Obstructive Sleep Apnea Detection Using Speech Signals with High Frequency Components

Kang-Gao Pang[1], Tai-Chiu Hsung[2]*, Guozhao Liao[1], Wing-Kuen Ling[1], Alex Ka-Wing Law[3], and Wing Shan Choi[3]
[1] Guangdong University of Technology, Guangzhou and 510006, China
[2] Chu Hai College of Higher Education, Hong Kong, China
[3] The University of Hong Kong, Hong Kong, China
Email: pangkanggao@163.com; richardhsung@chuhai.edu.hk; 2111903022@mail2.gdut.edu.cn; yongquanling@gdut.edu.cn; alex1209@connect.hku.hk; drwchoi@hku.hk

Abstract—In this study, the Obstructive Sleep Apnea (OSA) detection using speech signals during awake is considered. Traditional speech based OSA detection methods adopt traditional features (Formants, MFCC, etc.) on normal speech frequency range (<6kHz). However, it ignores the signal components outside this range that usually appear in pathological voices. In this paper, higher order traditional speech features (with more high frequency components) are adopted for detection. To better characterize OSA patients' speech, a high frequency feature set is proposed. It consists of the traditional speech features with optimized parameters and a new proposed feature: High frequency energy. Principal Component Analysis (PCA) based Sequence Forward Feature Selection (PCASFFS) are adopted as feature selection. In the simulation using 66 OSA patients' speech signals, it achieves an accuracy of 84.85% for multi-class (4 levels) detection with the proposed high frequency feature set using quadratic discriminant analysis classifier (QDA).

Index Terms—Obstructive sleep apnea, speech analysis, High Frequency Energy (HFE), machine learning

I. Introduction

Obstructive sleep apnea (OSA) is a prevalent sleep-related respiratory disease characterized by partial or total repetitive obstruction of the upper respiratory tract during sleep. The population it affects is mainly obese adult males, but it can also affect men and women of all ages and weights. Nowadays, the prevalence of OSA in adults is approximately 3% to 7% according to epidemiological data, and most OSA patients remain in an undiagnosed stage [1]-[2]. As the severity of a patient's OSA increases, his/her daytime sleepiness increases, which greatly affects their quality of life. At the same time, OSA patients are also prone to car accidents due to poor attention caused by poor sleep quality. On the other hand, OSA is also a risk factor for chronic diseases, such as hypertension, cardiovascular, congestive heart diseases and so on [3]-[7].

Nowadays, the gold standard for diagnosing OSA is Polysomnography (PSG). Several biosignals are recorded overnight including Electroencephalography (EEG), nasal airflow, Peripheral oxygen saturation (SpO2), abdominal and thoracic effort signals. They are then used to detect the average apnea and hypopnea events per hour during sleep hence Apnea hypopnea index (AHI). The OSA severity is classified into four levels according to AHI (AASM [8]), AHI<5 is considered as normal, 5⩽AHI<15 is considered as mild, 15⩽AHI<30 is considered as moderate, AHI⩾30 as severe. To perform PSG, it needs direct or allied health care professionals to hook up the sensors. The patient is required to sleep at least for 6 hours to obtain a reliable result. Moreover, some patients cannot tolerate the sensors after hooking up so cannot have enough sleep time during PSG test. Therefore, a more convenient, comfortable and inexpensive alternative way to measure OSA severity is desired for long.

One direction for estimating OSA severity is to use the speech signal of patients during awake. Many studies have shown that OSA is highly correlated with the anatomical and functional abnormalities [9], [10] of the upper airways, which may affect the acoustic characteristics of patient's speech. Many approaches have been proposed to predict the severity of OSA using speech analysis since 1980's. The script for patient to speak in the detection is sustained vowels recorded in 44.1kHz sampling rate. The adopted acoustic features are Linear Prediction Coefficients (LPC), formants, formants derivative, Formant bandwidth, Harmonic to noise ratio (HNR), Mel-frequency cepstral coefficients (MFCC), delta MFCC (ΔMFCC), ΔΔMFCC, log energy, jitter, shimmer and fundamental frequency [11]. Zigel *et al*. [12] used LPC, formants, MFCC features etc and achieves binary classification accuracy of 92.3% for 26 subjects. Kriboy *et al*. [13] found sustained vowel recordings in different body positions can highlight acoustic differences between OSA and non-OSA subjects. Kriboy *et al*. [14] further develop an OSA detection system based on this founding and it achieves binary classification accuracy of 84.67% for 35 subjects. Ben *et al*. [15] proposed a system which fuses short-term, long-term and sustained vowels acoustic features to estimate AHI and achieves diagnostic agreement of 67.3% between the

