



Estimating a Sleep Apnea Hypopnea Index Based on the ERB Correlation Dimension of Snore Sounds

Limin Hou¹, Qiang Pan^{1*}, Hongliang Yi^{2*}, Dan Shi¹, Xiaoyu Shi¹ and Shankai Yin²

¹ School of Communication and Information Engineering, Shanghai University, Shanghai, China, ² Department of Otolaryngology, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai, China

OPEN ACCESS

Edited by:

Kun Qian,
The University of Tokyo, Japan

Reviewed by:

Qinglin Meng,
South China University of
Technology, China
Jian Guo,
RIKEN Center for Computational
Science, Japan
Xiaodong Huang,
South China University of
Technology, China

*Correspondence:

Qiang Pan
qpan@laas.fr
Hongliang Yi
yihongli@126.com

Specialty section:

This article was submitted to
Health Informatics,
a section of the journal
Frontiers in Digital Health

Received: 03 October 2020

Accepted: 18 December 2020

Published: 01 February 2021

Citation:

Hou L, Pan Q, Yi H, Shi D, Shi X and
Yin S (2021) Estimating a Sleep Apnea
Hypopnea Index Based on the ERB
Correlation Dimension of Snore
Sounds.
Front. Digit. Health 2:613725.
doi: 10.3389/fdgth.2020.613725

This paper proposes a new perspective of analyzing non-linear acoustic characteristics of the snore sounds. According to the ERB (Equivalent Rectangular Bandwidth) scale used in psychoacoustics, the ERB correlation dimension (ECD) of the snore sound was computed to feature different severity levels of sleep apnea hypopnea syndrome (SAHS). For the training group of 93 subjects, snore episodes were manually segmented and the ECD parameters of the snores were extracted, which established the gaussian mixture models (GMM). The nocturnal snore sound of the testing group of another 120 subjects was tested to detect SAHS snores, thus estimating the apnea hypopnea index (AHI), which is called AHI_{ECD} . Compared to the AHI_{PSG} value of the gold standard polysomnography (PSG) diagnosis, the estimated AHI_{ECD} achieved an accuracy of 87.5% in diagnosis the SAHS severity levels. The results suggest that the ECD vectors can be effective parameters for screening SAHS.

Keywords: apnea hypopnea index, correlation dimension, non-linear acoustic characteristics, snore sound, sleep apnea hypopnea syndrome

INTRODUCTION

Snoring is one of the most important symptoms of Sleep Apnea Hypopnea Syndrome (SAHS) and carries much information for diagnosing the upper airway disorder (1). Snoring sounds can be recorded by a non-contact microphone using acoustical property analysis for the screening of SAHS (2, 3). The pitch and spectral characteristics of snoring have been widely applied (4, 5). The total airway response for a snore was extracted to examine SAHS by a higher-order statistics algorithm (6). Multiclass classification of snoring was acquired on the acoustic analysis of snore sounds (7). A genetic algorithm was applied to select the better features that can be extracted from the time and spectral domains of full-night recordings to determine the Apnea Hypopnea Index (AHI) value (8). The rhythmic variations in the snores were described to assess the AHI (9). Hidden Markov models with Mel frequency cepstral coefficients (MFCC) were used to classify subjects into different ranges of the AHI (10). Our previous work used snore spectral information to estimate the AHI (11). Traditional time and frequency analysis and the classic method for snore sounds were adopted in the studies mentioned above.

However, the irregular and turbulent airflow that is produced within the upper airway tissue vibrations that cause the snore, such as the intensity of respiratory airflow, vibration on the soft palate, thick tongue root, and epiglottic hypertrophy, etc. could be non-linear (12).

It was suggested that linear analysis methods were limited and that more useful information could be obtained using chaos theory to analysis the snore (13, 14). The largest Lyapunov exponent and entropy were calculated to classify snore-related sounds with multiclass system (15).

In this paper, a new correlation dimension was proposed to analyze the non-linear properties of snoring sounds for automatic AHI prediction. In contrast to the conventional correlation dimension, the all frequency region was divided into multi-sub-bands on an equivalent rectangular bandwidth (ERB) scale, and the correlation dimensions were calculated in each sub-band. Therefore, ERB correlation dimension (ECD) vectors were extracted rather than a single correlation dimension. The gaussian mixture model (GMM) was applied to build the ECD vector models. Whole-night snore sounds of patients were detected in our experiments, and then, the AHI_{ECD} values were estimated. The early experimental studies have been published in Chinese journals, when the experiment is the number of 60 snorers in reference (16). This research continues to now increased to 120 snorers. In other words, the experimental results of adding 60 people dropped slightly from the original 90 to 87%. It illustrates the robustness of new features. This study further adds a comparison with the classic feature MFCC. Compared to the polysomnography (PSG) diagnosis, our non-linear features achieved higher accuracy than the MFCC based snore spectrum information in the severity levels of the SAHS. The ECD vectors were found to characterize various severity levels of snores.

ERB CORRELATION DIMENSION

The phase space reconstruction technique has been widely used in the field of chaos and fractal theory (17), and it has been used in some applications in medical and speech signals (18–20). It could be more comprehensive disclosure of snore implied information by transforming them to high-dimensional space. The general representation of a snore is a time series. Let a one-dimensional discrete series $s(n)$ be denoted by the snore signal, that is get by sampling rate F_s .

