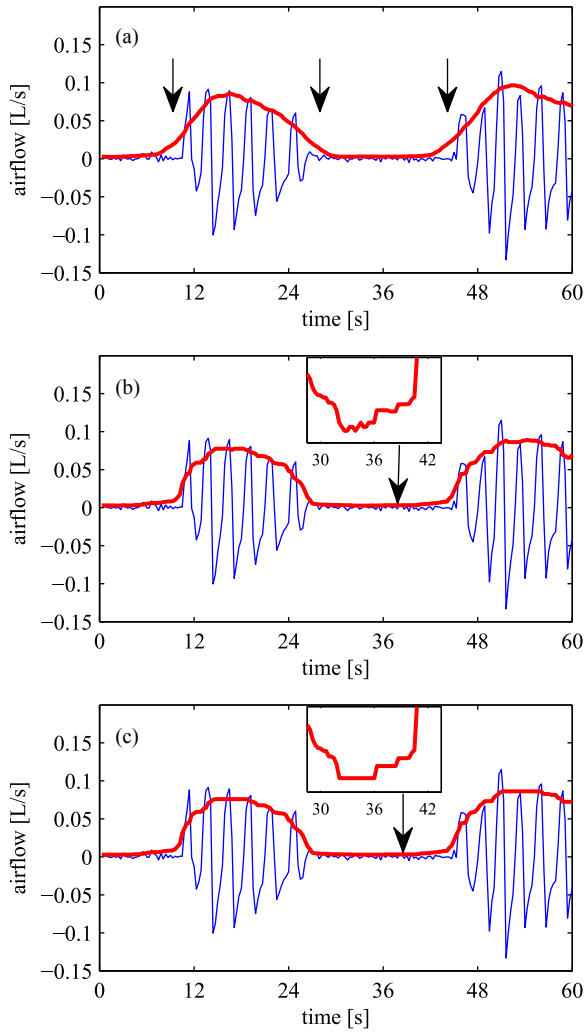

Fig. 8. Performance of (a) the classical Hilbert transform envelope detector and (c) its robust version in the presence of sleep apnea/hypopnea events. (b) Middle figure shows the intermediate detection results observed at the output of the SM filter. Thick line—envelope, thin line—airflow signal.

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 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-prefiltered signal samples $g(t)$, is given by

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

where $N = 2n + 1$ denotes the window size. The RM filter is more sensitive to the 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. 5–8 illustrate robustness of the proposed modified envelope detectors in two practically important cases.

Figs. 5 and 6 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.

Figs. 7 and 8 demonstrate another advantage of nonlinear filtering—preservation of sharp envelope "edges." When the linear low-pass 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 misclassification 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 [33]. As a result, the envelope "valleys" corresponding to SAH events usually have only one local minimum. This very much simplifies SAH identification as SAH episodes can be easily localized between the two successive local maxima.

Remark 1: Note that median windows centered at instants t and $t+1$ partially overlap. Therefore, since evaluation of $g(t)$ given by (11) has already sorted most of the samples that are required for evaluation of $g(t+1)$, the computation can be made much more efficient. Using the indexable skip list technique [34], the computational complexity of a median filter can be reduced from $O(M^2)$ comparisons to only $O(\log M)$ comparisons per time update [35]. The same technique can be used for realization of a fast RM filter.

Remark 2: Similarly as in the case of classical envelope detectors, the computational load of the proposed methods can be further reduced by downsampling the signal $f(t)$ prior to nonlinear filtration. To avoid aliasing effects, prior to downsampling, the signal $f(t)$ should be passed through a linear low-pass filter with an appropriately chosen cutoff frequency.

IV. Identification of SAH Events

We will call t the leftmost local maximum point of a discrete-time signal $x(\cdot)$ if for some $i \geq 1$, it holds that

$$\begin{aligned} &x(t) > x(t-1) \\ &x(t) = \ldots = x(t+i-1) > x(t+i). \end{aligned} \tag{13}$$

Likewise, t will be called the leftmost local minimum point of $x(\cdot)$ if for some $i \geq 1$, it holds that

$$\begin{aligned} &x(t) < x(t-1) \\ &x(t) = \ldots = x(t+i-1) < x(t+i). \end{aligned} \tag{14}$$

Note that while for piecewise-constant signals, such as those produced by the proposed nonlinear filter, the local signal maximum/minimum can be attained at more than one point, the leftmost local maximum/minimum points are always unique. Localization of the leftmost extremum points is straightforward and can be done by means of checking the conditions (13) and (14) in a sequential manner. Note also that between two consecutive leftmost local maximum points t_1 and t_2, $t_1 < t_2$, there exists one and only one leftmost local minimum point t_0, $t_1 < t_0 < t_2$.

