www.czasopisma.pan.pl www.journals.pan.pl

ARCHIVES OF ACOUSTICS
Vol. **45**, No. 1, pp. 141–151 (2020)
DOI: 10.24425/aoa.2020.132490

Research Paper

A Classification Method Related to Respiratory Disorder Events Based on Acoustical Analysis of Snoring

Can WANG[(1)], Jianxin PENG[(1)*], Xiaowen ZHANG[(2)*]

[(1)] *School of Physics and Optoelectronics*
South China University of Technology
Guangzhou, 510640, China

[(2)] *State Key Laboratory of Respiratory Disease*
Department of Otolaryngology-Head and Neck Surgery
Laboratory of ENT-HNS Disease, First Affiliated Hospital
Guangzhou Medical University
Guangzhou 510120, China
*Corresponding Author e-mail: phjxpeng@163.com, entxiaowen@163.com

(*received December 8, 2018; accepted October 3, 2019*)

Acoustical analysis of snoring provides a new approach for the diagnosis of obstructive sleep apnea hypopnea syndrome (OSAHS). A classification method is presented based on respiratory disorder events to predict the apnea-hypopnea index (AHI) of OSAHS patients. The acoustical features of snoring were extracted from a full night's recording of 6 OSAHS patients, and regular snoring sounds and snoring sounds related to respiratory disorder events were classified using a support vector machine (SVM) method. The mean recognition rate for simple snoring sounds and snoring sounds related to respiratory disorder events is more than 91.14% by using the grid search, a genetic algorithm and particle swarm optimization methods. The predicted AHI from the present study has a high correlation with the AHI from polysomnography and the correlation coefficient is 0.976. These results demonstrate that the proposed method can classify the snoring sounds of OSAHS patients and can be used to provide guidance for diagnosis of OSAHS.

Keywords: acoustical analysis; feature extraction; support vector machine; snoring sound.

1. Introduction

Obstructive sleep apnea hypopnea syndrome (OSAHS) is considered to be one of the most prevalent sleep-related breathing disorders, with an enormous effect on public health. It can be characterized by repeated episodes of complete or partial cessation of breathing while sleeping. It is also recognized as an independent risk factor for several clinical consequences, including daytime sleepiness (Rakel, 2009), cerebrovascular disease (Martínez-García *et al.*, 2012), systemic hypertension (Nieto *et al.*, 2000), atherosclerosis (Namtvedt *et al.*, 2013), autoimmune diseases (Kang, Lin, 2012), cognitive abnormalities (Adams *et al.*, 2001), stroke (Loke *et al.*, 2012), postoperative complications (Marin *et al.*, 2012), mental health problems (Hrubos-Strøm *et al.*, 2012) and impaired quality of life (Flemons, Tsai, 1997). Polysomnography (PSG) is currently the gold standard for OSAHS diagnosis (Pack, Gurubhagavatula, 1999). Unfortunately, PSG requires an overnight hospital stay in a specially equipped sleep suite, physically connected to more than 15 measurement channels via sensors (Ghaemmaghami, 2009). The PSG diagnosis method has many disadvantages: complex operation, unsuitable for mass screening, time consuming, inconvenient and expensive. Therefore, many researchers have focused on developing more portable and convenient methods for diagnosis and monitoring of sleep related breathing disorders (Hirotaka *et al.*, 2017; Marcel, 2017).

Snoring is the most common symptom of OSAHS and is caused by an obstruction in the airways due to a collapse in the soft tissues of the upper respiratory tract, which results in vibration of the soft tissues of the upper airway which creates the snoring sound. In the past two decades, there has been much attention given to acoustical analysis of snoring (Pevernagie *et al.*, 2010; Jin *et al.*, 2015). Xu *et al.* (2009) studied the acoustical characteristics of snoring for simple snoring patients and OSAHS patients. They found that there was a significant difference in power density ratio and central frequency at 800 Hz between patients with simple snoring and OSAHS. Israel *et al.* (2012) investigated the time domain and frequency domain characteristics of snoring by using an adaptive energy threshold algorithm and a Gaussian mixture model for snoring detection and extracted snoring characteristics such as the Mel Frequency Cepstral Coefficient (MFCC), pitch density and quiet interval time. A Bayesian model was used to classify simple snoring, moderate and severe sleep apnea with an accuracy of 80%. The Apnea-Hypopnea Index (AHI) is an index used to indicate the severity of sleep apnea and is represented by the number of apnea and hypopnea events per hour of sleep. Recently, many studies have sought to predict the AHI through analysis of the differences in the acoustical characteristics of subject snoring (Ng *et al.*, 2008; Xu *et al.*, 2015). Ng *et al.* (2008) quantified the differences in snoring resonance peak between pure snoring and sleep apnea in patients, and used the regression analysis method to fit the relationship between the first formant and AHI. The study achieved 88% sensitivity and 82% specificity in evaluating sleep apnea syndrome in patients. Xu *et al.* (2015) extracted the Earth Mover's Distance (EMD), a method to evaluate the dissimilarity between two multi-dimensional distributions in feature space using the distance measured between single features, to analyze the acoustical characteristics of snoring based on a frequency energy endpoint detection algorithm to detect snoring, and compared the AHI ($\mathrm{AHI}_{\mathrm{EMD}}$) results based on EMD with the results from PSG detection ($\mathrm{AHI}_{\mathrm{PSG}}$). This study obtained more than 93.3% sensitivity and 96.7% specificity for the identification of OSHAS patients.

The occurrence of snoring is not constant and fixed (Becker *et al.*, 1999), and not all snoring episodes are due to the same mechanisms during sleep (Pevernagie *et al.*, 2010). Some studies have already found significant differences between postapneic snores and other types of snores (Perez-Padilla *et al.*, 1993; Fiz *et al.*, 1996; Tobin *et al.*, 1998). Additionally, some snoring events were over-scored and some were under-scored, although the overnight average was reasonable, and thus, only a subset of snores may be useful for indicating OSAHS. However, acoustical analysis of snoring can discriminate between "simple snorers" and patients with OSAHS, but it can be difficult to estimate obstructive AHI accurately without distinguishing the special acoustical characteristics of snoring sounds before, during and after apnea and hypopnea episodes for OSAHS patients. This is likely the most critical reason that acoustical analysis of snoring is inadequate as a robust method for diagnosing OSAHS. Therefore, there is a requirement to establish appropriate and feasible methods that can be used for unequivocal classification of snoring sounds and to identify the special characteristics of snoring sounds in OSAHS patients (Jin *et al.*, 2015). Although methods for acoustical analysis of snoring as a diagnostic tool have become more mature, there is an urgent need for a rigorous, high-efficacy, single snoring event test with a large sample size to reflect the particular features of snoring that can be used to diagnose OSAHS.