Based on the Takens embedding theorem (21), the phase space reconstruction could transform a one-dimensional time series into a high-dimensional phase space vector $Y \in R^m$ as in Equation (1).

$$Y = [Y_1 \quad Y_2 \quad \dots \quad Y_i \quad \dots \quad Y_I]$$

Where

$$Y_i = [s(n_i) \quad s(n_i + \tau) \quad \dots \quad s(n_i + (m - 1)\tau)]^T \quad (1)$$

Here, τ is the time delay, and m is the embedding dimension. The reconstruction vector Y is an m -dimensional vector with I phase points. The appropriate time delay was selected according to the autocorrelative function (AR function) (22). The time delay τ is an integer multiple of the sampling interval: $\tau = n/F_s$.

The purpose of choosing the embedding dimension is to make the original chaos attractor and the reconstructed attractor

topology equivalent. We used the false nearest neighbor (FNN) method to determine the embedding dimension m (23). As the embedding dimension m increases, the orbit of the chaotic motion will gradually open, and the false nearest neighbors will be gradually eliminated, until the trajectory tends to be stable and the proper m is obtained (24). When the frame length was > 150 ms, the slope of the correlation integral curve increased very slowly. Finally in our snore work, the time delay of 0.75 ms and the embedding dimension of 15 were confirmed by the above method with a frame length of 150 ms.

The traditional correlation dimension has only a single parameter, it is difficult to make a more comprehensive analysis of complicated signals. Based on the ERB scale related to auditory perception (25, 26), several sub-bands were divided from the whole frequency band of the snore signal. Phase space reconstruction was performed in each of the sub-band signals of the snore. Then, the correlation dimension on these sub-bands were calculated, which obtained auditory sub-band ERB correlation dimension (ECD) vectors. The flow chart of extracting the ECD is shown in **Figure 1**.

Equation (2) is the correlation integral $C_q(I, r)$ of the q th subband, which calculates the probability that the distance of paired (Y_{iq}, Y_{jq}) is smaller than r .

$$C_q(I, r) = \frac{1}{I(I-1)} \sum_{i,j=1}^I \theta(r - |Y_{i,q} - Y_{j,q}|) \quad (2)$$

Where $\theta(\cdot)$ is the Heaviside function, and if $x < 0$, $\theta(x) = 0$; if $x > 0$, then $\theta(x) = 1$. The correlation dimension D is estimated based on the ratio of the logarithm of the correlation integral and the logarithm of the distance r , as in Equation (3).

$$D_q = \frac{\ln C_q(I, r + \Delta r) - \ln C_q(I, r)}{\ln(r + \Delta r) - \ln r} \quad (3)$$

Therefore, the correlation dimension of the q th subband was calculated by Equations (2, 3) based on the Grassberger-Procaccia (GP) algorithm (24). Finally, we get the ECD vector by arranging and integrating ERB subband' correlation dimension, as in Equation (4).

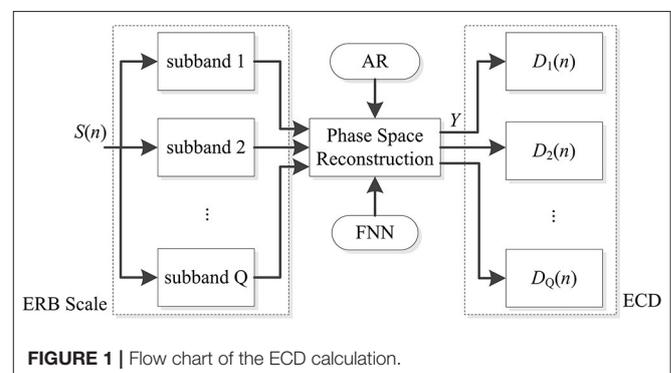


FIGURE 1 | Flow chart of the ECD calculation.

$$ECD(n) = [D_1(n) \ D_2(n) \ \dots \ D_q(n) \ \dots \ D_Q(n)] \quad (4)$$

In this study, the gap of the adjacent sub-band was one bandwidth of the ERB scale, the frequency range of 60 Hz to 4 kHz was divided into 24 sub-bands, that is $Q = 24$, and the 24-dimensional vector of the auditory sub-band ECDs were extracted.

Moore and Glasberg proposed the relationship between frequency and ERB scale (25, 26), as in Equation (5).

$$ERB(f) = 6.23f^2 + 93.39f + 28.52$$

$$ERBS(f) = 11.17628 * \ln\left(1 + \frac{46.06538f}{f + 14678.49}\right) \quad (5)$$

where f is physical frequency in kHz. $ERB(f)$ is the calculated rectangular bandwidth of the equivalent filter in Hz. $ERBS(f)$ is the ERB scale in physical frequency f in Hz.

SNORE ECD FEATURES

Snore Data

Snore sounds were recorded in the sleep monitoring laboratory in the Department of Otolaryngology of Shanghai Jiao Tong University Affiliated Sixth People's Hospital by a non-contact ambient microphone, and simultaneously, polysomnography (PSG) diagnosis was performed. The recording uses a non-contact microphone Sony EM-C10, which is hung on the head of the bed, about 30 cm away from the patient's nose and mouth. The recording sound card is Creative Audigy 4 Value, the desktop computer is Dell Inspiration 570, the recording software Adobe Audition 3.0, the sampling frequency is 8 kHz sampling, 16 bit quantization, and saved as WAV audio files. The recording duration is 7 h from 10:30 p.m. to 5:30 a.m. the next morning. In this test experiment, the half hour before the beginning and the end are removed, and 6 h of recording are used. The details of recording for snore sounds were the same as literature (11).

The AHI_{PSG} was the apnea hypopnea index as diagnosed by the gold standard PSG. The severity levels of SAHS were

determined using the AHI value. The subjects with $AHI > 30$, $15 < AHI \leq 30$, $5 \leq AHI \leq 15$, and $AHI < 5$ were classified as severe (S), moderate (M), mild (L) SAHS, and non-SAHS (N), respectively (27).