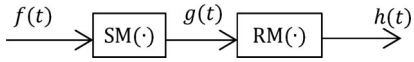

Fig. 9. Cascade made up of SM and RM filters.

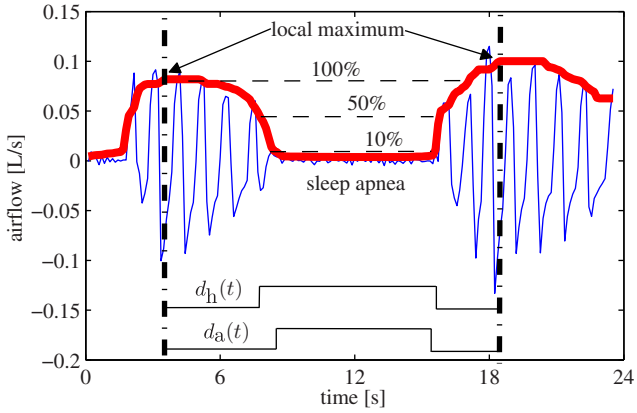

Fig. 10. Identification of SAH events based on the RESL envelope.

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 is found, 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, $t_1 < t < t_2$

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

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

where $d_{\mathrm{h}}(t)$ and $d_{\mathrm{a}}(t)$ denote binary pulses (sequences of binary values) indicating which samples in the analyzed 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. 10).

Remark: Since the signal observed at the output of the SM filter (the first component of our two-component nonlinear filter) usually exhibits many local maxima and minima, which can be attributed to local noise patterns rather than to the generic envelope properties [see Figs. 7(b) and 8(b)], it cannot be analyzed using the simple procedure described previously. The advantage of the RM filter, used at the second filtering stage, is that it produces smooth and easily interpretable signals that are free of such low-magnitude short-time fluctuations. Owing to this property of the robust envelope, the proposed analytic procedure yields results that stay in a good agreement with the

results of visual inspection of the airflow signal performed by an expert.

V. 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, nine male and six female [age: 53 ± 7 years (mean$\pm$standard deviation), duration of each study: 449 min] 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 (without square root transformation). Respiratory events were detected based on analysis of the airflow signal and scored, using the criteria proposed by AASM [1], [9], [10], 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 the 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 the 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). Under such conditions, the interscorer and intrascorer reliability is known to be very good [8].

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 recordings—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 low-pass 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 $N = 21$.

Tables I–III present qualitative and quantitative comparison of the competing methods.

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 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). The number of artifacts present in airflow recordings, based on subjective scoring of abnormally large peaks, ranged from 13 to 210 per recording. 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 analysis. The best results were obtained for Approach I (square-law-based method).

Table III shows results of the epoch-by-epoch examination. Evaluation was based on analysis of 30-s airflow epochs classified as positive, if at least 5 s of the epoch was affected by hypopneic/apneic activity. For each patient, the number of epochs was equal to 899. Four quality measures were used to assess the performance of the detectors: accuracy, sensitivity, specificity, and Cohen's coefficient of agreement (agreement beyond that expected by chance, usually referred to as kappa statistic). Again, the obtained results clearly indicate superiority of the square-law-based approach. Note that while the specificity index is usually higher for RemLogic than for the proposed approaches (due to a smaller number of false-positive decisions), its sensitivity index is pretty low (due to a much larger number of false negative decisions). The same tendency can be seen in Table II. In the case of Approaches I and II, both specificity and sensitivity reach satisfactory levels. Note also that the Cohen's coefficient of agreement κ evaluated for the robust square-law detector takes pretty large values—for 12 patients (out of 15) it holds that $\kappa \in [0.81, 1]$, which corresponds to the highest qualitative level of agreement strength, interpreted as "almost perfect agreement" [36] or "very good agreement" [37], and for the remaining three patients $\kappa \in [0.61, 0.80]$, which is regarded as "substantial agreement" [36] or "good agreement" [37].

Bland–Altman plots, comparing AHI scores provided by an expert with those resulting from the automated analysis, are shown in Fig. 11. In this popular graphical method, the differences between the two scoring techniques are plotted against their averages. Horizontal lines are drawn at the mean difference, and at the 95% limits of agreement, which are defined as the mean difference ± 1.96 times the standard deviation (SD) of the differences. Bland–Altman plots allow one to investigate the existence of any systematic difference (fixed bias) between the scores, and to identify possible outliers. According to the plots shown in Fig. 11, the proposed methods are nearly unbiased—mean biases are equal to -0.87 (SD $= 1.17$) and -1.11(SD $= 1.91$) event/hour of sleep for Approaches I and II, respectively. The differences remain close to the bias line, even for an increasing AHI index. For RemLogic, the bias is nonnegligible and equal to 4.78(SD $= 4.04$), which means that a risk that a practitioner could make the wrong decision and fail to refer a patient for more extensive testing is in this case higher.