In the study presented in this paper, snoring sounds related to respiratory disorder events are divided into four types of snoring: (1) snoring before and after a respiratory disorder event, (2) snoring during apnea, (3) snoring during hypopnea, and (4) simple snoring. A method for predicting AHI for OSAHS patients based on these four types of snoring recognition is investigated. Based on a previous study (Wang *et al.*, 2017), the overnight snoring signals for six patients with OSAHS are automatically extracted. The characteristics of the snoring spectrum such as the center of mass, spectral dispersion and spectral flatness are extracted. A support vector machine (SVM) is used to classify the four types of snoring. Finally, the AHI is predicted according to the type of snoring identified and the related respiratory disorder event.

2. Method

2.1. Subjects

Six habitual snorers referred for a full-night PSG participated in this study. Overnight sound recordings for the entire night were obtained from the First Affiliated Hospital of Guangzhou Medical University. The main outcome of a PSG test to assess the severity of OSAHS is the AHI. The duration of each recording was over 7 h. Table 1 lists the age, gender, AHIs from the PSG test, and body mass indices (BMI) of the six participants.

Table 1. Information of the six participants.

Patient number	Age	Gender	AHI	BMI
1	26	Male	14.7	22.3
2	37	Male	23.4	25.3
3	32	Male	22.9	26.5
4	50	Male	34	29.7
5	65	Male	69.1	28.2
6	63	Female	20.7	25.1

2.2. Recording of snoring sounds

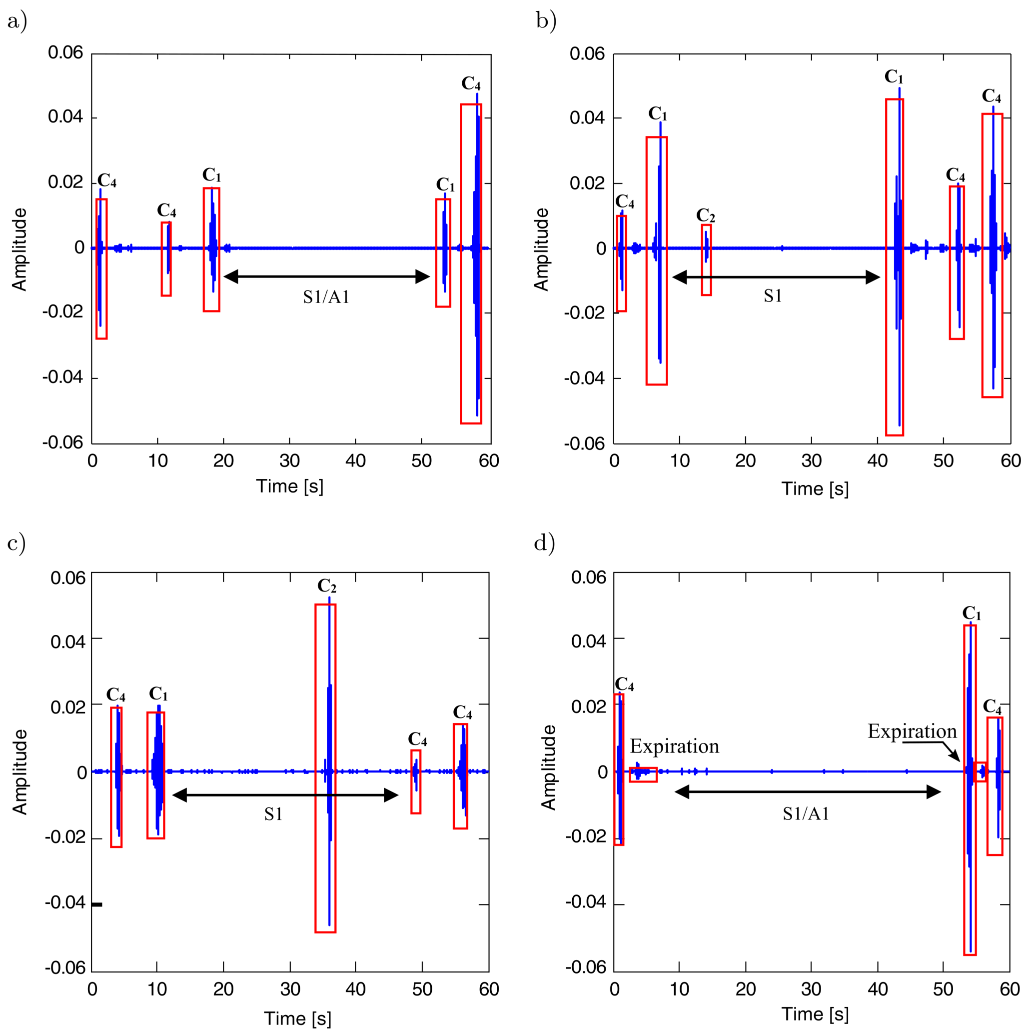
The recording room was a single ward. The air-conditioning was turned off and the double-glazed windows were closed during recording. The background noise in the room was less than 35 dB(A). A digital audio recorder (Roland, Edirol R-44, Japan) with a 40–20000 Hz ±2.5 dB frequency response and a directional microphone (RØDE, NTG-3, NSW, Australia) placed over the patient's mouth was used to record the snoring, at a distance of between 50–70 cm from the patient's mouth, depending on patient movements. The recorded snoring signal was digitized at a sampling rate of 44.1 kHz with 16-bit resolution.

2.3. Snoring classification and AHI prediction

The algorithm described in a previous study (Wang *et al.*, 2017) was used to automatically detect and extract each snoring sound from the overnight snoring recording for each OSAHS patient in order to create datasets for subsequent classification, with the snoring sounds identified by an ENT (ear-nose-throat) specialist.