The 213 subjects were consecutively recruited. In our experiment training phase, the snore episode was cut artificially from the sound of overnight recordings by a non-contact microphone on the bedhead, which included 93 subjects from 213. According to synchronized PSG nocturnal monitoring data, there were two types of snore episodes that were labeled. One was snoring sound labeled snore events by PSG diagnosis, which was only resounding and occurred periodically. The other was a loudly snoring sound appearance behind apnea or hypopnea events labeled by PSG. We called the former a simple snore (SIMP) and the latter a SAHS snore (SAHS). These are shown in Figure 2.

Another 120 subjects from 213 were as a test data set by their overnight recording of sounds. We removed the starting 30 min and the ending 30 min of recording. The remaining 6 h audio signal (11) were used for our test experiments as shown in Table 1.

The Largest Lyapunov Exponent

The largest Lyapunov exponent (LLE) of snores has been calculated to measure the rate of local divergence of nearby trajectories in the state space from dynamical systems theory (28). The LLE of all type snores are shown in Figure 3. The LLE of the four types of simple snore and SAHS snore are all positive. A few of the severe types have the Lyapunov exponent of SAHS snore < 0 , accounting for only 2.4% of the severe type of snoring episodes. The LLE of other types of snores did not appear negative, which also shows that the chaos of snoring is universal. This conclusion is consistent with other researches (13, 15, 29).

The mean of the LLE distribution of simple snore is greater than the LLE distribution of SAHS snore in same severity level. This phenomenon is common in the four types of snore signals. In moderate and severe levels, the difference between the two means of SIMP snore and SAHS snore is increasing. The results reveal that the orbital divergence speed of SIMP snore is greater than that of the SAHS snore, and is consistent with the other study (29). The LLE distribution suggests that unconscious airflow from simple snoring may have more freedom to roam,

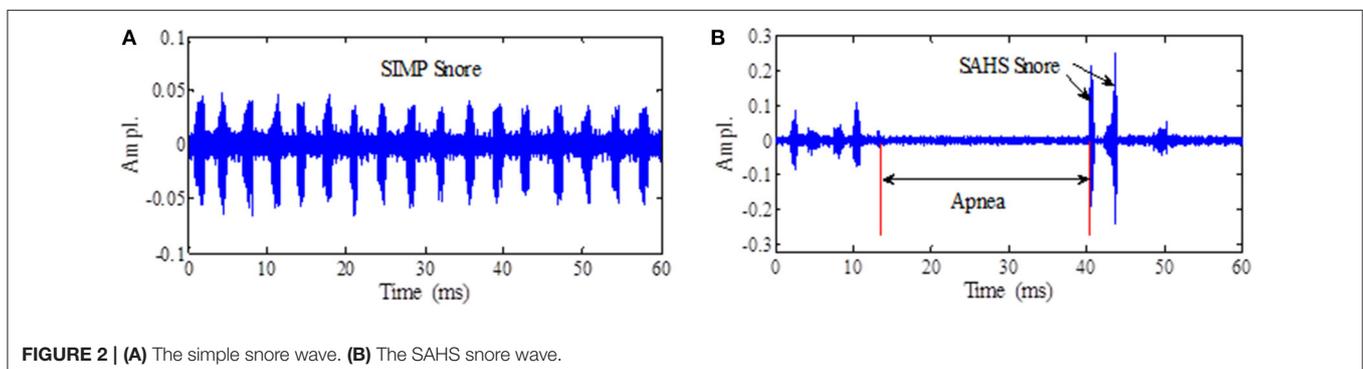
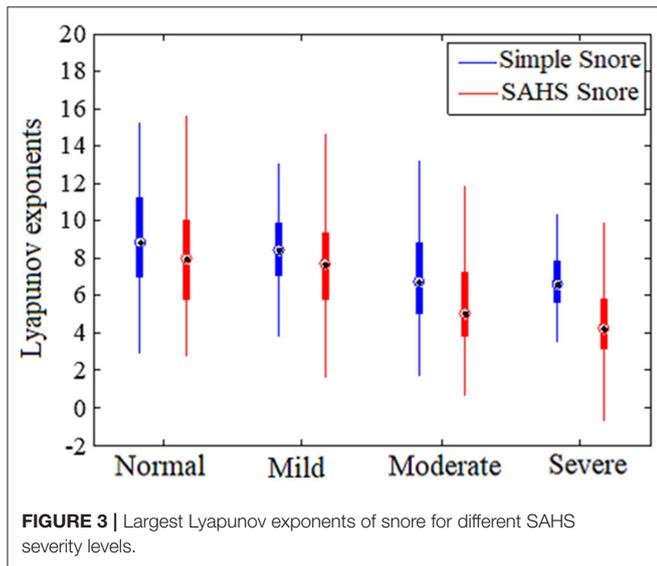


FIGURE 2 | (A) The simple snore wave. (B) The SAHS snore wave.

TABLE 1 | Snore data for training and test.

Training Data				
	N	L	M	S
Subjects (number)	10	23	24	36
Gender (M/F)	9/1	21/2	23/1	36/0
Age (years)	42.1 ± 8.5	46.2 ± 12.4	40.4 ± 13.2	45.6 ± 12.5
AHI _{PSG} (events/h)	2.4 ± 1.4	10.8 ± 3.5	24.5 ± 3.8	57.0 ± 16.8
SIMP Episodes (number)	339	919	480	430
SAHS Episodes (number)	55	376	480	916
Test Data				
Subjects (number)	30	30	30	30
Gender (M/F)	19/11	26/4	25/5	27/3
Age (years)	29.9 ± 8.6	40.7 ± 12.4	43.0 ± 13.8	38.9 ± 11.7
AHI _{PSG} (events/h)	1.9 ± 1.6	9.3 ± 3.0	22.2 ± 4.3	62.5 ± 17.8
Recording length (minutes)	360 × 30	360 × 30	360 × 30	360 × 30



while SAHS snoring may form a certain trend of airflow after being squeezed in the narrow upper airway.