Finally, it may be interesting to compare efficiency of Approach I (which is the recommended one) with that provided by other single-channel (airflow only) solutions, such as the SleepStrip and ApneaLink screeners mentioned in Section I. As reported in [38] (based on examination of 402 subjects), for a SleepStrip device, the sensitivity and specificity values ranged from 80% to 86% and from 57% to 86%, respectively. According to [39], the average sensitivity and average specificity obtained for an ApneaLink device (63 subjects) were equal to 76% and

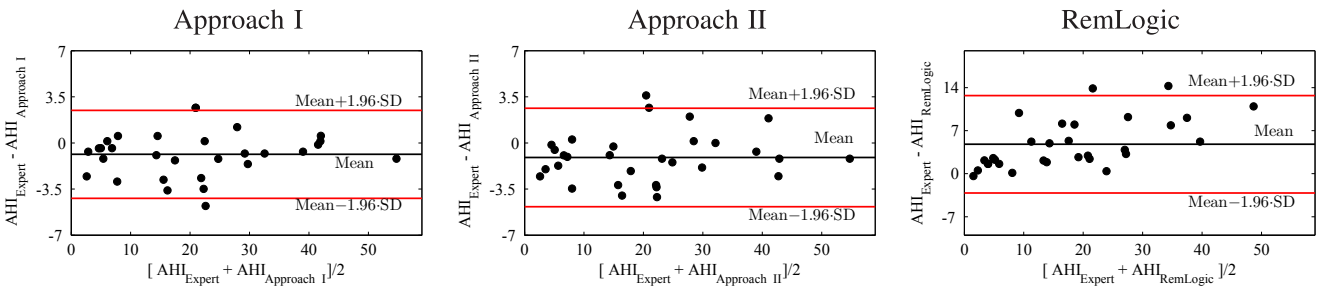

Fig. 11. Bland–Altman plots comparing AHI scores (for 30 patients) provided by an expert with those resulting from the automated analysis using Approaches I and II, and RemLogic (differences between the scores against their averages). Horizontal lines show the mean difference and the 95% limits of agreement (SD = standard deviation). Note the difference in the vertical scale between the plots for Approaches I/II and the plot for RemLogic.

TABLE I
QUANTITATIVE COMPARISON OF SCORES PROVIDED BY AN EXPERT WITH THE RESULTS YIELDED BY THE COMMERCIAL SYSTEM 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.

94%, respectively (for AHI$\geq$ 15). In the case of Approach I (30 subjects), sensitivity and specificity range from 77% to 100% and from 87% to 99%, respectively, and the corresponding average scores are equal to 90% (sensitivity) and 96% (specificity). Hence, even though more subjects should be examined to make such comparison statistically meaningful, the results obtained for Approach I look very promising. The same conclusion can be reached after comparing the Pearson correlation coefficients measuring similarity between the estimated AHI scores and the reference (PSG-based) AHI scores—the corresponding values are equal to 0.73 (SleepStrip), 0.89 (ApneaLink), and 0.99 (Approach I).

VI. DISCUSSION

Unlike "black box" approaches to detection of SAH events [11], [13]-[15], the proposed method refers directly to explicit SAH detection/classification rules formulated and approved by pulmonologists, where the concept of breathing baseline plays a crucial role. Since, during the verification stage, the robust envelope along with the corresponding baseline values and their 50/10 percentage levels can be superimposed on the analyzed signal (see Fig. 10), in each case an expert learns not only what decision was made by the automated scoring system, but also *why* such a decision was taken. Such a "visual feedback" is particularly useful in the presence of breathing artifacts. It makes verification both easier and faster.

Another important aspect of the proposed approach is its reproducibility. When the supervised learning methods, such as artificial neural networks [11] are used, the final form of the decision device depends on the training data, making the research results difficult to reproduce (unless the training dataset is made available). Additionally, many of the "black box" methods require setting a large number of design parameters, such as decision thresholds, adaptation gains, etc. In contrast with this, the proposed approach is explicit and easy to implement. In particular, note that since the concrete values of the analysis