Manuscript received June 10, 2021; revised December 21, 2021.
This work was supported by the Guangdong Provincial Key Laboratory of Intellectual Property and Big Data under Grant 2018B030322016.
Corresponding author email: richardhsung@chuhai.edu.hk.
doi:10.12720/jcm.17.1.49-55

speech-estimated AHI and the PSG-determined AHI for 198 subjects. Simply *et al*. [16] focus on analyzing breathing sounds to estimate OSA severity. The acoustic features adopted are MFCC, pitch peak, speech energy, average zero crossing rate (ZCR), noise energy, and achieves accuracy 76.5% for 90 subjects. Goldshtein *et al*. [17] proposed to use long-term feature set (28) and short-term feature set (100). The system achieves binary classification specificity and sensitivity of 83% and 79% for male OSA patients, and 86% and 84% for female OSA patients, respectively (total 96 subjects). The methods in [12], [13], [16], [17] distinguish OSA and non-OSA subjects whereas the methods proposed in [14], [15] estimate AHI value and the best achieved diagnostic agreement is 67.3%.

From the data collection collected by the School of Dentistry of the University of Hong Kong, it is found that the voice and speech of OSA patients show significantly different high-frequency spectrums at different degrees of severity. We propose fine-tuning the traditional features designed to work in the normal frequency range to work in the extended frequency range. The feature set in the normal frequency range proposed in [17] is adopted as the baseline. Some features are selected for the analysis of high-frequency components, such as the high-order coefficient of MFCC, the high-order formants and etc. We further propose a new feature dubbed as High Frequency Energy (HFE) for better characterizing OSA patients' speech.

In this paper, we study the high frequency components of speech signal for OSA severity estimation which is not considered in the previous studies. Our work focuses on classifying OSA patients into more detailed categories (normal, mild, moderate and severe). Multi-class OSA detection [18], [19] is better than binary detection. The latter cannot tell whether the severity is only mild or moderate to severe. OSA patients can see the urgency to seek for OSA treatment with Multiclass OSA detection. Moderate to severe OSA patients exhibit higher risks to have heart diseases, high blood pressure, stroke, depression, etc. From [19], patients with severe OSA are more likely to cause brain damage, where the integrity of matt fibers in many brain areas were significantly reduced. For moderate to severe OSA patients who underwent oral appliance treatment, frequent follow-ups are needed to determine whether there is a good clinical response to ensure the treatment effect. In next section, we will describe the materials and method, and the proposed detection system: signal pre-processing, feature extraction, feature selection and classification. In Section III, we will test the performance of the proposed system, and compare with the system using Goldshtein's feature sets [17]. Finally, we will make a conclusion in Section IV.

II. Material and Methods

A. Data Collection

Voice data of OSA patients of the Faculty of Dentistry, The University of Hong Kong (HKU) were collected along with overnight PSG studies. The collection of data was approved by the Institutional review board of HKU (UW11-122). Patients are requested to produce some sounds containing normal and sustained vowels (/a/, /e/, /i/, /o/, / u/). Vowels were used because they do not depend on language and speech content. The subjects were recruited from the patients with OSA which may undergo corrective jaw surgery, as shown in Table I. Each patient's OSA severity changed over time and their corresponding PSG studies at different time points (before and after surgery) are regarded as independent. Thirty-seven subjects with total 66 sets of data are collected. The voice data were recorded just before the PSG study with a high-fidelity recorder (Roland R-44 digital recorder, sampling rate was 96 kHz). Characteristics of the recordings is summarized in 0The data were divided into 4 groups according to the AHI value: normal group ($0 \leqslant AHI < 5$), mild group ($5 \leqslant AHI < 15$), moderate group ($15 \leqslant AHI < 30$), and severe group ($AHI \geqslant 30$).