ECD Calculation

According to the illustration in **Figure 1**, the ECD of the snore from the training data in **Table 1** were calculated. The distribution of the ECD vectors in each sub-band of the SIMP snore and SAHS snore of the N, L, M, and S levels, respectively, are shown in **Figures 4A–D**. The ECD vectors distinctively increased with the aggravation of the SAHS severity level in the middle and high-frequency sub-bands. Moreover, the distributions of the ECD vectors were not exactly the same for the SIMP snores and the SAHS snores at the same severity level,

and the ECDs of the SAHS snores were always higher than those of the SIMP snores.

In our study, the SIMP and SAHS snores of four levels (N, L, M, and S) were modeled using the Gaussian Mixture Model (GMM), which formed eight types, including N-simp, N-sahs, L-simp, L-sahs, M-simp, M-sahs, S-simp, and N-sahs. The ECDs of the training data in **Table 1** were extracted to model eight GMMs for the training phase (30), and are showed in **Figure 5**.

RESULTS AND DISCUSSION

Results

Mixture Number of GMM were assigned 2, 12, 12, and 8 for both SIMP and SAHS snore of N, L, M, and S level, respectively. GMM was solved by expectation-maximization (EM) algorithm. Two-fold cross-validation method was employed to evaluate the performance of clustering and classification of GMM regarding training data in **Table 1**. For each type of snore, the rate of being classified as different types is shown in **Figure 6**.

According to the PSG clinical diagnosis definition, AHI is the number of respiration events per hour of sleep. The ECD-calculated AHI is AHI_{ECD} as in Equation (6).

$$AHI_{ECD} = \frac{\text{Number of sleep respiratory events}}{\text{Duration of nocturnal sleep}} \text{ events/h (6)}$$

Figure 5 shows the testing phase, there were another 120 participants for the testing data, which consisted of 30 subjects for each severity level among N, L, M, and S in **Table 1**. Firstly, automatic endpoint detection was performed for snore signals of whole-night recordings to detect the snore sounds (16, 31). Thus, we obtained candidate respiratory events based on the unique rhythm of snores (16, 31). Then, the ECD vectors of these candidate respiratory events were extracted, and we calculated the probabilities of matching with eight GMMs. On the basis of the Bayesian maximum posterior probability rule, the maximum posterior probability winner among the eight GMMs was the snore type. When some snore episodes in a candidate respiratory event were classified as any SAHS snores among N, L, M, and S levels by the GMM, that candidate respiratory event was a true sleep respiratory event. Finally, the AHI_{ECD} score was estimated by the number of sleep respiratory events and the nocturnal sleep duration, as in Equation (6).

In the same way, we extracted another feature set that is MFCC, and estimated the AHI_{MFCC} score also. The MFCC is a classical feature and widely used automatic speech recognition. All experiment results in precision and recall were listed in **Table 2**.

Comparisons of AHI_{ECD}, AHI_{MFCC}, and AHI_{PSG} values of each subject, consistency with the severity of the gold standard PSG diagnosis was correctly screened. There are 120 testers in **Table 2**, including 30 people of four different severity levels. As a result of the MFCC classic feature test, 30 people who are non-SHS can correctly estimate. Twelve of the 30 mild patients were incorrectly classified as non-SHS or moderate SHS types. Eleven of 30 moderate patients were wrongly assorted as mild

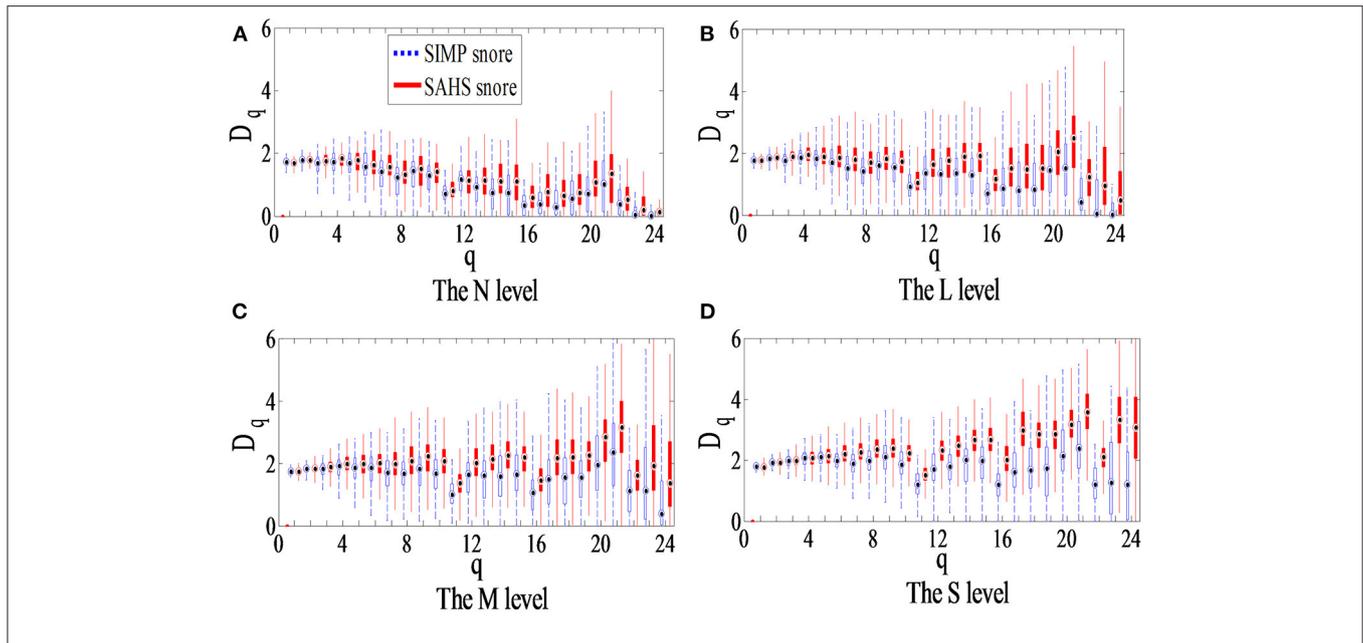


FIGURE 4 | Box plot of ECD vector for different SAHS severity levels. (A) The N level, (B) The L level, (C) The M level, and (D) The S level.