TABLE II
EVENT-BY-EVENT ANALYSIS

	Approach I					Approach II					RemLogic				
Patient	TP_{SAH}	TP_A	TP_H	FP	FN	TP_{SAH}	TP_A	TP_H	FP	FN	TP_{SAH}	TP_A	TP_H	FP	FN
1	**28**	**2***	**23**	8	**5**	25	**2***	21	9	8	16	1	12	**1**	17
2	**35***	**1***	**32***	10	**1***	**35***	**1***	**32***	14	**1***	22	**1***	21	**2**	14
3	**43**	9	**31**	**2**	**3**	42	**10**	29	11	4	23	8	11	4	23
4	**44**	**1***	**39**	**9**	**6**	39	**1***	34	19	11	21	**1***	20	17	29
5	**99**	**4***	**79**	27	**7**	95	**4***	75	34	11	28	3	17	**4**	78
6	**99**	16	**72**	**8**	**12**	98	15	71	16	13	82	**25**	49	12	29
7	**124**	67	**50**	13	**2**	122	**68**	46	21	4	84	65	12	**5**	42
8	**153**	110	**39**	4	**14**	151	110	37	6	16	132	**121**	9	**2**	35
9	**163**	13	**144**	**5**	**6**	158	**15**	139	20	11	99	9	80	10	70
10	**198**	23	**75**	24	**18**	179	20	53	**22**	37	168	**78**	42	23	48
11	**216**	**162**	**44**	12	**1**	212	161	41	18	5	186	158	25	**2**	31
12	**230**	21	**156**	17	**11**	216	25	142	25	25	158	**32**	93	**10**	83
13	**272**	**174***	**79**	25	**18**	266	**174***	71	32	24	214	155	32	**11**	76
14	298	194	**58**	15	13	**299**	**199**	54	32	**12**	199	102	16	**5**	112
15	**301**	**154**	**128**	**16**	**14**	286	149	113	18	29	220	130	58	25	95
Σ_{15}	**2303**	951	**1049**	195	**131**	2223	**954**	958	297	211	1652	889	497	**133**	782
Σ_{30}	**4283**	**1777**	**2104**	484	**276**	4160	1776	1954	664	399	3073	1600	1070	**366**	1486

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. The scores that are better than or equal to those yielded by the competing approaches are shown in boldface. Equal scores are marked with asterisks. TP_{SAH}: true positive apnea/hypopnea, TP_A: true positive apnea, TP_H: true positive hypopnea, FP: false positive, FN: false negative, Σ_{15}: sum for a representative group of 15 patients, Σ_{30}: sum for all patients. The number of misclassified events (hypopnea classified as apnea or *vice versa*) can be calculated as $M = TP_{SAH} - TP_A - TP_H$.

TABLE III
EPOCH-BY-EPOCH ANALYSIS

	Approach I				Approach II				RemLogic			
Patient	Acc (%)	Sens (%)	Spec (%)	κ	Acc (%)	Sens (%)	Spec (%)	κ	Acc (%)	Sens (%)	Spec (%)	κ
1	**97.78**	**80.49**	98.60	**0.76**	97.22	75.61	98.25	0.70	96.96	43.90	**99.53**	0.56
2	**97.55**	**94.23***	97.76	**0.80**	96.66	**94.23***	96.81	0.75	96.55	48.08	**99.53**	0.60
3	**98.67**	**91.07**	99.17	**0.89**	97.22	85.71	97.98	0.78	96.44	50.00	**99.53**	0.62
4	**97.55**	**87.72**	**98.22**	**0.81**	95.88	78.95	97.03	0.69	93.77	43.86	97.15	0.44
5	**93.88**	**87.74**	95.16	**0.79**	92.10	84.52	93.68	0.74	86.43	28.39	**98.52**	0.36
6	**95.88**	**84.67***	**98.13**	**0.85**	94.55	**84.67***	96.53	0.81	92.55	69.33	97.20	0.71
7	**94.66**	**95.11**	94.55	**0.85**	93.66	94.02	93.57	0.82	92.32	68.48	**98.46**	0.74
8	**96.66**	**92.64**	98.28	**0.92**	96.11	91.47	97.97	0.90	93.55	79.07	**99.38**	0.83
9	**97.22**	**91.41**	**98.86**	**0.92**	95.11	89.90	96.58	0.86	88.77	54.04	98.57	0.62
10	**94.80**	**92.02**	95.77	**0.86**	92.99	84.51	95.63	0.81	92.44	77.93	**96.94**	0.78
11	**95.77**	**94.75**	96.40	**0.91**	94.33	93.29	94.96	0.88	93.33	85.13	**98.38**	0.86
12	**93.88**	**93.67**	94.00	**0.87**	90.99	88.92	92.11	0.80	84.54	65.19	**95.03**	0.64
13	**92.99**	**92.53**	93.32	**0.86**	91.77	90.67	92.56	0.83	87.21	74.13	**96.56**	0.73
14	**96.33**	**96.28**	96.35	**0.92**	93.88	94.74	93.40	0.87	87.10	65.63	**99.13**	0.70
15	**91.32**	**91.85**	90.76	**0.83**	88.88	87.77	90.07	0.78	80.20	69.31	**91.92**	0.61
Ave_{15}	**95.67**	**91.08**	96.36	**0.85**	94.09	87.93	95.14	0.80	90.81	61.50	**97.72**	0.65
Ave_{30}	**95.11**	**90.34**	95.77	**0.82**	93.68	87.59	94.58	0.77	90.34	59.03	**97.02**	0.61