TABLE I: CHARACTERISTICS OF DATABASE

Group	NO. of sample	AHI
Normal	31	2.45±2.45(0-4.9)
Mild	13	9.1±3.8(5.3-12.9)
Moderate	10	20.1±4.8(15.3-24.9)
Severe	12	63.6±25.1(38.5-89.7)

B. Feature Extraction

For the feature extraction, traditional speech features proposed in Goldshtein's study [17] were adopted as baseline as shown in Table II. Based on the observation of high frequency components in spectrogram, the parameters of speech features were modified to operate in the entire frequency range. A new feature is proposed to capture more high-frequency characteristics named **"High frequency feature set".**

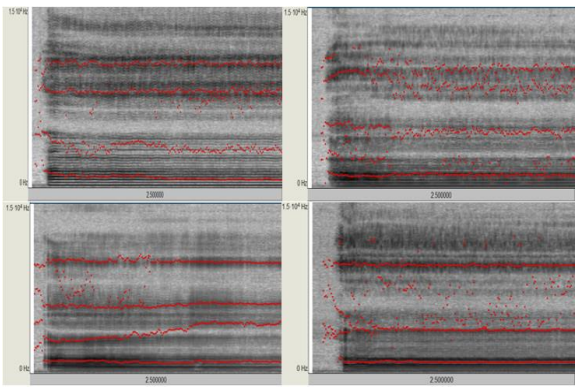
TABLE II: TRADITIONAL SPEECH FEATURES SET

#	Feature name	No. of features	Feature symbol
1	Fundamental frequency	1	F_0
2	Jitter	1	Jitt
3	Shimmer	1	Shimm
4	Vocal tract length	1	VTL
5	Harmonic to noise ratio	1	HNR
6	Formants	3	F_1-F_3
7	Formant derivative (F_2-F_1, F_3-F_2)	2	DF_{12}-DF_{23}
8	Formant band width	3	BW_1- BW_3
9	Long-term prediction	18	a_1-a_{18}
10	ARMA model	18	b_1-b_4, a_1-a_{14}
11	Mel-frequency cepstral coef.(MFCC)	16	c_1-c_{16}
12	ΔMFCC	16	Δc_1-Δc_{16}
13	ΔΔMFCC	16	$\Delta\Delta c_1$-$\Delta\Delta c_{16}$
14	Energy	1	E
15	ΔEnergy	1	ΔE
16	ΔΔEnergy	1	ΔΔE

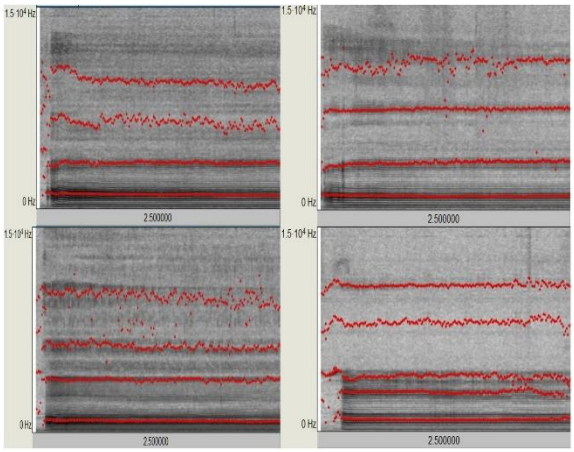
The recorded sound data were firstly pre-processed including segmentation, DC removal, and pre-emphasizing. Each fragment of a single voice (such as /a/, /e/,…) was windowed into a 40ms frame with 50% overlap using Hamming window in calculating 103 features for OSA detection, as shown in Table III. According to previous studies [9], [10], OSA is mainly related to anatomical abnormalities, so the feature set is selected which can show the difference between physiology and perception between OSA patients and normal people. For example, shapes of the vocal tract of OSA patients are different from that of normal person, which can be shown in Linear Prediction Coefficients (LPC), which can represent the changes envelope of the speech; At the same time, formant can be affected by the vocal tract shape; formant bandwidth (BW) can represent the changes in soft tissue characteristic; changes in harmonic noise ratio (HNR) can help us assess phonetic diseases [20]; MFCC can represent the signal spectrum in a perceptual aspect [21].

