The Methods of Acoustical Analysis of Snoring for the Diagnosis of OSAHS

Can Wang and Jianxin Peng*
School of Physics and Optoelectronics, South China University of Technology, China

Abstract
Obstructive sleep apnea hypopnea syndrome (OSAHS) is a serious respiratory disorder and the current detection method is Polysomnography (PSG). However, PSG is time consuming, high cost and inappropriate for the diagnosis at home. In recent years, the studies of acoustical analysis of snoring for the OSAHS diagnosis have got rapid development. Researchers have tried to explore a portable monitoring system that is affordable and can offer greater comfort to patients. In this review, we summarize the methods for the OSAHS diagnosis based on the acoustical analysis of snoring. The articles we selected show the acoustical analysis of snoring has a great potential in OSAHS diagnosis. At last, the future research on the acoustical analysis of snoring is prospected.

INTRODUCTION
Obstructive sleep apnea hypopnea syndrome (OSAHS) is a common sleep-related breathing disorder, characterized by the repeatable obstruction of the upper airway during sleep, causing the airflow in the upper airway to decrease or stop, and the clinical manifestation includes snoring at night with apnea and daytime sleepiness. The disease not only impaired the quality of life, but also easily leads to a series of complications, such as neurocognitive dysfunction, metabolic disorders, cardiovascular disease, respiratory failure and cardiopulmonary [1-4]. OSAHS is highly prevalent in adults, approaching 4-5% of men and 2-3% of women between the ages of 30-60 years [5], which threatens people's health and safety.

Polysomnography (PSG) is considered as the gold standard for diagnosis of OSAHS. The patient is required to sleep in the hospital for the whole night to acquire PSG testing whereby measurement equipment with 15 channels is mounted to his/her body. The physiological signals or parameters includes the electrocardiogram (ECG), electroencephalogram (EEG), electroculeogram (EOG), electromyography (EMG), nasal/oral airflow, body positions, body movements and the blood oxygen saturation of a patient which are monitored for the whole night to make the diagnosis [6]. However, PSG’s high cost, time consuming, labor-intensive and its complex operation nature have resulted many patients worldwide not to be treated on time. It has been estimated that more than 90% patients never accepted the related detection in developed countries [7]. Therefore, the researches on finding cheap and portable monitoring method to diagnoze OSAHS have been a hotspot in sleepmedicine at present.

Snoring is caused by the collapse of the soft tissue in the upper airway and the vibration of narrow soft tissues. In addition, the tongue falls back by the action of the gravity aggravating obstruction [8]. Although not everybody with OSAHS is a snorer, the majority of OSAHS people do snore [9]. Snoring appears to be the most intuitive characteristic symptoms of OSAHS patients, so it is often regarded as an important clinical characteristic and plays an indicative role in the diagnosis of OSAHS. As snoring is similar to voice, some researchers have done a lot of studies adopting methods similar to phonetics study [10-14], trying to achieve the diagnosis of OSAHS and determine its severity and the obstructive sites. They rely on the acoustical analysis of snoring as an assistant way for surgery and clinical treatment [11,14-17] and some portable devices are also used at home (home sleep testing, ambulatory cardio-respiratory screening devices) are available, standardized and widely accepted for sleep diagnosis [18].

Acoustical analysis of snoring as an affordable and noninvasive diagnostic method for OSAHS is promising [19]. The diagnostic process mainly includes the feature extraction of each single snoring period from the whole night sleep sounds and the features classification for identifying different snore sounds, as the detailed procedure as shown in Figure (1). The process of the acoustical analysis of snoring is divided into three primary parts: first, all the snores in the acoustic nocturnal signals are detected; then, the most discriminative features for the classification are determined, and finally, the pattern classification algorithms are used to classify the subjects into different classes.