Comparison of 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. Four quality measures were used to assess the performance of the proposed method: accuracy (Acc), sensitivity (Sens), specificity (Spec), and Cohen's coefficient of agreement (κ). The scores that are better than or equal to those yielded by the competing approaches are shown in boldface. Equal scores are marked with asterisks. Evaluation was based on analysis of 30-s airflow epochs classified as positive, if at least 5 s of the epoch was affected by hypopneic/apneic activity. For each examined patient, the number of epochs was equal to 899. Ave_{15}: average score for a representative group of 15 patients, Ave_{30}: average score for all patients.

window sizes M and N [which are the only design parameters of the nonlinear low-pass filter (11)–(12)] were recommended, both algorithms for robust envelope evaluation do not require any tuning.

When comparing results of the proposed approach with those yielded by RemLogic, one should remember that in the first case only the airflow signal is analyzed, i.e., information contained in other PSG channels (oximetry, EEG, body movement) is not taken into account. Figs. 12 and 13 illustrate two situations where the lack of such additional information is the source of false-positive decisions. In the first case (61% of all false positives when Approach I is used, 63% when Approach II is

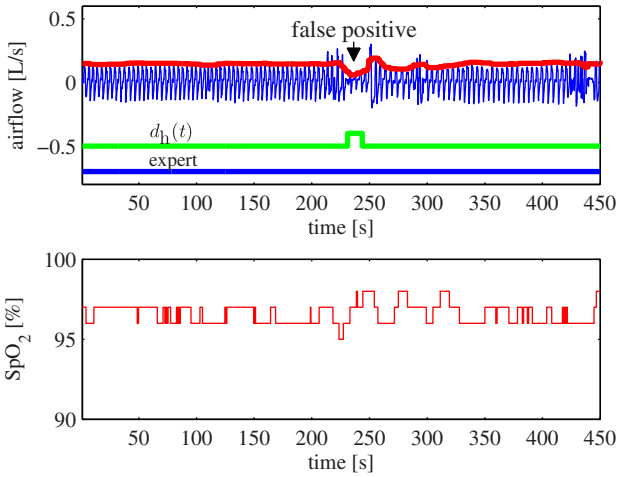

Fig. 12. False positive SAH detection (upper plot) caused by the lack of information about insufficient desaturation (lower plot).

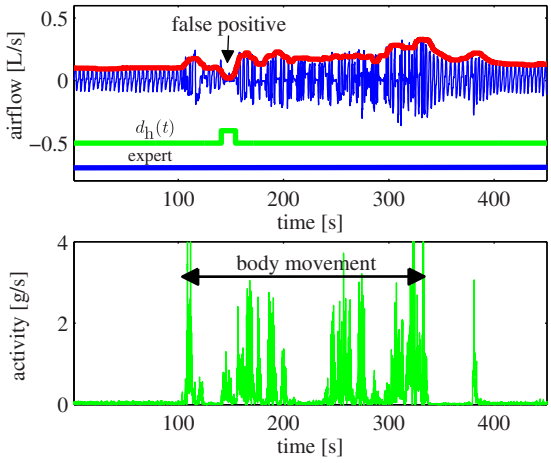

Fig. 13. False positive SAH detection (upper plot) caused by the lack of information about excessive body movement (lower plot).

used), shown in Fig. 12, the drop in the nasal pressure was classified as hypopneic activity even though there was no associated desaturation or arousal. In the second case (33% of all false positives when Approach I is used, 28% when Approach II is used), shown in Fig. 13, reduction of the airflow was accompanied by rapid body movements—in situations like this, pulmonologists regard the corresponding measurements as untrustworthy and temporarily suspend scoring.

Both types of errors pointed out previously (which constitute more than 90% of all false positives) can be easily eliminated by exploiting information drawn from other PSG channels. An example of an erroneous (false negative) decision that cannot be corrected in this way is shown in Fig. 14. In this case, the hypopneic events are not detected by the proposed algorithm because the corresponding baseline values are (slightly) understated. This suggests that the filtering scheme may need some corrections for unstable breathing patterns, such as the one depicted in Fig. 14.