Based on the PSG results, the ENT experts annotated the complete overnight snoring sounds of the patients with the four types of snoring sounds related to the respiratory disorder events that are shown in Fig. 1. The respiratory disorder events included apnea events and hypopnea events. In Fig. 1, S1 represents an apnea event and A1 represents a hypopnea event. C_1 represents a snoring sound before or after a respiratory disorder event (which includes snoring in the last breathing cycle before apnea, snoring in the first breathing cycle after apnea, snoring in the last breathing cycle before hypopnea, and snoring in the first breathing cycle after hypopnea), C_2 represents a snoring sound during apnea, C_3 represents a snoring sound during hypopnea and C_4 represents a common snoring sound (i.e. a simple snoring sound). All these annotations were used to verify the accuracy of the classification of snoring sounds and the prediction of AHI for OSAHS patients.

Since the AHI is the number of apnea or hypopnea events per hour, it can be predicted by the two respiratory disorder events. The detailed processing steps required to predict AHI based on the four types of snoring sounds are described below.



[Fig. 1]

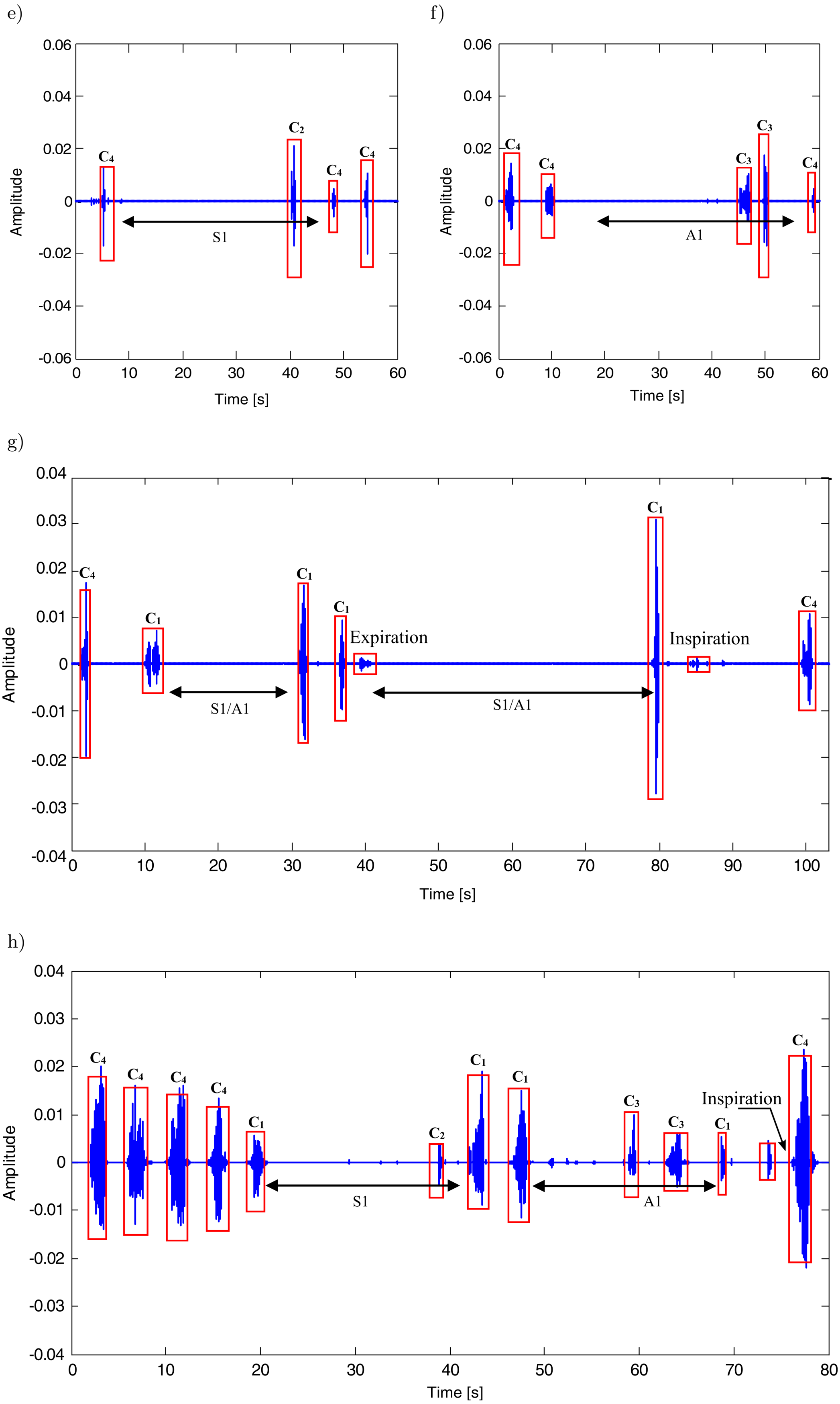


Fig. 1. Respiratory disorder events related to snoring sound definition diagram.

The first step is to determine the position of the C_1 snoring sound in the time sequence from left to right after the four types of snoring sounds are recognized from the full night's snoring recording:

1) If a C_4 snoring sound or a breathing sound occurs before the C_1 snoring sound, and the time interval (the time period from the end of the last snoring sound to the start of the snoring) is less than 10 s, the C_1 snoring sound is determined to be a snoring sound before a respiratory disorder event; if the time interval is more than 10 s, it is determined to be a snoring sound after a respiratory disorder event.

2) If a C_2 or C_3 snoring sound occurs before C_1, C_1 is determined to be a snoring sound after a respiratory disorder event.

3) If the snoring sound before C_1 is also C_1, and the time interval is less than 10 s, the second C_1 is determined to be a snoring sound before a respiratory disorder event; if the time interval is more than 10 s, it is determined to be a snoring sound after a respiratory disorder event.

The second step is to label C_1 snoring sounds that occur before a respiratory disorder event as C_{B1}, and label C_1 snoring sounds that occur after a respiratory disorder event as C_{A1}, according to a left-to-right time sequence of the snoring sound signals. No respiratory disorder event is considered to occur when there are common snoring sounds. A respiratory disorder event is identified between certain common snoring sounds including $C_{B1}C_{A1}$, $C_{B1}C_2C_{A1}$, $C_{B1}C_3C_{A1}$, $C_{B1}C_2$, $C_{B1}C_3$, C_2C_{A1}, C_3C_{A1}, C_{B1}, C_{A1}, C_2, C_3, $C_2C_2C_2$, and so on.

The final step is to calculate the total number of respiratory disorder events in the full overnight snoring sound signals in order to obtain AHI, the number of respiratory disorder events per hour.

