

Detection of compressed tracheal sound patterns with large amplitude variation during sleep

A. Kulkas · E. Rauhala · E. Huupponen · J. Virkkala ·
M. Tenhunen · A. Saastamoinen · S-L. Himanen

Received: 13 August 2007 / Accepted: 5 February 2008 / Published online: 21 February 2008
© International Federation for Medical and Biological Engineering 2008

Abstract The objective of the present work was to develop automated methods for the compressed tracheal breathing sound analysis. Overnight tracheal breathing sound was recorded from ten apnoea patients. From each patient, three different types of tracheal sound deflection pattern, each of 10 min duration, were visually scored, viewing the compressed tracheal sound curve. Among them, high deflection patterns are of special interest due to the possible correlation with apnoea–hypopnoea sequences. Three methods were developed to detect patterns with high deflection, utilizing nonlinear filtering in local characterization of tracheal sounds. Method one comprises of local signal maximum, the second method of its local range, and the third of its relative range. The three methods provided 80% sensitivity with 57, 91 and 93% specificity, respectively. Method three provided an amplitude-independent approach. The nonlinear filtering based methods developed here offer effective means for analysing tracheal sounds of sleep-disordered breathing.

Keywords Tracheal breathing sound · Apnoea · Automated detection · Nonlinear filtering · Compressed waveform

1 Introduction

Sleep apnoea is a common sleep disorder with a prevalence of 4% in adult men and 2% in adult women [20]. Sleep apnoea–hypopnoea syndrome (SAHS) is characterised by periods of cessation of breathing lasting over 10 s (apnoea) and periods of clearly reduced (<50%) breathing lasting over 10 s (hypopnoea) [3, 20]. These breathing abnormalities impair not only the daytime functions [17] but also physiological systems and, therefore, the early diagnosis of sleep apnoea–hypopnoea syndrome is of great importance.

Clinical assessment of SAHS has traditionally been based on conventional polysomnography. A detailed visual analysis of the polysomnographic signals gives information about respiratory, cardiovascular, and physiological problems [21] and it is considered to be the gold standard method for the diagnosis of SAHS [3]. The visual analysis provides a solid background for the diagnosis but it is time-consuming, costly, and subjective. Rapid development of technology has provided new recording methods and possibilities for automated analysis.

Previous studies have shown the promising usefulness of tracheal breathing sounds in the diagnostics of SAHS. Intensity changes of tracheal breathing sounds have been shown to correlate with apnoea and hypopnoea events [8]. Moreover, the tracheal sounds analysis can be used to analyse the severity of the SAHS using relative changes of the total tracheal sound power [18]. Apnoea events have been shown to be surrounded by power increases [16].

A. Kulkas (✉) · E. Rauhala · E. Huupponen · J. Virkkala ·
M. Tenhunen · A. Saastamoinen · S-L. Himanen
Department of Clinical Neurophysiology,
Medical Imaging Centre, Pirkanmaa Hospital District,
Tampere, Finland
e-mail: antti.kulkas@pshp.fi

E. Rauhala
Department of Clinical Neurophysiology,
Satakunta Central Hospital, Pori, Finland

J. Virkkala
Sleep Laboratory, Brain Work Research Center,
Finnish Institute of Occupational Health, Helsinki, Finland

Snoring is associated with SAHS [3] and has been referred to as the earliest and most consistent sign of upper airway dysfunction leading to SAHS [9]. Analysis of snoring sounds has been widely utilized in the assessment of sleep-disordered breathing [10, 11, 15]. Intensity levels, [25], pitch analysis [1], and formant frequencies [19] of snoring sounds have shown correlation between snoring and SAHS. Also, higher order spectra based algorithms [2] and a technique utilizing sub-band energy distributions [6] have been developed for snoring analysis. In a recent study, a database of applied acoustic measurements of oronasal respiration in children was used to compare apnoea-hypopnoea indices (AHI) with snoring indices [4]. Snoring density and loudness of snoring was correlated with increasing SAHS severity [4].

In long-term EEG monitoring, compressed EEG signals are often used to facilitate visual detection of very slow waveform alternations showing underlying trends [12]. Compressed signal analysis has also been applied in continuous scale sleep depth measurements [14, 23]. In these studies, spectral mean frequency was computed over the night showing the whole night sleep depth profile at one time.