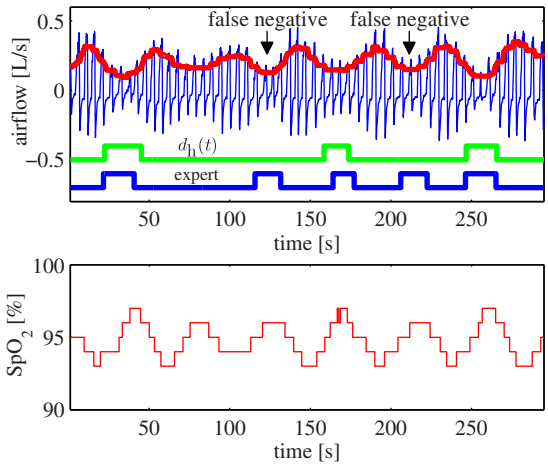

Fig. 14. False negative SAH detection caused by underestimation of the baseline value.

Detection rate of apnea events is another characteristic that should be improved. According to the results presented in Table II, detection rates of SAH events yielded by Approach I are pretty high (average detection rate = 93.6%, minimum detection rate = 84.9%). However, classification of apnea and hypopnea episodes is less successful. In particular, for four patients (#4, #5, #10, and #12), the apnea detection rate is lower than 30%. This means that some further work is needed to develop tools capable of distinguishing both types of events in a more reliable way.

In a group of patients suffering from an obstructive sleep apnea, duration of SAH events is an important diagnostic factor as it can be related to morbidity and mortality. For this reason, precision in determining the start and end points of each event is an important aspect of the compared approaches. Unfortunately, to objectively compare annotation accuracy of RemLogic with that of the proposed approaches, one should have access to unsupported (and hence unbiased) expert decisions based *only* on his/her analysis of PSG recordings. The currently used clinical routine, under which experts approve, reject, or correct decisions made by RemLogic, does not provide such unbiased reference data. Our general observation is that experts seldom correct start/end points proposed by RemLogic unless: 1) the need for changes is evident (gross over/understatement of the event duration) and 2) corrections have important qualitative implications (which usually happens in borderline situations, where the event duration is close to the decision threshold equal to 10 s). As a result, expert decisions on duration of breathing events are usually biased toward those made by the automated detection software.

VII. Conclusion

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 low-pass output filter with a cascade of two nonlinear filters—an SM filter and an RM

filter. Unlike the methods described in the literature, the proposed algorithms do not need initial training or optimization.

REFERENCES

[1] American Academy of Sleep Medicine Task Force, "Sleep-related breathing disorders in adults: Recommendations for syndrome definition and measurement techniques in clinical research," *Sleep*, vol. 22, pp. 667–689, 1999.

[2] W. Lee, S. Naqubadi, and M. H. Mokhlesi, "Epidemiology of obstructive sleep apnea: A population-based perspective," *Expert Rev. Respir. Med.*, vol. 2, pp. 349–364, 2008.

[3] American Thoracic Society, "Cardiorespiratory sleep studies in children: Establishment of normative data and polysomnographic predictors of morbidity," *Amer. J. Respir. Crit. Care Med.*, vol. 160, pp. 1381–1387, 1999.

[4] T. Young, M. Palta, J. Dempsey, J. Skatrud, S. Weber, and S. Badr, "The occurrence of sleep-disordered breathing among middle-aged adults," *New Engl. J. Med.*, vol. 328, pp. 1230–1235, 1993.

[5] T. Young, L. Evans, L. Finn, and M. Palta, "Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women," *Sleep*, vol. 20, pp. 705–706, 1997.

[6] V. Kapur, K. P. Strohl, S. Redline, C. Iber, G. O'Connor, and J. Nieto, "Underdiagnosis of sleep apnea syndrome in U.S. communities," *Sleep Breath*, vol. 6, pp. 49–54, 2002.

[7] Standards of Practice Committee of the American Sleep Disorders Association, "Practice parameters for the indications for polysomnography and related procedures," *Sleep*, vol. 20, pp. 406–422, 1997.

[8] S. Redline, R. Budhiraja, V. Kapur, C. L. Marcus, J. H. Mateika, R. Mehra, S. Parthasarthy, V. K. Somers, K. P. Strohl, L. G. Sulit, D. Gozal, M. S. Wise, and S. F. Quan, "The scoring of respiratory events in sleep: Reliability and validity," *J. Clin. Sleep Med.*, vol. 3, pp. 169–200, 2007.

[9] C. Iber, S. Ancoli-Israel, A.L. Chesson Jr., and S. F. Quan, *The ASSM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*, 1st ed.Westchester, IL, USA: American Academy of Sleep Medicine, 2007.

[10] R. B. Berry, R. Budhiraja, D. J. Gottlieb, D. Gozal, C. Iber, V. K. Kapur, C. L. Marcus, R. Mehra, S. Parthasarthy, S. F. Quan, S. Redline, K. P. Strohl, S. L. Davidson Ward, and M. M. Tangredi, "Rules for scoring respiratory events in sleep: Update of the 2007 AASM manual for the scoring of sleep and associated events," *J. Clin. Sleep Med.*, vol. 8, pp. 597–619, 2012.