2.4. Feature extraction

2.4.1. Spectral centroid

The spectral centroid is used to describe the weighted average frequency of the area under the power spectrum density (PSD) for a given frequency band (FitzGerald, Paulus, 2006). This feature can identify the location of major peaks and indicates where the "center of mass" of the spectrum is located

$$\text{Spectral Centroid} = \frac{\sum_i f_i X_i}{\sum_i X_i}, \tag{1}$$

where X_i is the energy amplitude corresponding to frequency f_i.

2.4.2. Spectral spread

Spectral spread reflects the concentration of a spectrum's energy around its spectral centroid. A smaller spectral spread value indicates that the energy distribution in the frequency domain is more concentrated and that most of the energy is near the spectrum centroid

$$\text{Spectral Spread} = \sqrt{\frac{\sum_i (f_i - \text{Spectral Centroid})^2 x_i}{\sum_i x_i}}. \tag{2}$$

2.4.3. Spectral flatness

Spectral Flatness is the ratio of the geometric mean of the magnitude spectrum to the arithmetic mean of the magnitude spectrum (Peeters, 2004),

$$\text{Spectral Flatness} = \frac{\sqrt[N]{\prod_{s=1}^{N} p_s}}{\frac{1}{N}\sum_{s=1}^{N} p_s}, \tag{3}$$

where p_s represents the total energy of the s-th frequency band, and the spectrum flatness is between 0 and 1. For a completely flat power spectrum, i.e. white noise, the spectrum flatness has a value of 1.

2.4.4. Positive and negative amplitude sum (PN_+), positive and negative amplitude difference (PN_-), positive and negative amplitude ratio (PNAR)

Each snoring segment $x(n)$ is divided into frames to obtain $x_k(n)$, with each frame length equal to 20ms and 50% overlap. The maximum positive amplitude of the k-th frame signal $x_k(n)$ is:

$$P_k = \max[x_k(n)], \quad k = 1, \cdots, K, \tag{4}$$

where k is the total number of frames in the potential snoring segment. Similarly, the maximum negative amplitude of the k-th frame signal $x_k(n)$ is:

$$N_k = \max[-x_k(n)], \quad k = 1, \cdots, K. \tag{5}$$

PN_+, PN_- and PNAR can be defined respectively as:

$$\mathrm{Var}(P_k + N_k), \tag{6}$$

$$\mathrm{Var}(P_k - N_k), \tag{7}$$

$$\mathrm{Var}\left(\frac{P_k + N_k}{P_k - N_k}\right), \tag{8}$$

where $\mathrm{Var}(\cdot)$ is the variance of the values in brackets.

2.4.5. Power ratio at 500 Hz (PR_{500})

The power ratio compares the relative power below and above a specific frequency, which is defined as follows (Xu *et al.*, 2009):

$$\mathrm{PR}_{500} = \frac{\sum\limits_{f_i=0}^{500\,\mathrm{Hz}} P_x(f_i)}{\sum\limits_{f_i=0}^{f_c} P_x(f_i)}, \tag{9}$$

$$p_x(f_i) = \operatorname*{mean}_k P_{xx}(f_i,k), \tag{10}$$

where f_c is the cutoff frequency, $P_{xx}(f_i,k)$ is the PSD for the k-th frame, and $P_x(f_i)$ is the average PSD for each potential snoring segment.

2.4.6. Maximum power ratio (MPR)

The maximum power ratio (MPR) reflects the degree of sound jitter, and for 500 Hz is defined as (Xu *et al.*, 2009):

$$\mathrm{MPR}_{500} = \frac{\sum\limits_{f_i=0}^{500\,\mathrm{Hz}} P_x(f_i)}{\sum\limits_{f_i=0}^{f_c} P_{xx}(f_i,k)}. \tag{11}$$

2.4.7. Spectral entropy (SE)

The spectral entropy (SE) is used to measure the flatness of the PSD, defined as:

$$\mathrm{SE} = -\sum_f P_x(f_i)\ln(P_x(f_i)). \tag{12}$$

2.4.8. Sample entropy

The Sample entropy (Richman, Moorman, 2000) is the measure of the complexity of a time series and is calculated as follows:

1) For an N-point time series, the m-dimensional vector $x_m(i)$ is defined as:

$$x_m(i) = \{x(i), x(i+1), \cdots, x(i+m)\}, \quad 1 \le i \le N. \tag{13}$$

2) The distance between any two m-dimensional vectors $x_m(i)$, and $x_m(j)$ is:

$$\begin{gathered} d[[X_m(i), X_m(j)] = \max[|x(i+k) - x(j+k)|], \\ 0 \le k \le m-1, \quad i \ne j, \quad 1 \le i, \quad j \le N-m. \end{gathered} \tag{14}$$

3) The following formulas are defined:

$$B_i^m(r) = \frac{1}{N-m-1} B_i, \tag{15}$$

$$B^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^m(r), \tag{16}$$

$$A_i^{m+1}(r) = \frac{1}{N-m-1} A_i, \tag{17}$$

$$A^{m+1}(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} A_i^{m+1}(r), \tag{18}$$

$$r = a \times \mathrm{SD}, \tag{19}$$

where r is the number of template matches, B_i is the number of $d[X_m(i), X_m(j)]$ values less than r, A_i is the number of $d[X_{m+1}(i), X_{m+1}(j)]$ values less than r, a is an empirical value, and SD is the standard deviation of the signal in the time domain.

4) The sample entropy is estimated to be:

$$\mathrm{SampEn}(m,r) = \lim_{N\to\infty}\left\{-\ln\left[\frac{A^m(r)}{B^m(r)}\right]\right\}. \tag{20}$$

For a finite length sequence, N is a finite value and the sample entropy is calculated as:

$$\mathrm{SampEn}(Zx,r,N) = -\ln\left[\frac{A^m(r)}{B^m(r)}\right]. \tag{21}$$

2.4.9. Frequency energy spectrum features

The frequency domain (40–2000 Hz) is divided into three frequency bands: B_1(40–300 Hz), B_2(301–850 Hz), and B_3(851–2000 Hz). The maximum sound intensity, the mean sound intensity and the maximum corresponding sound intensity at the corresponding frequency of each frequency band are calculated in the frequency domain. The mean sound intensity corresponds to the frequency spectrum characteristic.

For each snoring sound, 30 features including the spectral centroid, the spectral spread and the spectral flatness are extracted. The detailed features are shown in Table 2.