TABLE III: HIGH FREQUENCY FEATURES SET FOR OSA DETECTION

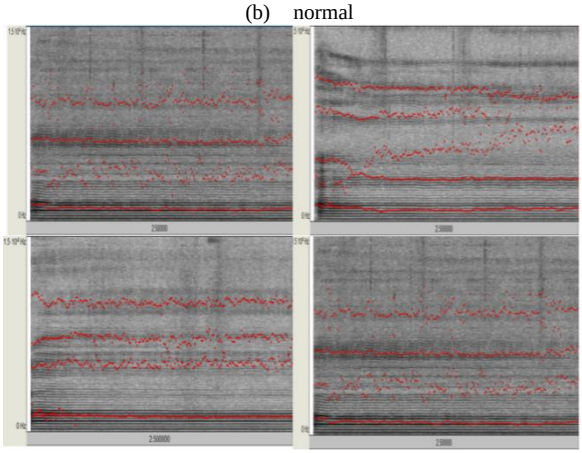
#	Feature name	No. of features	Feature symbol	Statistical value
1	Formant	5	F_1-F_5	mean, std
2	Formant derivative	4	ΔF_1-ΔF_4	mean
3	Formant bandwidth	5	BW_1-BW_5	mean, std
4	Mel-Frequency Cepstral Coef.(MFCC)	18	c_1-c_{18}	mean
5	ΔMFCC	18	Δc_1-Δc_{18}	mean
6	ΔΔMFCC	18	$\Delta\Delta c_1$-$\Delta\Delta c_{18}$	mean
7	Linear-Prediction Coef.(LPC)	18	a_1-a_{18}	whole
8	Energy	1	E	whole
9	Fundamental frequency	1	F_0	mean, std
10	Harmonic to noise ratio	1	HNR	mean, std
11	Log Energy	1	loc	mean
12	High Frequency Energy	1	HFE	whole



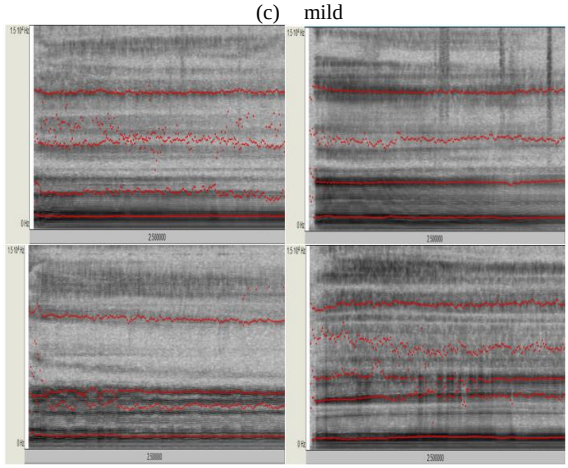
(a) Severe



(b) normal



(c) mild



(d) moderate

Fig. 1. Spectrogram (generated using Praat [22]) of /a/ sound of OSA patients with various severity levels