[11] P. Várady, T. Micsik, S. Benedek, and Z. Benyó, "A novel method for the detection of apnea and hypopnea events in respiration signals," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 9, pp. 936–942, Sep. 2002.

[12] H. Nakano, T. Tanigawa, T. Furukawa, and S. Nishima, "Automatic detection of sleep-disordered breathing from a single-channel airflow record," *Eur. Respir. J.*, vol. 29, pp. 728–736, 2007.

[13] S. I. Rathnayake, I. A. Wood, U. R. Abeyratne, and C. Hukins, "Nonlinear features for single-channel diagnosis of sleep-disordered breathing diseases," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 8, pp. 1973–1981, Aug. 2010.

[14] B. Koley, "Automated detection of apnea and hypopnea events," in *Proc. IEEE 3rd Int. Conf. Emerging Appl. Inf. Technol.*, Kolkata, India, 2012, pp. 85–88.

[15] B. Koley and D. Dey, "Adaptive classification system for real-time detection of apnea and hypopnea events," in *Proc. Point-of-Care Healthcare Technol. Conf.*, Bangalore, India, 2013, pp. 44–45.

[16] T. Penzel, J. McNames, P. De Chazal, B. Raymond, A. Murray, and G. B. Moody, "Systematic comparison of different algorithms for apnoea detection based on electrocardiogram recordings," *Med. Biol. Eng. Comput.*, vol. 40, pp. 402–407, 2002.

[17] P. De Chazal, C. Henegham, E. Sheridan, R. Reilly, P. Nolan, and M. O'Malley, "Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnea," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 6, pp. 686–696, Jun. 2003.

[18] A. H. Khandoker, J. Gubbi, and M. Palaniswami, "Automated scoring of obstructive sleep apnea and hypopnea events using short-term electrocardiogram recordings," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 6, pp. 1057–1067, Nov. 2009.

[19] M. Bsoul, H. Minn, and L. Tamil, "Apnea MedAssist: Real-time sleep apnea monitor using single-lead ECG," *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 3, pp. 416–427, May. 2011.

[20] A. Burgos, A. Goñi, A. Illarramendi, and J. Bermudez, "Real-time detection of apneas on a PDA," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 4, pp. 995–1002, Jul. 2010.

[21] D. Álvarez, R. Hornero, J. V. Marcos, and F. Del Campo, "Multivariate analysis of blood oxygen saturation recordings in obstructive sleep apnea diagnosis," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 12, pp. 2816–2824, Dec. 2010.

[22] A. Yadollahi, E. Giannouli, and Z. Moussavi, "Sleep apnea monitoring and diagnosis based on pulse oximetry and tracheal sound signals," *Med. Biol. Eng. Comput.*, vol. 48, pp. 1087–1097, 2010.

[23] B. Xie, and H. Minn, "Real-time sleep apnea detection by classifier combination," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 3, pp. 469–477, May 2012.

[24] M. El Shayeb, L.-A. Topfer, T. Stafinski, L. Pawluk, and D. Menon, "Diagnostic accuracy of level 3 portable sleep tests versus level 1 polysomnography for sleep-disordered breathing: A systematic review and meta-analysis," *Canadian Med. Assoc. J.*, vol. 186, pp. E25–E51, 2014.

[25] A. Otero, P. Félix, and M. R. Álvarez, "Algorithms for the analysis of polysomnographic recordings with customizable criteria," *Expert Syst. Appl.*, vol. 38, pp. 10133–10146, 2011.

[26] A. Bartolo, B. D. Clymer, R. C. Burgess, and J. P. Turnbull, "Automatic on-line detection of apneas and hypopneas," in *Proc. IEEE 19th Annu. Conf. Eng. Med. Biol. Soc.*, Chicago, IL, USA, 1997, pp.1066–1069.

[27] S. I. Rathnayake, U. R. Abeyratne, C. Hukins, and B. Duce, "Modelling of polysomnographic respiratory measurements for artefact detection and signal restoration," *Physiol. Meas.*, vol. 29, pp. 999–1021, 2008.

[28] S. Haykin, *Communication Systems*, 5th ed. New York, NY, USA: Wiley 2009.

[29] S. L. Marple, Jr., "Computing the discrete-time analytic signal via FFT," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2600–2603, Sep. 1999.

[30] A. Potamianos, and P. Maragos, "A comparison of the energy operator and the hilbert transform approach to signal and speech demodulation," *Signal Process.* vol. 37, pp. 95–120, 1994.