In this article, we aim to review the recent diagnostic
realized real-time analysis and signal enhancement of snore basis function depending on the characteristics of the snore, used wavelet series to deduce and construct a new wavelet basis function is variable. The Wavelet Transform varying non-stationary signals.

which is a very useful tool to analyze the instantaneous and time-signal by using the multi-scale operation of scaling and translation, domain. Wavelet Transform is able to separate the fine details in a wavelets offer a simultaneous localization in time and frequency of a non-stationary signal to be analyzed. Compared to FFT, properties of signal). Wavelet Transform allows the components to analyze the snoring sound of patients is desirable. Fast Fourier Transform (FFT) is an analysis method earliest applied in the snoring frequency domain analysis. For non-stationary signals, the locality in frequency domains by the FFT is poor, it can’t give the detailed results of acoustical analysis of snoring. In recent years, many methods such as the Wavelet transform, the Hilbert-Huang Transform and blind analysis of snoring. In recent years, many methods such as the FFT is a powerful tool for analyzing the components of a stationary signal (no change in the properties of signal) while is less useful in analyzing non-stationary signal (change in the properties of signal). Wavelet Transform allows the components of a non-stationary signal to be analyzed. Compared to FFT, wavelets offer a simultaneous localization in time and frequency domain. Wavelet Transform is able to separate the fine details in a signal by using the multi-scale operation of scaling and translation, which is a very useful tool to analyze the instantaneous and time-varying non-stationary signals.

Due to the great flexibility of wavelet transform, additional wavelet basis function is variable. The Wavelet Transform used wavelet series to deduce and construct a new wavelet basis function depending on the characteristics of the snore, realized real-time analysis and signal enhancement of snore [21]. This problem bases on how to construct a proper wavelet basis function that can be applied to a more accurate acoustical analysis of snoring.

Hilbert-Huang transform

Hilbert-Huang Transform (HHT) is a new analysis method used for nonlinear and non-stationary signals. It decomposes the signals into intrinsic mode function by using empirical mode decomposition based on the local characteristic time scale of signals. Each intrinsic mode function through Hilbert-Huang transform gets the ultimate Hilbert spectrum, which reflects the inherent characteristics of the signals [22]. For the acoustical analysis of snoring, Zhang [23] applied HHT method to establish the signals’ Hilbert spectrum and marginal spectrum. The results showed that HHT had higher time-frequency resolution than the time-frequency distribution established by wavelet transform. HHT avoided the difficulty of the choice of the basis function and non-adaptive limitation of wavelet transform, though there are still some problems such as the optimization of empirical mode decomposition algorithm and boundary problem.

Independent component analysis

Blind source separation is mainly used to remove the signals interference of uncorrelated sources and to recover the source signals we need [24]. Independent component analysis (ICA) is an efficient signal processing method dedicated to solve blind source separation and attempts to confirm transformation to ensure the independence of each component as much as possible [25]. It is a computational method for separating a multivariate signal into additive subcomponents. Vrins et al. [26], explored blind source separation to analyze snoring signals and possible to extract snoring signals by ICA, it gave the encouraging results of the method in their application and also stressed the obstacles, which came from hardware equipment that could cause a bad influence on the signals separation. Moreover, a basic ICA model is classic linear, instantaneous and noiseless, it doesn’t conform to the most realistic model, so the method applied to the snore analysis still needs to be improved.

Higher order statistics analysis

Usually, the statistical characteristics of a Gaussian random variables or a Gaussian random process can be completely expressed by first order and second order statistics. Snoring sounds are non-stationary in nature [27] using low order statistics analysis probably causes the loss of phase information, besides, the signal is classic linear, instantaneous and noiseless, it doesn’t conform to the most realistic model, so the method applied to the snore analysis still needs to be improved.
divided into two categories at present: one uses signal processing method including the short-time energy threshold method \[29\], double-threshold end-point detection \[30\], snoring enhancement method based on autocorrelation character \[21\] and so on. The algorithms of these methods are relatively easy, but the detection accuracy are not high. The other is based on the category theory, such as artificial neural network method \[31\], support vector machines method \[32\] and Gaussian mixture model method \[33\]. These methods are efficient approach with high precision, become the research hotspot of snoring detection recently, but the algorithms are complex, which require an improved algorithm to achieve a higher efficiency and accuracy.

**SNORING FEATURES**

It’s important to find the most discriminative features for the classification of subjects. In the following sections, the major features appeared in many research are reviewed.