Also, experience has been gained on sleep stage profile estimation, for instance, with a neural network approach [5].

For tracheal breathing sounds, the visual analysis of compressed signals has not been widely applied previously. Our pilot study of tracheal sound analysis showed that it is possible to perform a visual analysis of tracheal breathing sounds based on the compressed waveform of the data [13, 22]. While viewing the sound curve of the whole night as local minimum and maximum traces with a time resolution of 20 s, different deflection patterns can be identified from the compressed waveform [13]. Examples of the raw sound data during different deflection patterns, denoted as thin, thick, and plain, are shown in Fig. 1a and the compressed sound data in Fig. 1b. The compressed waveforms show that during the thin pattern, the extreme points of the amplitudes are higher than during the thick and plain patterns but low in variation. In a thick section, loud noises seem to be accompanied by silent sections in-between, which may correspond to apnoeas [13]. These seem to lead to high deflection in the compressed tracheal sound waveform during the thick section. During plain sections, the maximum amplitude levels are low and nearly constant, which leads to low variation and low absolute maximum amplitudes in the compressed sound waveforms. The objective of the present work was to develop detection methods for the analysis of tracheal sound curve to identify the thick patterns of high deflection and to differentiate them from thin and plain patterns.

2 Materials and methods

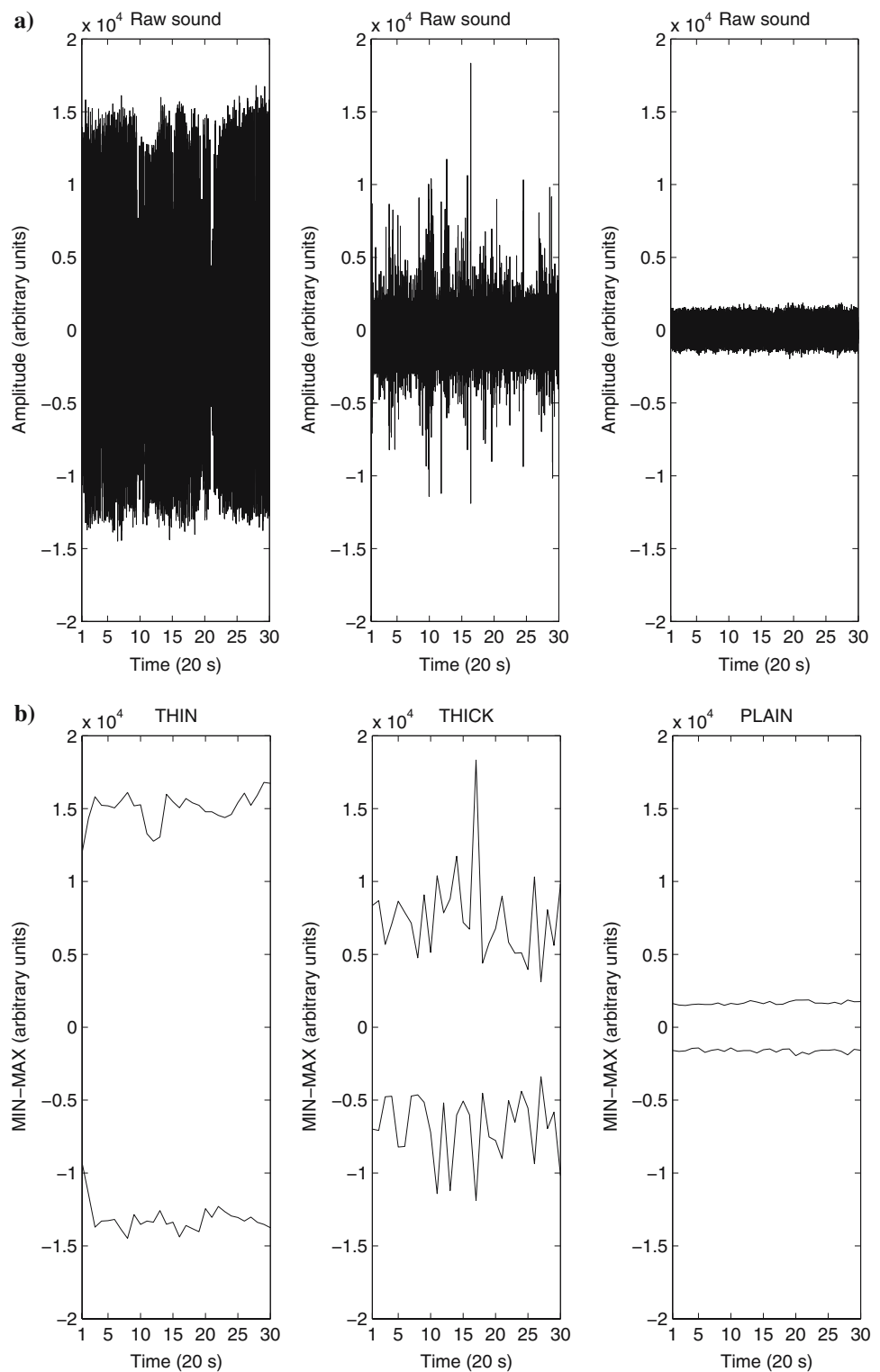
2.1 Recordings

The analysed data in the study consist of polysomnographic recordings from nine male and one female patient. The mean \pm standard deviation (SD) age of the patients was 40.9 ± 12.1 years, the mean body mass index \pm SD was 28.4 ± 6.2 kg/m², and the mean AHI \pm SD was 16.6 ± 9.9 events/h. All the polysomnograms were recorded in the sleep laboratory of Tampere University Hospital in the Pirkanmaa Hospital District and the