2.5. Support Vector Machine

The Support Vector Machine (SVM) proposed by Cortes and Vapnik (1995) is a machine learning method developed from statistical learning theory. The basic idea is to maximize the classification interval in the feature space, for the linearly inseparable training data set $T = \{(x_1,y_1)(x_2,y_2)\cdots(x_N,y_N)\}$, where N is the number of potential snoring segments, $x_i \in X \in R^n$, $y_i = \{-1, 1\}$, $i = 1, 2, \cdots, N$. Introducing the relaxation variable ξ_i and the penalty coefficient C, the original optimization problem of the linearly inseparable training data is:

$$\begin{gathered} \min_{\omega,b,\xi} \ \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{N}\xi_i, \\ \text{s.t.} \quad y_i(\omega\cdot x_i + b) \ge 1-\xi_i, \quad \xi_i \ge 0, \quad i = 1, 2, \cdots, N, \end{gathered} \tag{22}$$

where ω and b are the classification hyperplane normal vector and constant terms, $C > 0$.

By introducing the Lagrange function and the Lagrange multiplier α_i, SVM classification attempts to

Table 2. A set of features for snoring sounds.

Time-domain features	
PN_+, PN_-, PNAR	positive/negative amplitude ratio, positive/negative amplitude ratio, positive/negative amplitude ratio
SampEn_{50}, SampEn_{100}	Sample entropy($a = 50$, $a = 100$)
Frequency-domain features	
SC_{mean}, SC_{var}	The mean/variance of spectral centroid
SS_{mean}, SS_{var}	The mean/variance of spectral spread
SF_{mean}, SF_{var}	The mean/variance of spectral flatness
$B_i - I_{\max}$	B_i – maximal sound intensity
$B_i - I_{\text{mean}}$	B_i – mean sound intensity
$B_i - f_{\text{peak}}$	B_i – peak sound frequency
$B_i - f_{\text{mean}}$	B_i – mean sound frequency
PR_{100}, PR_{300}, PR_{500}, PR_{800}	100 Hz/300 Hz/500 Hz/800 Hz power ratio
MPR_{500}, MPR_{800}	500 Hz/800 Hz maximal power ratio
SE	Spectral entropy

B_1 = (40–300) Hz, B_2 = (301–850) Hz, B_3 = (851–2000) Hz

construct an optimal classification hyperplane and then transforms it into a dual quadratic programming problem according to the Karush-Kuhn-Tucker (KKT) condition

$$\min_{\alpha} \quad \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j \left(x_i \cdot x_j\right) - \sum_{i=1}^{N}\alpha_i, \tag{23}$$

$$\text{s.t.} \quad \sum_{i=1}^{N}\alpha_i y_i = 0, \quad 0 \le \alpha_i \le C, \quad i = 1, 2, \cdots, N.$$

The optimal solution α^* is obtained by solving the dual quadratic programming problem as described above, b^* is obtained from $b^* = y_i - \sum_{i=1}^{N}\alpha_i^* y_i \left(x_i \cdot x_j\right)$. By using the kernel function instead of the inner product function $\varphi(x) \cdot \varphi(z)$ in the high-dimensional feature space, the final classification decision function is obtained

$$f(x) = \text{sign}\left(\sum_{i=1}^{N}\alpha_i^* y_i K\left(x_i, x_j\right) + b^*\right). \tag{24}$$

The radial basis function can be defined as follows:

$$K(x, z) = e^{-g\|x-z\|^2}, \tag{25}$$

where g is the parameter for the radial basis function.

SVM was originally designed for the second class of classification problems. Its classification method can be used to construct multiple classifiers to manage multi-class problems. This experiment uses a one-to-one method to realize four types of classification problems, that is, one type of SVM is trained between each set of two classes in the training set. For K types, two $K(K-1)/2$ SVMs are constructed. When classifying unknown samples, each classifier discriminates by category, and finally uses a "voting mechanism" to judge the final category.

3. Results

3.1. Training and test data

The automatic detection model for snoring that was established in a previous study (Wang *et al.*, 2017) is used to analyze a full night of snoring for the six patients in this experiment. The snoring detection results for each of the six patients are shown in Table 3. Sen, Spe, AUC, PPV and Acc represent the sensitivity, specificity, area under the curve, positive predictive value and accuracy values, respectively.

Table 3. Classification results of snoring sounds for 6 patients.

Patient number	Sen [%]	Spe [%]	PPV [%]	AUC	Acc [%]
1	95.98	93.49	99.37	0.9421	95.77
2	90.17	93.63	94.18	0.9272	91.78
3	97.34	73.19	90.84	0.8283	90.87
4	98.15	87.05	95.51	0.8991	95.24
5	93.39	90.44	93.22	0.9156	92.16
6	97.98	77.56	82.48	0.8658	88.16

The inspiratory phase for the six OSAHS patients was labeled as C_1, C_2, C_3, and C_4 by ENT experts for the full night of snoring sounds. There were a total of 14000 snoring sounds between all six patients in the study, including 878 C_1 type snoring sounds (6.27%), 262 C_2 type snoring sounds (1.87%), 691 C_3 type snoring sounds (4.94%), and 12169 C_4 type snoring sounds (86.92%). Since the proportion of C_4 type snoring sounds was much larger than all other snoring sounds, 300 C_4 snoring sounds from each patient were randomly selected from the whole night of snoring sounds, to avoid these sounds from influencing the quantitative difference between the four types of snoring sounds on the classification results in this study. 878 C_1, 262 C_2, 691 C_3 and 1800 C_4 snoring sounds constituted the test sample, and 2000 further snoring samples were randomly selected to find the optimal parameters. After the optimal parameters were set, a 5-fold cross validation method was used to verify the classification performance of the system.

3.2. Classification and forecast results

The 30 feature parameters of 3631 snoring segments were calculated respectively. The Mann-Whitney non parametric test for feature values of the four types of snoring sounds was tested by using the SPSS 19 software. The results showed that most of the features had a better ability to discriminate the four types of snoring sounds ($p < 0.05$). There was no significant difference between C_2 and C_3, C_2 and C_4, C_3 and C_4 for the 40–300 Hz band peak sound intensity ($B_1 - f_{\text{peak}}$) ($p > 0.05$). There was also no significant difference between C_1 and C_2, C_1 and C_3 for the 301–850 Hz band mean sound intensity ($B_2 - f_{\text{mean}}$), and between C_1 and C_2, C_2 and C_3, C_2 and C_4 for the PR_{500}. Therefore, further classification will exclude the above three features and thus 27 features were finally extracted for each snoring segment.