In previous studies [12]-[17], voice data were usually collected in 44.1kHz sampling rate. It was usually further down-sampled to lower sampling rate below 16kHz for efficient processing. Traditional speech processing methods operate below 5-6kHz as the maximum frequency of ordinary people's voices is usually below 6kHz. The characteristics of high frequency were therefore ignored in previous studies. However, we can clearly see voices of OSA patient exhibited different characteristics in different severity levels at frequency higher than 6kHz from the spectrogram, as shown in Fig. 1. In the figure, four patients' sustained /a/ sounds (target 5sec) for each severity levels

were illustrated. The recording sampling rate was 96kHz and the signals were resampled to 24kHz. The display duration was set to 2.5 sec for all the sounds in the figures. Order 5 formant tracking with frequency limit of 12kHz applied. It can be observed that, as the severity increases, there are more mixed formant peaks with varying bandwidth. The detected peaks are usually fluctuating during sound production. In contrast, normal people's /a/ sounds exhibit clearer and more stable formant tracks. We conjecture that sleep apnea events induce the sound production organ to produce more high frequency components during awake, which may hence be detected from OSA patients' voices. Traditionally, lower order formants (F_1- F_3), lower order MFCC (16) are adopted. In this paper, higher order formants (F_1-F_5) and higher order MFCC were included to reveal more high-frequency characteristics. Furthermore, "ratio of high frequency energy to total energy" was proposed as a new feature. The target frequency band was from 5kHz to 10kHz. Energy ratio was adopted because it is invariant to inter- and intra-patient voice energy variations.

As the input signal may change its frequency and cross the 5kHz boundary, it is desirable for the calculation of high frequency components to be adaptive. Hence, Hilbert Huang transform (HHT) [23] was adopted. Firstly, ten intrinsic mode function (IMF) of input signal s were computed, and then Hilbert transform was applied to get the corresponding instantaneous frequency: the amplitude function $a_g(n,k)$, and the frequency function $f_g(n,k)$.

The proposed High frequency energy ratio (HFE) is defined as:

$$g(n,k) = \{f_g(n,k), a_g(n,k): \quad for\ n = 0,1,\dots,N; k = 1,\dots,10\} \quad (1)$$

where $g(n,k)$ is the Hilbert transform of kth IMF of input signal s at abscissa n;

$$\mathrm{a}'_{\mathrm{g}} = \begin{cases} a_g(n,k) & for \quad f_g(n,k) > 5kHz \\ 0 & otherwise \end{cases} \quad (2)$$

$$\mathrm{E}_{5\mathrm{k}} = \sum\nolimits_{n,k} a'_g(n,k) \quad for\ n = 0,1,\dots,N; k = 1,\dots,10 \quad (3)$$

$$\mathrm{E}_{\mathrm{all}} = \sum\nolimits_{n,k} a_g(n,k) \quad for\ n = 0,1,\dots,N; k = 1,\dots,10 \quad (4)$$

$$\mathrm{HFE} = \frac{E_{5k}}{E_{all}} \quad (5)$$

Along with the traditional speech features, the adopted features in this study are shown in Table III, which are categorized into the following 3 groups:

1. Time domain features: Log Energy
2. Cepstral features: Mel-Frequency Cepstral Coefficients (MFCC); ΔMFCC; ΔΔMFCC
3. Spectral features: Formant; Bandwidth (BW); Linear Prediction Coefficients (LPC); Harmonic noise ratio (HNR); High Frequency Energy (HFE)

The voice signal was framed into 40ms blocks with 50% overlap using Hamming window. Signal features were calculated for each frame, except the #7, #8 and #12 features. The statistical values of the signal features are calculated: mean and/or standard deviation (std). And they are used to compose the final feature vector $[x_i^n]$, where i, n is the number of the subject and feature index, respectively. The total number of elements was 103. The feature vector was then used for severity classification.

C. Feature Selection

To avoid over-fitting in training the classifier [24], the dimensions of the feature vector need to be reduced. In previous studies [12]-[17], single feature selections, which are to use the original feature set for feature selection, were considered, such as SFFS. In this paper, a different feature selection approach was adopted to optimize the classifier. We designed a feature selection method - Principal Component Analysis (PCA) based Sequential Forward Feature Selection (PCASFFS): PCA along with sequential forward feature selection (SFFS). PCA was firstly applied on the original feature set. It centered the feature vectors and then used singular value decomposition (SVD) to find the loadings for getting a reduced dimension feature set [25]. Then, SFFS was applied to select features from the dominant PC feature set. SFFS consists of three steps: 1. inclusion, 2. testing, and 3. exclusion. In the first step, SFFS will select the best performing features to include, and then add the remaining features one by one to test whether improves. If the feature under test improves the performance, it will be added. If there is no improvement in performance, it is excluded.