**Time-domain features**

**Pitch period:** The fundamental frequency is defined as the lowest frequency of a periodic waveform. Pitch period is a fundamental frequency of an audio waveform, which is one of the essential parameter for the description of speech signals. The autocorrelation method \[34\], cepstrum method \[35\] harmonic product spectrum method \[36\] and pitch estimation algorithm based on higher order statistics \[28\] are the available methods to evaluate the snoring pitch period, which reflects the difference of snore sounds of different types to some extent \[19\].

**Snoring sound intensity/sound pressure level:** Snoring sound intensity is defined as the sound power per unit area. The research found that the snoring sound intensity is connected with apnea hypoventilation index (AHI), represented by the number of apnea and hypopnea events per hour of sleep and used to indicate the severity of sleep apnea), and the greater the snoring sound intensity, the more severe the OSAHS \[37,38\]. Sound pressure level is the logarithmic form of sound pressure and people’s sense of sound pressure is proportional to sound pressure level instead of sound pressure. Peng and Xu’s study \[12\] selected the sound pressure level parameters based on the A-weighted equivalent sound level and accumulative percentile sound level 10, 50, 90, which were significantly different between simple snoring and OSAHS.

**Inter-event silence:** Apnea symptom is easily observed from snoring recording. If respiratory apnea occurs in sleep and the interval duration between two adjacent snoring event lasts 10~60s, this interval duration are called inter-event silence \[39\]. Inter-event silence is related to AHI and reflects the seriousness of OSAHS. Besides, body posture during sleep may affect the acoustic characteristics of snores, such as snoring intensity, but inter-event silence is the feature not affected by snoring intensity. In the study of Ben-Israel, Tarasiuk and Zigel indicated that the inter-event silence was the best feature for predicting AHI \[39,40\].

**Frequency-domain features**

**Formants:** The formants of snoring are the resonant in the upper airway, manifested in the spectral peaks of the sound spectrum and reflected the complex characters of upper airway. Linear prediction coefficient method \[41\] is widely used to estimate the snoring formant. It is shown that the formant frequency may carry the important information of snoring, which can differentiate apnea snorers from benign snorers \[42\].

**Energy spectral density features:** Energy spectral density represents the relation that the signal energy changes with frequency. Spectrum estimation usually employs the FFT \[43\]. Some researches proposed the energy spectral features: the maximal intensity, the mean intensity, the peak frequency and the mean frequency in different frequency bands of sleep sounds were change with AHI, which reflect the severity of OSAHS \[44,45\].

**Power ratio 800 (PR800):** Power ratio 800 (PR800) is defined as the ratio of spectral power below 800 Hz to that from the band 800 Hz to the cut-off frequency. The first formant and PR800 were reported to well reflect structure information of upper airway \[29\]. Some researchers have pointed out that PR800 could achieve the recognition between mild OSAHS and benign snorers. They also have illustrated the great performance of PR800 in diagnosis of OSAHS \[46\]. Furthermore, PR800 itself also provide a promising feature to diagnose the obstructive sites \[47\].

**Mel cepstability:** The name derived from ‘mel frequency cepstrum coefficient stability’, is the feature used frequently in speech recognition. In the snoring analyses aspect, Mel Cepstability is the method of measuring entire night snoring spectrum’s stability. The feature is defined as the sum of variances of 12 MFCCs related to the frames with the highest energy in each snore

\[
\text{MelCepstability} = \frac{\sum_{i=1}^{12} \text{var}(c_i)}{\sum_{k=1}^{\infty} |\sum_{i=1}^{12} S_k[i]|^2} \quad (1)
\]

Where \(z\) means all snore, \(c_i\) is a vector of the \(i\)th MFCC of each snore, and \(\sum |S_k[i]|^2\) is the total energy of the \(k\)th snore. Compared to healthy subjects, the muscle of OSAHS patients in upper airway is unstable, so correspondingly having the smaller Mel Cepstability \[39\].