study was approved by the local ethical committee. The digital polysomnographic recorder Embla N7000 and the Somnologica studio software (Medcare[®], Iceland) were used as the recording system. The recording montage consisted of six electroencephalographic derivations (Fp1-A2, Fp2-A1, C3-A2, C4-A1, O1-A2, O2-A1), two electro-oculographic channels and submental electromyography (EMG), electrocardiogram (ECG), nasal pressure, thermistor, thoracic and abdominal respiratory movements, body position, anterior tibialis muscle EMG, blood oxygen saturation (SaO₂) and pulse rate by a finger pulse oximeter (Nonin XPOD[®], Nonin Medical Inc., USA), transcutaneous measurement of carbon dioxide (TcCO₂, Tina TCM4, Radiometer, Denmark), the Emfit sleep mattress (Emfit Ltd, Vaajakoski, Finland) and a tracheal sound recorder. A sampling rate of 1 Hz was used for oximetry (SaO₂ and pulse) and TcCO₂ measurements, 10 Hz for respiratory movements, 500 Hz for ECG, 11,025 Hz for tracheal sound, and 200 Hz for all other signals.

Tracheal sound recording was performed with a small electret microphone, Panasonic WM-60A (Matsushita Electric Industrial Co. Ltd, Kadoma Osaka, Japan) The microphone has a 3 mm deep conical air cavity of 25 mm diameter. The sensitivity of the microphone is 10 mV/Pa and the frequency range is 20 Hz–20 kHz [24]. The microphone is attached to the skin in the suprasternal notch with an adhesive tape ring and with additional taping on the top. The measured breathing sound signal is amplified with a preamplifier unit. After that, the signal is fed into an external sound card USB Sound Blaster Audigy 2 NX (Creative Labs, Singapore) for 24-bit A/D conversion followed by USI-01 USB isolator (MESO, Mittweida, Germany) providing galvanic isolation between the patient and the recording equipment. The -3 dB cut-off frequency of the anti-aliasing filter is 4.5 kHz and the attenuation at 6 kHz is 70 dB. SuperHeLSA software (Pulmer, Helsinki, Finland) provides the raw data from the sounds recorded over the trachea. The data is converted into the Embla data format.