[31] T. S. Huang, *Two-Dimensional Digital Signal Processing II: Transform and Median Filters.*, New York, NY, USA: Springer-Verlag, 1981.

[32] K. Najarian, and R. Splinter, *Biomedical Signal and Image Processing*, 2nd ed., New York, NY, USA: CRC Press, 2012.

[33] I. Pitas, and A. N. Venetsanopoulos, *Nonlinear Digital Filters: Principles and Applications.*, New York, NY, USA: Springer-Verlag, 1990.

[34] W. Pough, "Skip lists: A probabilistic alternative to balanced trees," *Commun. ACM*, vol. 33, pp. 668–676, 1990.

[35] R. Hettinger. (2010). Regaining lost knowledge [Online]. Available: https://rhettinger.wordpress.com/tag/running-median

[36] J. R. Landis, and G. G. Koch, "Measurement of observer agreement for categorical data," *Biometrics*, vol. 33, pp. 159–174, 1977.

[37] D.G. Altman, *Practical Statistics for Medical Research.*, London, U.K.: Chapman Hall, London, 1991.

[38] T. Shochat, N. Hadas, M. Kerkhofs, A. Herchuelz, T. Penzel, J. H. Peter, and P. Lavie, "The SleepStrip™: An apnoea screener for the early detection of sleep apnoea syndrome," *Eur. Respir. J.*, vol. 19, pp. 121–126, 2002.

[39] T. K. Eman, D. Steward, D. Einhom, N. Gordon, and E. Casal, "Validation of ApneaLink™ for the screening of sleep apnea: A novel and simple single-channel recording device," *J. Clin. Sleep Med.*, vol. 3, pp. 387–392, 2007.

Marcin Ciołek received the M.Sc. degree in automatic control from the Gdańsk University of Technology, Gdańsk, Poland, in 2010, where is currently working toward the Ph.D. degree.

Since 2011, he has also been an Assistant Professor in the Department of Automatic Control, Faculty of Electronics, Telecommunications and Computer Science, Gdańsk University of Technology. His professional research interests include speech, music, and biomedical signal processing.

Maciej Niedźwiecki (SM'13) was born in Poznań, Poland, in 1953. He received the M.Sc. and Ph.D. degrees from the Technical University of Gdańsk, Gdańsk, Poland, and the Dr.Hab. (D.Sc.) degree from the Technical University of Warsaw, Warsaw, Poland, in 1977, 1981, and 1991, respectively.

He spent three years as a Research Fellow with the Department of Systems Engineering, Australian National University, 1986–1989. He is currently a Professor and the Head of the Department of Automatic Control, Faculty of Electronics, Telecommunications and Computer Science, Gdańsk University of Technology, Gdańsk. His research interests include system identification, statistical signal processing, and adaptive systems. He is the author of the book *Identification of Time-varying Processes*, New York, NY, USA: Wiley, 2000.

Dr. Niedźwiecki served as a Vice Chairman of Technical Committee on Theory of the International Federation of Automatic Control (IFAC) during 1990–1993. He is currently an Associate Editor for IEEE TRANSACTIONS ON SIGNAL PROCESSING, a Member of the IFAC committees on modeling, identification and signal processing and on large scale complex systems, and a Member of the Automatic Control and Robotics Committee of the Polish Academy of Sciences.

Stefan Sieklicki was born in Warsaw, Poland, in 1952. He received the M.Sc. degree from the Technical University of Gdańsk, Gdańsk, Poland, in 1976, and the Ph.D. degree from the Technical University of Prague, Prague, Czech Republic, in 1986.

Since 1977, he has been a Lecturer in the Department of Automatic Control, Faculty of Electronics, Telecommunications and Computer Science, Gdańsk University of Technology, Gdańsk. His research interests include biomedical signal processing and automation in agriculture.

Jacek Drozdowski was born in 1973. He received the M.D. and Ph.D. degrees from the Medical University of Gdańsk, Gdańsk, Poland, in 1998 and 2005, respectively.

He is currently an Assistant Professor in the Department of Pneumology and Allergology, Medical University of Gdańsk. His research interests include sleep breathing disorders.

Janusz Siebert was born in 1953. He received the M.D., Ph.D., and D.Sc. degrees from the Medical University of Gdańsk, Gdańsk, Poland, in 1978, 1989, and 2001, respectively.

He was the Dean of Medical Faculty, Medical University of Gdańsk, and the Director of the University Center for Cardiology. He is currently a Professor and the Head of the Department of Family Medicine, Medical University of Gdańsk, and a Professor at the Gdańsk University of Technology, Gdańsk. His research interests include the development of noninvasive techniques for evaluation of circulatory and respiratory systems, cardiac protection during cardiac surgery procedures and diagnosis of cardiac arrhythmias.