**Psychoacoustic parameters**

In the sight of psychoacoustics, snore sounds belong to noise, psychoacoustic analysis as an environment noise analysis tool also has a great potential in the diagnosis of OSAHS. Three typical psychoacoustic parameters (loudness, sharpness, roughness) are used to analyze three different snoring sounds (primary snoring; upper airway resistance syndrome; obstructive sleep apnea syndrome) \[48,49\]. The results showed that the acoustical analysis of snoring by psychoacoustic parameters provided a promising way to differentiate the different snoring types.

**The nonlinear features with chaos theory**

It seems that majority studies have focused on the snoring detection by the linear features, such as zero-crossings, MFCCs and formants, but SRSs is complex in nature \[50\], it’s difficult to solve the classification of SRSs depending on traditional linear feature. Some researches indicate the chaotic characteristics of
snoring sounds [51,52]. The largest Lyapunov exponent (LLE) is one of the indices of chaotic nature and quantifies the dynamic stability of a system; Yılmaz and Aniks [53] evaluate the LLE of apnea/hypopnea patients and simple snorers. The results showed the LLE are quite different between apnea/hypopnea patients group and simple snorers group. Moreover, the randomness of a system is called entropy, there also showed that entropy can be used in classification of SRSs [54]. It is likely to achieve more accurate results by linear and nonlinear analysis together.

Other features
To increase the accuracy of the snoring feature extraction, multiple features are often extracted. The Azarbarzin and Moussavi [37] reported that gender, BMI, and height were the parameters that did change the characteristics of snoring sounds significantly. In the process of snoring feature extraction the effect of anthropometric parameters should be comprehensively considered [55]. Beyond that, the ECG signal could give an opportunity to improve the diagnostic accuracy [56,57].

SNORING CLASSIFICATION METHODS

The classification technology can be applied to determine the snoring types and the snoring site by features analysis of snoring. Bayesian classification, k neighbor method (KNN) and support vector machine (SVM) are common methods to realize classification.

Bayesian classification
Bayesian classification is a simple and effective classification algorithm for feature object based on the Bayes principle. The method may be divided into three main steps. First, calculating the priori probabilities of objects’ features. Second, determining the posteriori probability according to Bayesian formula. Last, Bayesian decision is established to identify the classes by the minimum error or the minimum risk principle. The results indicated the good precision in the diagnosis of OSAHS and disease severity. We can also find improved Bayesian classification method to improve the algorithm accuracy appreciably.

Gaussian mixture model
Gaussian mixture model (GMM) was developed to estimate the probability distribution by using multiple Gaussian distribution functions:

\[ p = \sum_{i=1}^{N} w_i f_i (\vec{v}) \]  

Where, \( f_i \) is a Gaussian distribution function. \( \vec{v} \) is the feature vector extracted, the weighted coefficient \( w_i \) between the different Gaussian distribution should satisfy the normalizing condition:

\[ \sum_{i=1}^{N} w_i = 1 \]  

A GMM is often used to model speech, the studies [57] hypothesized that speech signal properties of obstructive sleep apnea patients would be different than those subjects not having obstructive sleep apnea. Each phoneme, which was the basic element of speech, could be represented as a single cluster on the feature space. A GMM classifier was trained and the maximum likelihood estimator was adopted to estimate model parameters. The result achieved a good performance of the specificity and sensitivity, at least 79%, but with a high algorithm complexity and a low real-time [58].

K-nearest neighbor algorithm
K-nearest neighbor (KNN) is one of the simplest machine learning algorithms. It is a mature classification method in theory, which can realize the binary-classification and multiple classifications of subjects. The input of this algorithm is the extracted feature vectors which corresponding to the points on the feature space and the classification results is its output. The distance of two points on the feature space indicates the neighbor degree. To make the right decision, it often adopts the majority vote principle. KNN was used by Mikami et al. [36,59], for classifying oral/nasal snore sounds. Although the KNN method is relatively simple and the parameter only one (k; the number of neighbors), it is easier to obtain a nonlinear classification boundary and don’t have to adjust many hyper parameters in advance, which reduces the complexity of classification boundary. As a result, over 89% of oral and nasal snores are successfully classified, but KNN method appropriate to larger number of samples, for small sample, it fails to identify [36,60].