The visual analyses of the data were done in consensus by two experienced clinical neurophysiologists using

Fig. 1 Three 10 min sections of raw and compressed tracheal sound signal. **a** The raw data during thin, thick, and plain patterns. **b** The compressed waveforms of the local extreme points of the raw tracheal sound data with a time resolution of 20 s. In this example, the thin section seems to consist of densely repeating loud sounds. In the thick section, loud sounds seem to be appearing more seldom. During plain section, there are virtually no loud sounds present



Somnologica Studio (Medcare® Flaga, Reykjavik, Iceland). For these purposes, a compressed sound curve of the whole recording was visually viewed as local minimum and maximum traces with a time resolution of 20 s (Fig. 1b). Viewed this way, the thick, thin, and plain

patterns can be seen. Based on our pilot study, the thick patterns seem to correlate promisingly to apnoea-hypopnoea sequences [13]. From each subject, one continuous 10 min thick, thin, and plain sound pattern was visually marked. The mean \pm SD AHI during these 10 min marked

thick sound curves was 65.4 ± 14.3 events/h, during the thin sections 0 ± 0 events/h, and during plain sound curve 0 ± 0 events/h.

2.2 Detection methods

The developed detection methods aim at replicating the visual analysis to find the visually marked thick sound curve patterns and to distinguish them from the plain and thin sound curve sections. The automated analysis is performed with a time resolution of 1 s using a sliding window centred at second k of length $w\alpha$ seconds. As the first step, the local maximum, denoted as $A(w\alpha)_k$, of the raw sound signal, denoted as x , is determined as:

$$A(w\alpha)_k = \max(x_{\{k-w\alpha/2\} \times fs : \{k+w\alpha/2\} \times fs}), \tag{1}$$

where fs represents the sampling frequency of the raw sound data (11,025 Hz).

Figure 2 shows examples of the values of $A(w\alpha)_k$ during thin, thick, and plain tracheal sound sections.

As the second level of nonlinear filtering, $A(w\alpha)_k$ is further processed to characterise its behaviour within the second window of length $w\beta$ centred at time index of second k . Local extreme points of the second level are now determined as:

$$U(w\alpha, w\beta)_k = \max(A(w\alpha)_{k-w\beta/2:k+w\beta/2}), \tag{2}$$

$$L(w\alpha, w\beta)_k = \min(A(w\alpha)_{k-w\beta/2:k+w\beta/2}). \tag{3}$$

Using these second level extreme points $U(w\alpha, w\beta)_k$ and $L(w\alpha, w\beta)_k$, a measure of local range, denoted as $R(w\alpha, w\beta)_k$, is then readily obtained as:

$$R(w\alpha, w\beta)_k = U(w\alpha, w\beta)_k - L(w\alpha, w\beta)_k. \tag{4}$$

Finally, a relative range, denoted as $r(w\alpha, w\beta)_k$, is calculated from the previously defined values readily as:

$$r(w\alpha, w\beta)_k = R(w\alpha, w\beta)_k / U(w\alpha, w\beta)_k. \tag{5}$$

The values of the $r(w\alpha, w\beta)_k$ range from 0 to 1, regardless of the $A(w\alpha)_k$ and $R(w\alpha, w\beta)_k$. Examples of the values of $R(w\alpha, w\beta)_k$ during thin, thick, and plain tracheal sound sections are presented in Fig. 3. Figure 4 shows an example of the values of $r(w\alpha, w\beta)_k$ during thick, thin, and plain tracheal sound sections.

In order to reach the best possible correspondence with visual scoring, the adjustable method parameters $w\alpha$ and $w\beta$ of $A(w\alpha)_k$, $R(w\alpha, w\beta)_k$, and $r(w\alpha, w\beta)_k$ should be selected in such a way that the thick sound curve sections are maximally separated from the thin and plain sound curve sections. In the present work, window length $w\alpha$ was selected to vary from 5 to 60 s and window length $w\beta$ from 5 to 100 s with 5 s steps. In the calculations given by