D. Multi-Class OSA detection

After feature selection, the severity can be predicted from using a classifier. Quadratic Discriminant Analysis Classifier (QDA) and Naive Bayes Classifier (NBC) [26] were selected in the experiment for predicting four severity levels: normal, mild, moderate and severe. The performances of classifier combinations were assessed by using K-fold cross-validation and leave-one-out (LOO) methods [27]. The classification error L is defined as:

$$\mathrm{L} = \sum_{j=1}^{n} w_j I\left\{y'_j \neq y_j\right\} \quad (6)$$

which is a weighted fraction of misclassified observations where y_j , y'_j, n are the ground truth severity level, predicted severity level (with the maximal posterior probability) and the sample size respectively; $I\{.\}$ is the indicator function. The weight for observation j is w_j according to the corresponding prior class probability and they are normalized so they sum to 1, i.e. $\sum_{j=1}^{n} w_j = 1$.

III. Expriments

In the experiment, it is interesting to see which combination of feature sets, feature selection and classifier

delivers the best performance. We tested the proposed systems with 66 sets of OSA patients' vowels (/a/, /e/, /i/, /o/, /u/). As shown in Fig. 2, the input speech vowel /a/ was firstly preprocessed for features extractions. Then, the features were selected and used to predict the OSA severity with a classifier.

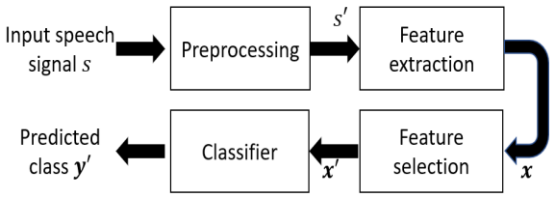


Fig. 2. OSA detection system block diagram.

The features in Goldshtein's study [17] were adopted for multi-class OSA detection. The best achieved accuracy was 80.3% while most of the other results are around 72%. It was regarded as baseline for comparing to various combinations of feature sets, feature selection and classifiers, as follows:

1. For the feature selection: basic SFFS and PCA+SFFS which is cascaded PCA and SFFS.
2. For the classifier: Naive Bayes classifier (NBC), quadratic discriminant analysis classifier (QDA). K-fold and LOO were adopted both in training and testing.
3. For the feature set: a) traditional feature set; b) high frequency feature set and with HFE and c) high frequency feature set without HFE.

The classification errors of the proposed system and traditional feature set with different model settings are shown in Table IV and Table V, respectively. From the tables, we can see that the performance of the high-frequency feature set without HFE is better than the traditional feature set. The performance of the high-frequency feature set with HFE is better than the high-frequency feature set without HFE. From the table, it can be seen that additional feature selection mode of PCA performed better than the single feature mode selection. The classification performance of the second QDA was better than that of NBC. It is also shown that the newly added feature HFE improved performance in all the combination for the classification tasks.

TABLE IV: MULTI-CLASS OSA DETECTION ERROR *L* WITH HIGH FREQUENCY FEATURE SET

Feature selection method+classifier		Validation			
		Resubstitution	***LOO***	***K-fold***	
				K=5	***K=10***
without HFE	SFFS+NBC	0.3333	0.4545	0.4545	0.4394
	PCA+SFFS+NBC	0.1515	0.2121	0.2121	0.197
	SFFS+QDA	0.2878	0.4091	0.4545	0.4091
	PCA+SFFS+QDA	0.1212	0.2121	0.2271	0.2273
WithHFE	SFFS+NBC	0.2727	0.3636	0.4091	0.3636
	PCA+SFFS+NBC	0.1515	0.1667	0.1818	0.1818
	SFFS+QDA	0.2878	0.3636	0.3485	0.3788
	PCA+SFFS+QDA	**0.106**	**0.1667**	**0.1515**	**0.1818**

IV. CONCLUSION AND FUTURE WORK

In this experiment, we attempt to achieve multi-class OSA detection using speech. Parameters of traditional speech features [17] are modified to operate in the entire frequency range for better detect OSA. Higher order formants and MFCC, high frequency energy-HFE are adopted in the proposed system. Then, we look for the best combination of feature selection models (PCA+SFFS) and classifier for the multi-class OSA detection. Using a quadratic discriminant analysis classifier, it achieved classification accuracy of 84.85%. The results of this paper show that the proposed feature (including HFE) set improve the speech based OSA detection. It shows potential to be a screening tool for OSA patients.