Artificial neural network
Artificial neural network (ANN) is an information processing system to imitate the structure and function of the brain’s neural network [61]. It forms network structures by a large number of processing units (neurons), and reflects the basic characteristic of the brain. On the one hand, the learning process obtains knowledge from the external environment. On the other hand, the internal neurons store knowledge and information obtained by the learning process. ANN has a high speed information processing and strong self-regulation ability, which can be used in the classification of different snoring types. De Silva, Abeyratne and Hukins [28] applied a neural network-based pattern recognition algorithm for obstructive sleep apnea / non-obstructive sleep apnea classification, the method resulted in a sensitivity of 91 ± 6% and a specificity of 89 ± 5% for test data, achieved the approximate diagnosis of OSA. The complex structure, long training time and tending to occur over fitting phenomenon become the limitation for ANN to diagnose OSA. It is still at the stage of study and exploration around the world [60].

Support vector machine
Support vector machine (SVM) is a new method of machine learning, proposed by Vapnik in 1995. It is a two-class model developed in statistical theory and the basic thought is maximize margin on the feature space. SVM has many applications in...
pattern recognition. Mikami et al. [59], on the establishment of a previous study, adopted Support Vector Machine (SVM) classifier to classify oral and nasal snore sounds. Snoring is nonlinear, based on the spectral properties, they selected seven kernel functions (linear, polynomial, sigmoid, Gaussian, Laplacian, Chi-square and Kullback-Leibler) to identify oral and nasal snore sounds. The results elucidated that at least over 93% classification accuracy was obtained, the highest accuracy reached over 95% by the use of KL kernel and there was not so much difference among seven kernel functions. Moreover, compared to previous study, the classification accuracy has increased about 5% by the use of SVM. The SVM seems to have a good performance for two-class, but it still worth improving for the difficulty of solving kernel functions and multi-class problems [62].

CONCLUSIONS

Snoring is a prevalent disorder, increases with age and is serious in adults. In this study, we review the present state of acoustical analysis of snoring for the OSAHS diagnosis, the acoustical analysis of snoring and detection methods, the features extraction and snoring classification methods are respectively summarized and analyzed. These methods have their own advantages and shortcomings and some of them have achieved good accuracy in the diagnosis of OSAHS compared the PSG method. It is believe that the choice of proper analysis method given the best reflection on the snoring characteristics and the use of improved algorithms from the pattern recognition algorithms above for snoring detection and classification can obtain a better result in the diagnosis of OSAHS.

The practice has so far proven that acoustical analysis of snoring is a noninvasive, convenient and promising method for the diagnosis of OSAHS. However, more research and development are required to clarify whether acoustical analysis of snoring can be realized for diagnosis in the home environment. To improve the diagnostic accuracy, in addition to a large number of subjects, more than one night snoring signals have to be recorded. What’s more, the use of a standardized method to record snoring and reduce the influence of snoring sound intensity are needed in order to compare studies of OSAHS and snoring to advance this field.

ACKNOWLEDGEMENTS

This work is supported by Science and Technology Planning Project of Guangdong Province, China (Grant No. 2013B060100005).

REFERENCES

2. Ip MS, Lam B, Ng MM, Lam WK, Tsang KW, Lam KS. Obstructive sleep apnea is independently associated with insulin resistance. Am J Respir Crit Care Med. 2002; 165: 670-676.  


23. Zhang Y. Snoring signal spectrum analysis based on hilbert-huang
transform, Computer Engineering & Applications. 2011; 47: 11-104.
29. Xu Zhiyong, Qian Qun, Xu Hujie, Huang Weining. The airway monitoring method based on resonance and peak power ratio of snoring and changes, China: CN 10519901A. 2013.
52. Mikami T. Detecting nonlinear properties of snoring sounds for sleep apnea diagnosis, In Bioinformatics and Biomedical Engineering, the 2nd International Conference on Bioinformatics and Biomedical Engineering 2008; 1173-1176.
56. Czopek K. Significance of snoring and other sounds appearing during the night based on the ECG, In Computing in Cardiology Cin C. 2012; 637-640.