Principal Component Analysis (PCA) is a dimension-reduction tool that can be used to reduce a large set of variables to a small set which will still contain most of the information in the larger set. PCA was used to analyze the 27 features of the snoring samples. All principal components are linear combinations of the set of 27 features. The individual contribution percentage and the cumulative contribution percentage of each of the first ten principle components are shown in Fig. 2. Since the cumulative contribution percentage of the first ten principal components is 93.82%, the first ten principal components were used for classification of the four types of snoring sounds.

The experimental dataset was classified using a nonlinear SVM, and a one-on-one SVM multi-classification algorithm was chosen. The radial basis functions were selected as the kernel functions and the training data was used for learning. The grid search (GS), genetic algorithm (GA) and particle swarm op-

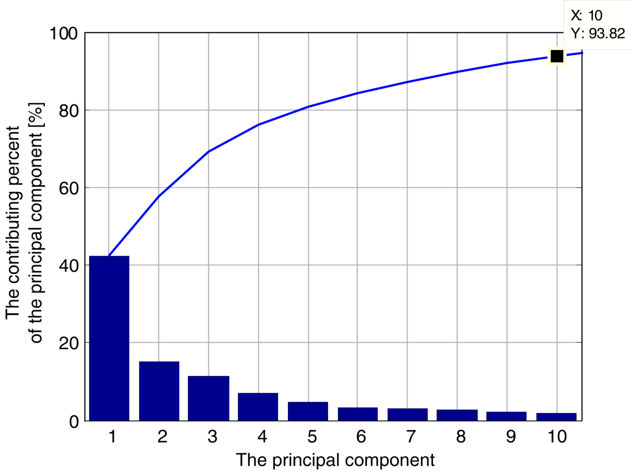


Fig. 2. Principal component analysis.

timization (PSO) algorithm were used to find the optimal parameters C and g, and these optimal parameters were then used to test the experimental dataset. The results of the classification accuracy under the optimal parameter conditions are shown in Table 4. It can be seen from Table 4 that the optimal parameters C and g obtained by each of the three optimal methods were different, but they have almost identical classification results. The total recognition accuracy rate for the four types of snoring sounds under each of the three optimal methods was at least 91.14%. The C_4 snoring sounds had the highest identification rate with a classification sensitivity of more than 98.15%. The identification rates for C_2 and C_3 were lower than that of C_1 and C_4 and the classification sensitivity was between 78.86 ~ 82.20%. Therefore, the SVM classification method can achieve a better classification result for recognition of these four types of snoring sounds.

After setting the optimal parameters for the experiment, the classification performance of the system was verified by using 5-fold cross-validation. The results of each verification were recorded to classify the 3631 snoring sounds which were combined with the 10369 simple snoring sounds. The position of the four types of snoring sounds was used as the basis for calculating the number of apnea events or hypopnea events per hour of the six patients with OSAHS, and to predict the AHI_{TEST} values. A comparison of the results with the AHI_{PSG} values detected by the PSG is shown in

Table 4. Classification results of four types of snoring sound by SVM.

	Optimal parameters		Running time	Classification sensitivity [%]				Overall accuracy [%]
	C	g		C_1	C_2	C_3	C_4	
GS	181.02	0.18	41 min 45 s	88.01	80.49	82.20	98.15	91.38
GA	4.91	0.86	37 min 05 s	88.27	79,67	81.55	98.40	91.20
PSO	23.46	0.52	35 min 09 s	87.24	78.86	82.20	98.28	91.14

Fig. 3 and shows consistency between the experimental prediction of AHI_{TEST} and the PSG-detected AHI_{PSG} value. Figure 4 shows a scatter plot of AHI_{TEST} and AHI_{PSG} for the six patients and regression analysis was used to show a significant positive correlation between AHI_{TEST} and AHI_{PSG} ($r^2 = 0.953$, $p < 0.001$).

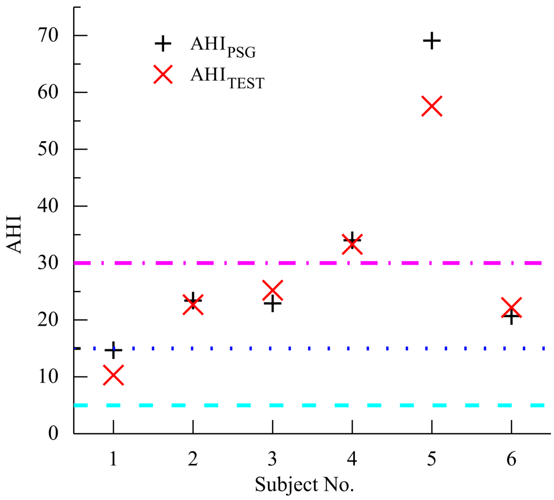


Fig. 3. Comparison of AHI_{TEST} and AHI_{PSG}.

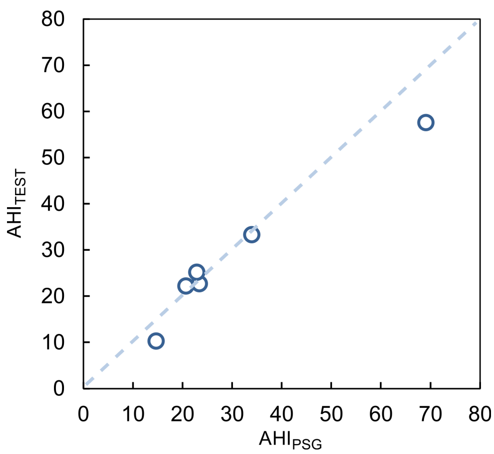


Fig. 4. The relationship between AHI_{TEST} and AHI_{PSG}.

4. Discussions

In this study, 30 feature parameters were automatically extracted from overnight snoring sound recordings of six patients with OSAHS. 27 of these 30 features could be used to successfully discriminate the four types of snoring sounds. Principal component analysis showed that the first ten principal components carried 93.82% of the information in the set of 27 features. These results demonstrate that there were differences in the snoring sounds for the full night's recordings of OSAHS patients, and the acoustical characteristics of the whole night of snoring sounds should be used instead of partial recordings for identification and classification of snoring sounds to obtain more accurate results.