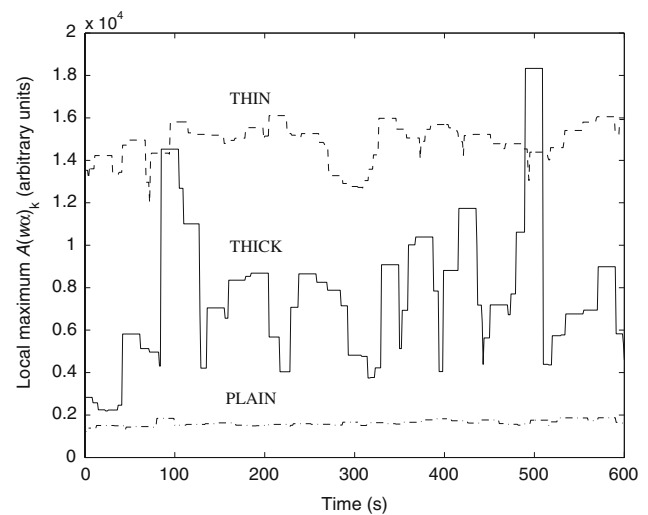


Fig. 2 Local maximum of the amplitude values, denoted as $A(w\alpha)_k$, are presented during thin, thick, and plain tracheal sound sections, of those seen in Fig. 1. The thin section shows highest values followed by the thick section. The lowest values are seen during the plain section. The local maximum was obtained with $w\alpha$ of length 20 s

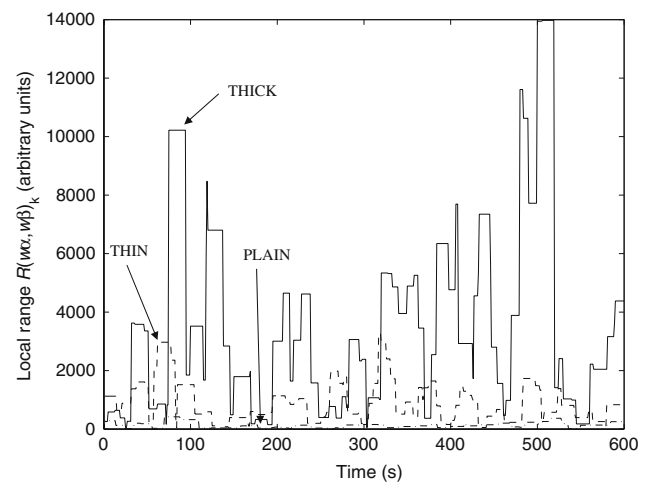


Fig. 3 The local range of the maximum amplitude values, denoted as $R(w\alpha, w\beta)_k$, are represented during thin, thick, and plain tracheal sound sections, of those seen in Fig. 1. The highest values of the local range can be seen during the thick sound curve. The local range was obtained with $w\alpha$ and $w\beta$, both of length 20 s

Eqs. 1–3, the quantities $w\alpha/2$ and $w\beta/2$ were rounded to the nearest integer.

The first detection method utilises $A(w\alpha)_k$ in detecting a thick sound curve. The thick sound curve is detected during the seconds k when the amplitude threshold, denoted as λ_A , is exceeded:

$$A(w\alpha)_k > \lambda_A. \tag{6}$$

In the second detection method, $R(w\alpha, w\beta)_k$ is used. The thick sound curve is detected during the seconds k when the range threshold, denoted as λ_R , is exceeded:

$$R(w\alpha, w\beta)_k > \lambda_R. \tag{7}$$

The third detection method utilises $r(w\alpha, w\beta)_k$. The thick sound curve is detected during the seconds k when the relative range threshold, denoted as λ_r , is exceeded:

$$r(w\alpha, w\beta)_k > \lambda_r. \tag{8}$$

2.3 Performance evaluation

The values of $A(w\alpha)_k$, $R(w\alpha, w\beta)_k$ and $r(w\alpha, w\beta)_k$ were calculated for the whole night at a time resolution of 1 s. The performance evaluation was done considering specifically all the marked 10 min (thick, thin, and plain) sound sections from each patient, totalling 5 h. Results were counted at 1 s precision. A true positive finding was calculated if the detection of thick curve was simultaneous to visually scored thick section. If a thick curve was detected and there was no corresponding visually analysed thick section, a false positive finding was counted.

Receiver operator characteristics (ROC) curves were determined for all the three methods using all the combinations of $w\alpha$ and $w\beta$ resulting in a total of 240 ROC curves per method. In the detection methods one and two, thresholds λ_A and λ_R ranged from 0 to 20,000 with 200 steps, which led to 101 different values. In the detection method three, threshold λ_r ranged from 0 to 1 with 0.01 steps, which led to 101 different threshold values. The best ROC of each method was found based on the area under curve (AUC).