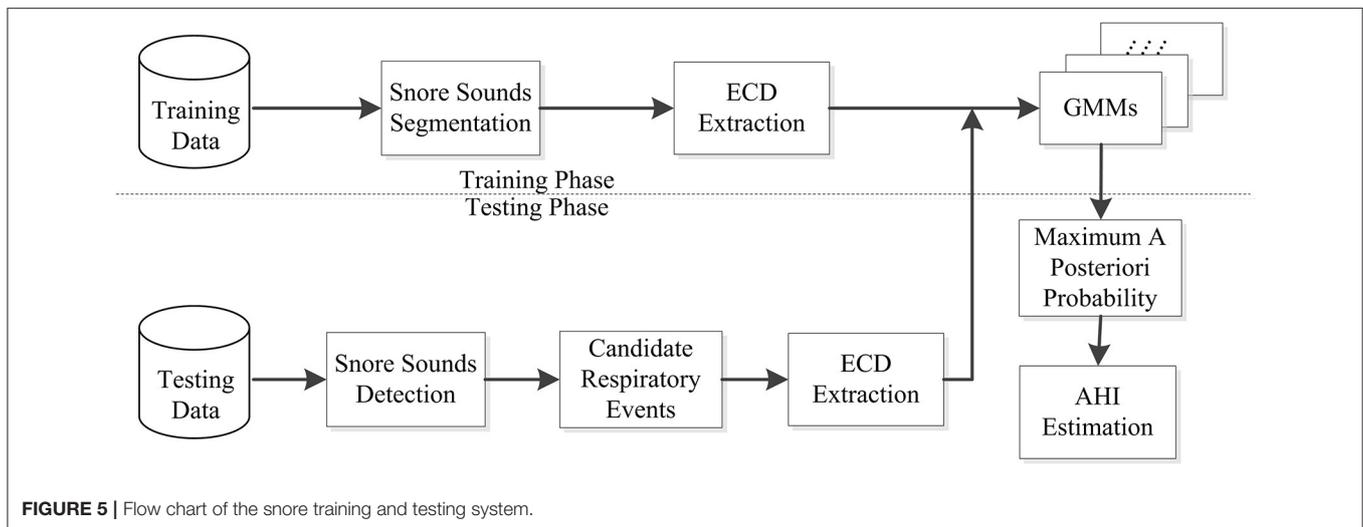


FIGURE 5 | Flow chart of the snore training and testing system.

or severe SAHS types. Thirty patients with severe patients, one of whom were mistakenly estimated as moderate SAHS patients. Compared to AHI_{PSG} in the diagnosis of the SAHS severity level, and the AHI_{MFCC} estimation achieved the mean precision and recall of 79.25 and 80.00%, respectively, as shown in **Table 2**.

As a result of the ECD feature test, two out of 30 non-SAHS people were mistakenly estimated to be mild patients. Thirty patients with mild patients, eight of whom were incorrectly estimated to be non-SAHS or moderate patients. Of the 30 patients with moderate disease, five of them were incorrectly estimated to be mild or severe SAHS patients. Thirty people with

severe patients were correctly estimated. The AHI_{ECD} estimation using our proposed method achieved, respectively, the mean precision and recall of 87.74 and 87.50% compared to AHI_{PSG} in the diagnosis of the SAHS severity level as shown in **Table 2**.

The precision and recall of AHI_{ECD} are higher than AHI_{MFCC} in mild and moderate levels especially.

Comparisons results of AHI_{ECD} and AHI_{MFCC} , both features are good at both ends (i.e., non-SAHS and severe patients). However, for patients with mild SAHS and moderate SAHS, the number of errors by using MFCC is higher than ECD feature. The precision and recall of AHI_{ECD} are higher than AHI_{MFCC} in mild

and moderate level especially. New fractal features achieve better results than classical spectral features. Relative literature (16), this work increased the number of patients in test experiments from 60 to 120 and adds to compares them with the classic feature MFCC. Therefore, the experimental results of this paper are almost the same as the initial experiments, once again confirming the advantages of the new features.

Discussion

Most of the previous studies of detecting snore used the acoustic characteristics of the speech signal of the active pronunciation (4–11), that is limit. Because the snoring contained more breathing sounds than speech. This airflow has more randomness and is generated by passive vocalization. Thereby, we proposed a new feature from the chaos and fractal theory to characterize the irregular extent of snore.

The AHI_{ECD} by new features is closer to clinical diagnosis results than AHI_{MFCC} by conventional parameters. The distribution scatter plot of AHI_{PSG} vs. AHI_{MFCC}, AHI_{ECD} is shown in Figure 7. The black asterisk represents the result

of PSG diagnosis AHI_{PSG}, the green pentagram represents AHI_{MFCC}, and the purple circle represents AHI_{ECD}, and the red dotted line represents the boundary of different severity. The cohen's kappa coefficient of AHI_{ECD} and AHI_{PSG} consistency is 0.833, and AHI_{MFCC} and AHI_{PSG} consistency is 0.733.