Previous studies have shown that methods based on snoring sound analysis can reach accuracies of between 75.1–92.5% for the detection of snoring sounds for OSAHS patients (Młynczak *et al.*, 2017; Emoto *et al.*, 2018). Młynczak *et al.* (2017) reported that their system achieved a mean of 88.8% accuracy in the differentiation of normal breathing and snoring. Emoto *et al.* (2018) was able to classify low-intensity snoring/breathing episodes (SBEs) and low-intensity non-SBEs that may occur during actual sleep with an average accuracy of 75.10% using artificial neural network analysis. Using ten-fold cross validation, Kim *et al.* (2018) achieved an accuracy of 88.3% in the four group classification and an accuracy of 92.5% in the binary classification. This study demonstrated that their models can be used to estimate the severity of sleep disordered breathing. Hirotaka *et al.* (2017) indicated that the accuracy of their device was 90.7% based on hysteresis extraction. In their study, the four types of snoring were classified by SVM and the classification performance of the system was verified by using five-fold cross validation. Table 4 shows that the overall accuracy rate obtained by the three optimization methods was more than 91.14%. These results indicated a higher accuracy, but the accuracy is still not enough to be used for clinical applications for automatic snoring detection.

The C_4 snoring sound had a better recognition rate than the other three snoring sound types. The classification results can achieve a better recognition rate for all four types of snoring sounds. The optimal parameters C and g under the three optimization methods were different but with similar respective classification results. This shows that the optimal parameters of the SVM model are not unique. A C parameter value that is too high may lead to an over-learning problem with a high time-cost for large samples, and it was shown that there is only a small difference in classification accuracy when a moderate optimization parameter C value was used in the particle swarm optimization algorithm, thus minimizing the computation time. Using the particle swarm optimization algorithm to optimize the SVM parameters had better classification effects on the recognition of the four types of snoring sounds. Follow-up studies should explore new features that can reflect the natural differences in the four types of snoring sounds in order to improve the recognition rate of snoring sounds during apnea and hypopnea.

Ben-Israel *et al.* (2012) estimated AHI_{EST} using a multivariate linear regression model with five features (Mel-Cepstability, Running Variance, Apneic Phase Ratio, Inter-Event Silence and Pitch Density) as the independent variables. They found that AHI_{EST} is correlated with the AHI value from PSG ($r^2 = 0.81$, $p < 0.001$). Xu *et al.* (2015) explored the relationship between AHI obtained from the Earth mover's distance (EMD) calculations and PSG monitoring in the Chinese Han population and found a significant positive correlation between AHI_{PSG} and AHI_{EMD} ($r^2 = 0.950$, $p < 0.001$). This study conducted a classification exercise using a full night's recording of snoring in OSAHS patients and predicted the AHI value as well as the respiratory disorder events through recognition of the four types of snoring sounds. The predicted AHI_{TEST} value is essentially in agreement with AHI_{PSG} from PSG and thus AHI_{TEST} could be used to determine the severity of OSAHS patients. These results indicate that acoustical analysis of snoring could meet the growing screening and diagnosis requirements of patients with OSAHS. Further studies should obtain more snoring sounds from regular snoring patients and OSAHS patients of different genders to improve the accuracy of OSAHS diagnosis and severity based on the four types of snoring sounds, and provide a credible reference that can be adopted for clinical use.

5. Conclusions

In this study, a classification method and an AHI prediction method were presented using the SVM method, based on full overnight snoring sounds related to respiratory disorder events. 30 feature parameters were automatically extracted from the full overnight snoring sound recordings of six patients with OSAHS and it was found that 27 of 30 features could be used to classify the four types of snoring sounds. Principal component analysis demonstrated that the first ten principal components reached 93.82% cumulative variance contribution. The accuracy rate for the recognition of the snoring sounds for the whole night obtained by GS, GA and PSO methods was more than 91.14%. The optimization parameter C of the PSO algorithm was moderate and had the lowest computation time of the three methods. The prediction AHI based on recognition of the four types of snoring sounds related to respiratory disorder events were highly correlated with the results from PSG and may be useful to determine the severity of OSAHS patients. These results demonstrate that the proposed method can classify the snoring sounds of OSAHS patients and provide guidance for diagnosis of OSAHS.

Acknowledgements

This work was supported by National Natural Science Foundation of China (81570904, 11974121).

References

1. Adams N., Strauss M., Schluchter M., Redline S. (2001), Relation of measures of sleep-disordered breathing to neuropsychological functioning, *American Journal of Respiratory and Critical Care Medicine*, **163**(7): 1626–1631, doi: 10.1164/ajrccm.163.7.2004014.
2. Becker H.F. *et al.* (1999), Breathing during sleep in patients with nocturnal desaturation, *American Journal of Respiratory and Critical Care Medicine*, **159**(1): 112–118, doi: 10.1164/ajrccm.159.1.9803037.
3. Ben-Israel N., Tarasiuk A., Zigel Y. (2012), Obstructive apnea hypopnea index estimation by analysis of nocturnal snoring signals in adults, *Sleep*, **35**(9): 1299–305, doi: 10.5665/sleep.2092.
4. Cortes C., Vapnik V.N. (1995), Support vector networks, *Machine Learning*, **20**(3): 273–297, doi: 10.1007/BF00994018.
5. Emoto T., Abeyratne U.R., Kawano K., Okada T., Jinnouchi O., Kawata I. (2018), Detection of sleep breathing sound based on artificial neural network analysis, *Biomedical Signal Processing and Control*, **41**, 81–89, doi: 10.1016/j.bspc.2017.11.005.
6. FitzGerald D., Paulus J. (2006), Unpitched percussion transcription, [in:] A. Klapuri, M. Davy [Eds], *Signal processing methods for music transcription*, Springer, Boston, MA, doi: 10.1007/0-387-32845-9_5.
7. Fiz J.A. *et al.* (1996), Acoustic analysis of snoring sound in patients with simple snoring and obstructive sleep apnea, *European Respiratory Journal*, **9**(11): 2365–2370, doi: 10.1183/09031936.96.09112365.
8. Flemons W.W., Tsai W. (1997), Quality of life consequences of sleep-disordered breathing, *Journal of Allergy and Clinical Immunology*, **99**(2): S750–S756, doi: 10.1016/S0091-6749(97)70123-4.
9. Ghaemmaghami H., Abeyratne U.R., Hukins C. (2009), Normal probability testing of snore signals for diagnosis of obstructive sleep apnea, *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Minneapolis, MN, pp. 5551–5554, doi: 10.1109/IEMBS.2009.5333733.
10. Hirotaka H., Masakazu T., Syunsuke T., Toshikazu S., Hiroshi Y. (2017), Validation of a new snoring detection device based on a hysteresis extraction algorithm, *Auris Nasus Larynx*, **44**(5): 576–582, doi: 10.1016/j.anl.2016.12.009.
11. Hrubos-Strøm H. *et al.* (2012), Sleep apnoea, anxiety, depression and somatoform pain: a community-based high-risk sample, *European Respiratory Journal*, **40**(2): 400–407, doi: 10.1183/09031936.00111411.
12. Jin H. *et al.* (2015), Acoustic analysis of snoring in the diagnosis of obstructive sleep apnea syndrome: a call for more rigorous studies, *Journal of Clinical Sleep Medicine*, **11**(7): 765–771, doi: 10.5664/jcsm.4856.
13. Kang J.H., Lin H.C. (2012), Obstructive sleep apnea and the risk of autoimmune diseases: a longitudinal population-based study, *Sleep Medicine*, **13**(6): 583–588, doi: 10.1016/j.sleep.2012.03.002.
14. Kim T., Kim J. W., Lee K. (2018), Detection of sleep disordered breathing severity using acoustic biomarker and machine learning techniques, *Biomedical Engineering OnLine*, **17**(1): 16, doi: 10.1186/s12938-018-0448-x.