3 Results

Examples of the local maximum, local range, and relative range values during thick, thin, and plain sound sections are seen in Figs. 2, 3 and 4, respectively. In Fig. 2, the thin and thick sections show clearly higher values of $A(w\alpha)_k$ than the plain section reflecting the increased loudness of the sound during thin and thick sections. In Fig. 3, $R(w\alpha, w\beta)_k$ clearly shows the largest values during the thick section and, therefore, the differentiation from thin and plain sections seems possible. In Fig. 4, $r(w\alpha, w\beta)_k$ shows good differentiation of thick section from thin and plain sections.

Overall, the best thick pattern detection results were obtained with method three, based on the relative range, followed closely by method two utilising the local range. The poorest results were obtained with method one with a local maximum amplitude. The best results were obtained using a window length $w\alpha$ of 60, 25, 20 s and $w\beta$ of 5, 100, 90 s, in methods one, two, and three, respectively.

ROC curves representing the outcomes of these method versions are depicted in Fig. 5. The corresponding AUC

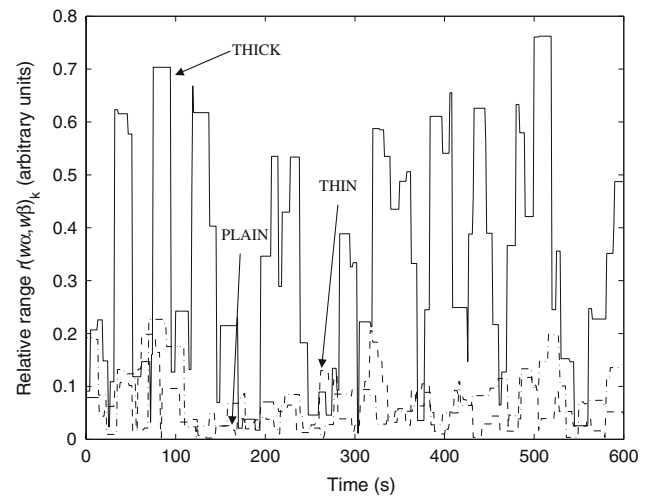


Fig. 4 Relative range values, $r(w\alpha, w\beta)_k$, are presented during thin, thick, and plain tracheal sound sections, of those seen in Fig. 1. The highest values are obtained during the thick sound curve. The values during the thin and plain sections are lower and of the same magnitude. The relative range was obtained with $w\alpha$ and $w\beta$, both of length 20 s

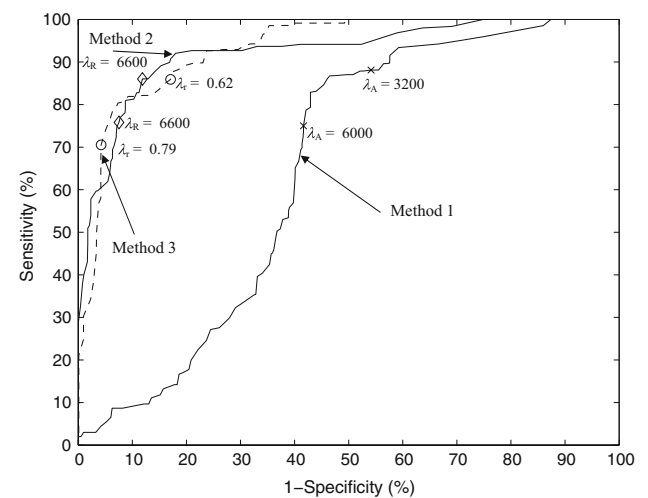


Fig. 5 The outcome of the thick sound curve detection with the three developed methods. On the ROC curves, two threshold values of interest are indicated with crosses in method one, with diamonds in method two, and with circles in method three. The ROC curves were obtained using the window lengths $w\alpha$ and $w\beta$ providing the overall best thick pattern detection performance, being 60 and 5 s, 25 and 100 s, 20 and 90 s in methods one, two, and three, respectively

values for the methods one, two, and three were 0.65, 0.92, and 0.93, respectively. As an example of a thick sound curve detection performance, the overall best method was method three, which provided sensitivity and specificity of 80 and 93% while the methods one and two provided specificities of 57 and 91% with 80% sensitivity, respectively.