The Bland-Altman-plot is depicted in Figure 8. The ordinate represents the difference of AHI_{PSG} and AHI_{ECD} with pinkish, the difference of AHI_{PSG} and AHI_{MFCC} with green. The mean and variance of difference of AHI_{ECD} and AHI_{PSG} were smaller than AHI_{MFCC}. Compared with PSG, 92.50% (111/120) of AHI_{ECD} falls within the consistency limit of 1.96 times variance, higher than 88.33% (106/120) of AHI_{MFCC}. This further suggests AHI_{ECD} estimated by ERB correlation dimensions is more accordance with AHI_{PSG} than AHI_{MFCC}.

In terms of the severity of SAHS, especially the N-type and the severe type, their frequency spectrum has obvious differences. Therefore, the MFCC parameters maintain a good performance for the judgment of these two types. The ECD feature is the same. However, for the intermediate types of mild and moderate, the accuracy of MFCC's outcome drops sharply. The ECD we proposed is much better than MFCC in these two types.

This paper presents a method to measure the degree of disorder of the snoring signal like noise, which were new features called the ECD vectors. The correlation dimensions of the high frequency sub-bands were larger than those of the low-frequency sub-bands in ECD vector. The maximal correlation dimension appeared in the 21st ERB sub-band as shown in Figure 2. This finding suggested that the SAHS snores contain much more irregular and fast-changing components in the high frequency range. The ECD vectors could reveal information that is consistent with the characteristic of the time-domain waveform of the snore. When the upper airway is blocked, the airway becomes shorter. When the upper airway rushes open, the airflow in the narrow area is squeezed, and the turbulent airflow is released. However, the snore spectrum is attenuated to a smaller magnitude in the high frequency range, and thus, it is difficult to give an appropriate description for the mild and moderate level of a snore, which could be too small to distinguish different level. These non-linear methods are expected to provide useful information for better understanding of irregular snoring sounds (13, 14). MFCC includes only magnitude of snore spectral, but our ECD feature completes information in snore sound. When the upper airway is obstructed, the shortening of the airway leads to an increase in the medium and high frequency components, the airflow in the narrow area is squeezed, and some the rapid

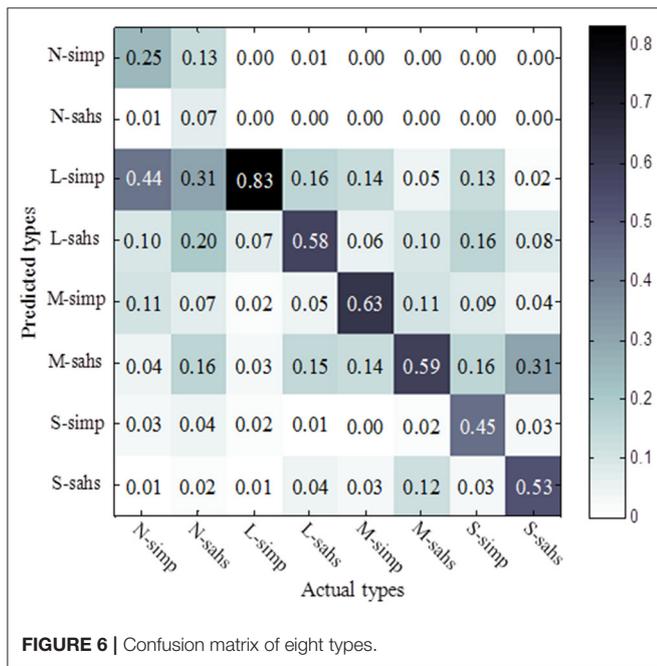


FIGURE 6 | Confusion matrix of eight types.

TABLE 2 | Precision and recall of AHI_{MFCC} and AHI_{ECD} compared AHI_{PSG}.

Levels	N	L	M	S	Total Correct	Mean
Subjects(number)	30	30	30	30	120	
Precision of AHI _{MFCC}	85.71%	64.28%	70.37%	96.66%	96	79.25%
Recall of AHI _{MFCC}	100%	60%	63.33%	96.66%	96	80.00%
Precision of AHI _{ECD}	100%	81.48%	75.75%	93.75%	105	87.74%
Recall of AHI _{ECD}	93.33%	73.33%	83.33%	100%	105	87.50%

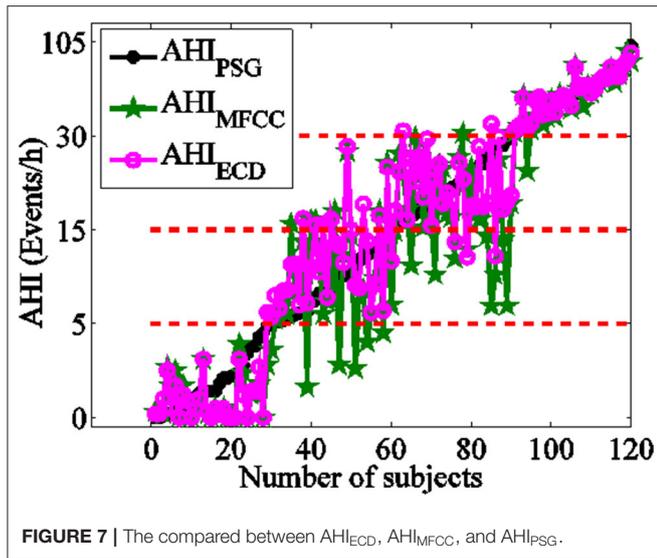


FIGURE 7 | The compared between AHI_{ECD} , AHI_{MFCC} , and AHI_{PSG} .