15. Loke Y.K., Brown J.W., Kwok C.S., Niruban A., Myint P.K. (2012), Association of obstructive sleep apnea with risk of serious cardiovascular events: a systematic review and meta-analysis, *Circulation: Cardiovascular Quality and Outcomes*, **5**(5): 720–728, doi: 10.1161/CIRCOUTCOMES.111.964783.

16. Marin J.M. *et al.* (2012), Association between treated and untreated obstructive sleep apnea and risk of hypertension, *JAMA*, **307**(20): 2169–2176, doi: 10.1001/jama.2012.3418.

17. Martínez-García M.A. *et al.* (2012), Cardiovascular mortality in obstructive sleep apnea in the elderly: role of long-term continuous positive airway pressure treatment: a prospective observational study, *American Journal of Respiratory and Critical Care Medicine*, **186**(9): 909–916, doi: 10.1164/rccm.201203-0448OC.

18. Młyńczak M., Migacz E., Migacz M., Kukwa W. (2017), Detecting breathing and snoring episodes using a wireless tracheal sensor – a feasibility study, *IEEE Journal of Biomedical and Health Informatics*, **21**(6): 1504–1510, doi: 10.1109/JBHI.2016.2632976.

19. Namtvedt S.K. *et al.* (2013), Impaired endothelial function in persons with obstructive sleep apnoea: impact of obesity, *Heart*, **99**(1): 30–34, doi: 10.1136/heartjnl-2012-303009.

20. Ng A.K., Koh T.S., Baey E., Lee T.H., Abeyratne U.R., Puvanendran K. (2008), Could formant frequencies of snore signals be an alternative means for the diagnosis of obstructive sleep apnea?,*Sleep Medicine*, **9**(8): 894–898, doi: 10.1016/j.sleep.2007.07.010.

21. Nieto F.J. *et al.* (2000), Association of sleep-disordered breathing, sleep apnea, and hypertension in a large community-based study. Sleep heart health study, *JAMA*, **283**(14):, 1829–1836, doi: 10.1001/jama.283.14.1829.

22. Pack A.I., Gurubhagavatula I. (1999), Economic implications of the diagnosis of obstructive sleep apnea, *Annals of Internal Medicine*, **130**(6): 533–534, doi: 10.7326/0003-4819-130-6-199903160-00019.

23. Peeters G. (2004), *A large set of audio features for sound description (similarity and classification) in the CUIDADO project*, CUIDADO IST Project Report. 54, version 1.0, pp. 1–25.

24. Perez-Padilla J.R., Slawinski E., Difrancesco L.M., Feige R.R., Remmers J.E., Whitelaw W.A. (1993), Characteristics of the snoring noise in patients with and without occlusive sleep apnea, *American Review of Respiratory Disease*, **147**(3): 635–644, doi: 10.1164/ajrccm/147.3.635.

25. Pevernagie D., Aarts R.M., Meyer M.D. (2010), The acoustics of snoring, *Sleep Medicine Review*, **14**(2): 131–144, doi: 10.1016/j.smrv.2009.06.002.

26. Rakel R.E. (2009), Clinical and societal consequences of obstructive sleep apnea and excessive daytime sleepiness, *Postgraduate medicine*, **121**(1): 86–95, doi: 10.3810/pgm.2009.01.1957.

27. Richman J.S., Moorman J.R. (2000), Physiological time-series analysis using approximate entropy and sample entropy, *American Journal of Physiology-Heart and Circulatory Physiology*, **278**(6): H2039–H2049, doi: 10.1152/ajpheart.2000.278.6.H2039.

28. Tobin M.J., Mador M.J., Guenther S.M., Lodato R.F., Sackner M.A. (1998), Variability of resting respiratory drive and timing in healthy-subjects, *Journal of Applied Physiology*, **65**(1): 309–317, doi: 10.1152/jappl.1988.65.1.309.

29. Wang C., Peng J., Song L, Zhang X. (2017), Automatic snoring sounds detection from sleep sounds via multifeatures analysis, *Australasian Physical & Engineering Sciences in Medicine*, **40**(1): 127–135, doi: 10.1007/s13246-016-0507-1.

30. Xu H. *et al.* (2015), Nocturnal snoring sound analysis in the diagnosis of obstructive sleep apnea in the Chinese Han population, *Sleep Breath*, **19**(2): 599–605, doi: 10.1007/s11325-014-1055-0.

31. Xu H., Yu L., Huang W., Chen L., He Y. (2009), A preliminary study of acoustic characteristics of snoring sound in patients with obstructive sleep apnea/hypopnea syndrome (OSAHS) and with simple snoring, *Journal of Audiology and Speech Pathology*, **17**(3): 235–238, http://en.cnki.com.cn/Article_en/CJFDTOTAL-TLXJ200903013.htm.