4 Discussion

In the present initial study, we developed three new computational methods aiming to mimic the visual compressed sound signal observations. To our knowledge, computational compressed sound signal analysis has not been studied previously. Here we studied rather long compressed sound curve sections including specifically the visually scored thick, thin, and plain patterns. Based on our previous work, thick patterns seem to correlate promisingly to apnoea–hypopnoea sequences [13]. In the present work, we focused on the detection of these thick patterns irrespective of the underlying patient conditions during all these patterns, which are topics of additional studies.

Nonlinear filtering was applied here to determine the local maximum amplitude, local range, and relative range in the detection methods. The developed methods convey a large amount of information of the sound signal dynamics due to the relatively long window lengths used. In addition, the developed methods provide a clear data reduction (99.98%) and are computationally efficient, which support the analysis of large datasets.

The developed method number one is dependent on the absolute local maximum sound amplitude, $A(w\alpha)_k$. It was not able to differentiate well the thin and thick sections due to partly similar sound waveform behaviour of these sections. Because both of these sections contain loud sound periods, differentiation between these sections based on absolute amplitudes alone is not feasible. However, absolute amplitude could differentiate well between quiet and loud sound periods inside analysed sections. The method two utilised the local range of the maximum sound amplitudes, $R(w\alpha, w\beta)_k$. Thin sections typically consist of densely repeating loud sounds whereas plain sections are mainly composed of low to moderate sounds, varying at breath cycle time scale. During thin and plain sections, the values of $R(w\alpha, w\beta)_k$ are rather small due to the uniformity of the sections. It is known that apnoeas cause reduction in the sound intensities [8, 18]. Therefore, it seems quite possible that the thick sections could consist of loud sounds, during breath cycles, followed by somewhat longer silent sections, which together lead to higher $R(w\alpha, w\beta)_k$ values. The relative range, $r(w\alpha, w\beta)_k$, of method three was designed to be independent of absolute sound amplitudes, adding to the applicability of this method. The results showed that the methods two and three were overall the best, being the most potential approaches for SAHS diagnosis.

Regarding the time scales of the analysed phenomena, the visual analysis was based on the local maximum and minimum using the previously determined time resolution of 20 s [13]. There exists only one absolute time scale in apnoea definition, the 10 s minimum duration of

respiratory pause [3, 20]. However, the mean duration of apnoea events is somewhat longer, about 40 s [20]. The analysis window lengths tested here ranged from 5 to 100 s with 5 s steps, which covered well the range of interest.

A method of breath cycle determination has previously been developed applying pulsatile pressure signal acquired by a pressure sensor under a pillow for the analysis of sleep apnoea [7]. Possibly, in future studies, estimation of individual breath cycles from tracheal sound may also provide additional information.

Analysing the sound signal in compressed form seems to provide an effective means for analysing breathing sounds during sleep. The methods of compressed sound signal analysis could provide fast overall characterisation of breathing patterns over the night. Computational methods could perhaps be applied to a larger clinical population to support the SAHS diagnosis and treatment planning, as more is learned of these phenomena. The tracheal sound analysis is considered to have a great potential in the evaluation of sleep-disordered breathing patterns.

Acknowledgments This study was financially supported by the National Technology Agency of Finland, the Research fund of the Tampere University Hospital, the Jenny and Antti Wihuri foundation, the Tampere Tuberculosis foundation, the Emil Aaltonen foundation, as well as the Finnish Cultural foundation.