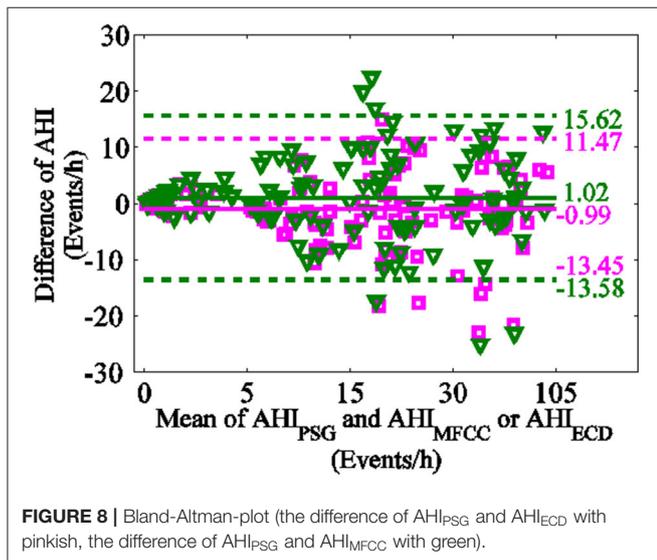


FIGURE 8 | Bland-Altman-plot (the difference of AHI_{PSG} and AHI_{ECD} with pinkish, the difference of AHI_{PSG} and AHI_{MFCC} with green).

change component increases. The Fourier transform shows a characteristic of global decline and local prominence. Compared with MFCC, the ERB enlarged partially and highlighted the anomaly of the mid and high frequency components.

Inspired by the non-linear frequency scale and MFCC characteristics of the Mel spectrum, we use ERB to set the sub-band frequency interval to 8, 4, 2, and 1 ERB bandwidth, so that 3, 6, 12, and 24 subbands are obtained severally in formula (5). The obvious differentiation the snoring sounds of different severity appears when dividing three sub-bands but the details are not enough to distinguish well. As the number of subbands increases, more and more details provide a richer diversity of different severity level of SAHS. According to the distribution of the auditory filter, as it is divided into about 20 subbands in 4 kHz, a set of features is more effective. We adopted one ERB bandwidth and 24 subbands are obtained in formula (5).

No matter how many take the ERB scale, ECD features exhibits SAHS severity is directly proportional to the relationship, that is, the more severe the SAHS, the faster the ECD rises in the middle and high frequency regions, shown as in Figure 4.

However, the calculation of the correlation dimension was time-consuming. This limitation requires us to optimize the algorithm for the correlation dimension. The nature of the correlation dimension on the number of more subbands may need further study.

CONCLUSIONS

Based on the previous experiment, we prove the chaotic nature of snoring sound by the LLE and perfect a new method for estimating the AHI value of SAHS using the correlation dimension vector for snore sounds, which was superior to the conventional spectrum analysis. The ECD vectors might be closely related to the SAHS severity level and reveal the effect of different SAHS severities on the upper airway. The correlation dimension of the sub-bands reveals the inherent information of the mid and high frequencies, while the Fourier transform has its limitations. Chaos provides many quantitative parameters for exploring the nature of this internal information. It could be a study about correlation between fractal dimension and internal physical properties of sleep respiratory sound. There is a positive effect on the development of a medical supplementary diagnosis and in-home healthcare in the internet era.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary materials, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Department of Otolaryngology, Shanghai Jiao Tong University Affiliated Sixth People's Hospital. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

All the co-authors contributed to this work. LH was responsible for all scheme design and about correlation dimension algorithm work of this paper. QP organized the writing work of this paper and MFCC programming work. HY and SY participated on the clinical diagnosis by PSG and actively discussed in this work. DS programmed the EM algorithm work of this paper. XS programmed the result analysis work of this paper.

FUNDING

This study was funded by the Science and Technology Commission of Shanghai Municipality (No. 13441901600) and National Natural Science Foundation of China (Nos. 61525203 and 61572308).

REFERENCES

1. Pevernagie D, Aarts RM, De Meyer M. The acoustics of snoring. *Sleep Med Rev.* (2010) 2:131–44. doi: 10.1016/j.smr.2009.06.002
2. Dafna E, Tarasiuk A, Zigel Y. Automatic detection of whole night snoring events using non-contact microphone. *PLoS ONE.* (2013) 12:e84139. doi: 10.1371/journal.pone.0084139
3. Camacho M, Robertson M, Abdullatif J, Certal V, Kram YA, Ruoff CM, et al. Smartphone apps for snoring. *J Laryngol Otol.* (2015) 10:974–9. doi: 10.1017/S0022215115001978
4. Yang Y, Qin Y, Haung W, Peng H, Xu H. Acoustic characteristics of snoring sound in patients with obstructive sleep apnea hypopnea syndrome. *J Clin Otorhinolaryngol Head Neck Surg.* (2012) 8:360–3.
5. Ng AK, Koh TS, Baey E, Lee TH, Abeyratne UR, Puvanendran K. Could formant frequencies of snore signals be an alternative means for the diagnosis of obstructive sleep apnea? *Sleep Med.* (2008) 8:894–8. doi: 10.1016/j.sleep.2007.07.010
6. Karunajeewa AS, Abeyratne UR, Hukins C. Multi-feature snore sound analysis in obstructive sleep apnea–hypopnea syndrome. *Physiol Meas.* (2011) 1:83–97. doi: 10.1088/0967-3334/32/1/006
7. Solà-Soler J, Fiz JA, Morera J, Jané R. Multiclass classification of subjects with sleep apnoea–hypopnoea syndrome through snoring analysis. *Med Eng Phys.* (2012) 9:1213–20. doi: 10.1016/j.medengphy.2011.12.008
8. Dafna E, Tarasiuk A, Zigel Y. OSA severity assessment based on sleep breathing analysis using ambient microphone. In: *35th Annual International Conference of EMBC.* Osaka:IEEE (2013). p. 2044–7. doi: 10.1109/EMBC.2013.6609933
9. Ben-Israel N, Tarasiuk A, Zigel Y. Obstructive apnea hypopnea index estimation by analysis of nocturnal snoring signals in adults. *Sleep.* (2012) 9:1299–305. doi: 10.5665/sleep.2092
10. Herath DL, Abeyratne UR, Hukins C. Hidden Markov modelling of intra-snore episode behavior of acoustic characteristics of obstructive sleep apnea patients. *Physiol Meas.* (2015) 12:2379–404. doi: 10.1088/0967-3334/36/12/2379
11. Xu H, Wei S, Yi H, Hou L, Zhang C, Chen B, et al. Nocturnal snoring sound analysis in the diagnosis of obstructive sleep apnea in the Chinese Han population. *Sleep Breath.* (2015) 2:599–605. doi: 10.1007/s11325-014-1055-0
12. Azarbarzin A, Moussavi Z. Nonlinear properties of snoring sounds. *Int Conf Acoust.* (2011) 4316–9. doi: 10.1109/ICASSP.2011.5947308
13. Sakakura A. Acoustic analysis of snoring sounds with chaos theory. *Int Congr Ser.* (2003) 1257:227–30. doi: 10.1016/S0531-5131(03)01170-1
14. Mikami T. Detecting nonlinear properties of snoring sounds for sleep apnea diagnosis. In: *2nd International Conference on Bioinformatics and Biomedical Engineering.* Shanghai:IEEE (2008). p. 173–6. doi: 10.1109/ICBBE.2008.621
15. Ankişhan H, Yilmaz D. Comparison of SVM and ANFIS for snore related sounds classification by using the largest Lyapunov exponent and entropy. *Comput Math Methods Med.* (2013) 2013:238937. doi: 10.1155/2013/238937
16. Hou LM, Shi D, Liu HC, Zhang WT. Screening of SAHS snore based on ERB correlation dimension. *J Appl Sci.* (2017) 2:181–92. doi: 10.3969/j.issn.0255-8297.2017.02.005
17. Grassberger P, Procaccia I. Dimension and entropy of strange attractors from a fluctuating dynamic approach. *Phys D.* (1984) 1–2:34–54. doi: 10.1016/0167-2789(84)90269-0
18. Hekmatmanesh A, Micaeli M, Sadeghniai-Haghighi K, Wu HP, Nazeran H. Sleep spindle detection and prediction using a mixture of time series and chaotic features. *Adv Electr Electron Eng.* (2017) 3:435–47. doi: 10.15598/aeec.v15i3.2174
19. Subramaniam K, Clark AR, Hoffman EA, Tawhai MH. Metrics of lung tissue heterogeneity depend on BMI but not age. *J Appl Physiol.* (2018) 2:328–39. doi: 10.1152/jappphysiol.00510.2016
20. Hou LM, Deng DC, Wang SZ. Improvement of speaker identification performance using nonlinear feature. *Patt Recog Artif Intell.* (2006) 6:776–81. doi: 10.1360/crad20061223
21. Takens F. Detecting strange attractors in turbulence in dynamical systems and turbulence, Warwick (1980). *Lect Notes Math.* (1981) 898:366–81. doi: 10.1007/BFb0091924
22. Rosenstein MT, Collins JJ, De Luca CJ. Reconstruction expansion as a geometry-based framework for choosing proper delay times. *Phys D.* (1994) 1–2:82–98. doi: 10.1016/0167-2789(94)90226-7
23. Cao L. Determining minimum embedding dimension from scalar time series. *Springer US.* (2002) 2:43–60. doi: 10.1007/978-1-4615-0931-8_3
24. Grassberger P, Procaccia I. Measuring the strangeness of strange attractors. *Phys D.* (1983) 1–2:189–208. doi: 10.1016/0167-2789(83)90298-1
25. Moore BC, Glasberg BR. Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *J Acoust Soc Am.* (1983) 3:750–3. doi: 10.1121/1.389861
26. Glasberg BR, Moore BCJ. Derivation of auditory filter shapes from notched-noise data. *Hear Res.* (1990) 1–2:103–38. doi: 10.1016/0378-5955(90)90170-T
27. Editorial Board of Chinese Journal of Ot, Subspecialty Group of Pharyngology. Guideline for the diagnosis and surgical treatment of obstructive sleep apnea hypopnea syndrome. *Chin J Otorhinolaryngol Head Neck Surg.* (2009) 2:95–6. doi: 10.3760/cma.j.issn.1673-0806.2009.02.003
28. Rosenstein MT, Collins JJ, De Luca CJ. A practical method for calculating largest Lyapunov exponents from small data sets. *Phys D.* (1993) 65:117–34. doi: 10.1016/0167-2789(93)90009-P
29. Kizilkaya M, Ari F, Demirgunes DD. Analysis of snore sounds by using the largest Lyapunov exponent. *J Concr Appl Math.* (2011) 9:146–53.
30. Nakamura K, Toda T, Saruwatari H, Shikano K. Speaking-aid systems using GMM-based voice conversion for electrolaryngeal speech. *Speech Commun.* (2012) 1:134–46. doi: 10.1016/j.specom.2011.07.007
31. Hou LM, Zhang CH, Yin SK, Yi HL, Meng LL. *A Device of Screening OSAHS Based on Recording Snore Sound.* Chinese Patent CN103735267A. Amsterdam: Elsevier Science (2014).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Hou, Pan, Yi, Shi, Shi and Yin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.