References

1. Abeyratne U, Wakwella A, Hukins C (2005) Pitch jump probability measures for the analysis of snoring sounds in apnea. *Physiol Meas* 26:779–798
2. Abeyratne U, Karunajeewa A, Hukins C (2007) Mixed-phase modeling in snore sound analysis. *Med Biol Eng Comput* 45:791–806
3. American Academy of Sleep Medicine (1999) Sleep related breathing disorders in adults: recommendation for syndrome definition and measurement techniques in clinical research. *Sleep* 22:667–689
4. Brietzke S, Mair E (2007) Acoustical analysis of pediatric snoring: what can we learn. *Otolaryngol Head Neck Surg* 136:644–648
5. Caffarel J, Gibson G, Harrison J et al (2006) Comparison of manual sleep staging with automated neural network-based analysis in clinical practice. *Med Biol Eng Comput* 44:105–110
6. Cavusoglu M, Kamasak M, Eroglu O et al (2007) An efficient method for snore/nonsnore classification of sleep sounds. *Physiol Meas* 28:841–853
7. Chen W, Nemoto X, Kanemitsu Y et al (2005) Unconstrained detection of respiration rhythm and pulse rate with one under-pillow sensor during sleep. *Med Biol Eng Comput* 43:306–312
8. Cummiskey J, Williams T, Krumpke P et al (1982) The detection and quantification of sleep apnea by tracheal sound recordings. *Am Rev Respir Dis* 126:221–224
9. Dalmasso F, Prota R (1996) Snoring: analysis, measurement, clinical implications and applications. *Eur Respir J* 9:146–159
10. Fiz J, Abad J, Jane R et al (1996) Acoustic analysis of snoring sound in patients with simple snoring and obstructive sleep apnoea. *Eur Respir J* 9:2365–2370

11. Gavriely N, Jensen O (1993) Theory and measurements of snores. *J Appl Physiol* 74:2828–2837
12. Hellström-Westas L, Rosen I, de Vries L et al (2006) Amplitude-integrated EEG classification and interpretation in preterm and term infants. *NeoReviews* 7:76–86
13. Himanen S-L, Huupponen E, Kulkas A et al (2007) Compressed tracheal sound analysis as a screening method for obstructive apneas and hypopneas. *Sleep Med* 8:S91
14. Huupponen E, Saastamoinen A, Joutsen A et al (2005) Antero-posterior difference in EEG sleep depth measure is reduced in apnea patients. *J Med Syst* 29:527–538
15. Issa F, Morrison D, Hadjuk E et al (1993) Digital monitoring of sleep-disordered breathing using snoring sound and arterial oxygen saturation. *Am Rev Respir Dis* 148:1023–1029
16. Kaniusas E, Pfützner H, Saletu B (2005) Acoustical signal properties for cardiac/respiratory activity and apneas. *IEEE Trans Biomed Eng* 52:1812–1822
17. Melamed S, Oksenberg A (2002) Excessive daytime sleepiness and risk of occupational injuries in non-shift daytime workers. *Sleep* 25:315–322
18. Nakano H, Hayashi M, Oshima E et al (2004) Validation of a new system of tracheal sound analysis for the diagnosis of sleep apnea-hypopnea syndrome. *Sleep* 27:951–957
19. Ng A, Koh T, Baey E et al (2008) Could formant frequencies of snore signals be an alternative means for the diagnosis of obstructive sleep apnea? *Sleep Med* (in press)
20. Penzel T, McNames J, deChazal P (2002) Systematic comparison of different algorithms for apnoea detection based on the electrocardiogram recordings. *Med Biol Eng Comput* 40:402–407
21. Roux F, D'Ambrosio C, Mohsenin V (2000) Sleep-related breathing disorders and cardiovascular disease. *Am J Med* 108:396–402
22. Saarenpää-Heikkilä O, Lapinlampi A-M, Himanen S-L (2007) Possible connection with thin signal curve in compressed tracheal sound analysis and flow limitation pattern in children. *Sleep Med* 8:S91–S92
23. Saastamoinen A, Huupponen E, Värrä A et al (2007) Systematic performance evaluation of a continuous-scale sleep depth measure. *Med Eng Phys* 29:1119–1131
24. Sovijärvi A, Helistö P, Malmberg L et al (1998) A new versatile PC-based lung sound analyzer with automatic crackle analysis (HeLSA); repeatability of spectral parameters and sound amplitude in healthy subjects. *Technol Health Care* 6:11–22
25. Van Brunt D, Lichstein K, Noe S et al (1997) Intensity pattern of snoring sounds as a predictor for sleep-disordered breathing. *Sleep* 20:1151–1156