In the current preliminary study, 66 sets of OSA patients' voice data were collected. Only sustained vowels are used as input for the detection. We are studying on using more sustained voices and continuous speech, which may further reflect the characteristics of vocal tract. It is planned to collect further data from OSA patients to further study the voice features for better OSA detection. On the other hand, it is also interesting to look for more specific clinical correlations of speech features to OSA severity, such as before and after jaw surgery.

TABLE V: MULTI-CLASS OSA DETECTION ERROR *L* WITH HIGH FREQUENCY FEATURE SET

Feature selection method+classifier	Validation			
	Resubstitution	***LOO***	***K-fold***	
			K=5	***K=10***
PCA+SFFS+NBC	0.1212	0.197	0.2727	0.2424
PCA+SFFS+QDA	0.197	0.2879	0.2879	0.2879

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

This study was designed and coordinated; data was collected by Tai-Chiu Hsung and Wing Shan Choi. Kang-Gao Pang performed data preprocessing, design and develop the main algorithm. All authors performed data analysis, read and approved the final manuscript.

REFERENCES

[1] N. M. Punjabi, "The epidemiology of adult obstructive sleep apnea," *Proceedings of the American Thoracic Society*, vol. 5, no. 2, pp. 136-143, 2008.

[2] K. P. Pang and D. J. Terris, "Screening for obstructive sleep apnea: an evidence-based analysis," *American Journal of Otolaryngology*, vol. 27, no. 2, pp. 112-118, 2006.

[3] T. Young, P. E. Peppard, and D. J. Gottlieb, "Epidemiology of obstructive sleep apnea," *Amer. J. Respir. Crit. Care Med.*, vol. 165, no. 9, pp. 1217–1239, 2002.

[4] A. Tarasiuk, S. Greenberg-Dotan, T. Simon, A. Tal, A. Oksenberg, and H. Reuveni, "Low socioeconomic status is a risk factor for cardiovascular disease among adult

obstructive sleep apnea syndrome patients requiring treatment," *Chest*, vol. 130, no. 3, pp. 766–773, 2006.

[5] J. M. Parish and V. K. Somers, "Obstructive sleep apnea and cardiovascular disease," *Mayo Clinic Proceedings*, vol. 79, no. 8, pp. 1036-1046, 2004.

[6] S. Redlins and K. P. Strohl, "Recognition and consequences of obstructive sleep apnea hypopnea syndrome," *Clinics in Chest Medicine*, vol. 19, no. 1, pp. 1–19, Mar. 1998.

[7] G. Coccagna, A. Pollini, and F. Provini, "Cardiovascular disorders and obstructive sleep apnea syndrome," *Clinical and Experimental Hypertension*, vol. 28, no. 3–4, pp. 217–224, 2006.

[8] R. B. Berry, R. Brooks, C. E. Gamaldo, *et al*., "American academy of sleep medicine," *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*, Version 2.4. Darien, IL: American Academy of Sleep Medicine, 2017.

[9] P. K. Monoson and A. W. Fox, "Preliminary observation of speech disorder in obstructive and mixed sleep apnea," *Chest*, vol. 92, no. 4, pp. 670–675, 1987.

[10] A. W. Fox, P. K. Monoson, and C. D. Morgan, "Speech dysfunction of obstructive sleep apnea. A discriminant analysis of its descriptors," *Chest*, vol. 96, no. 3, pp. 589–595, 1989.

[11] R. R. W. Rabiner and Schafer, *Theory and Applications of Digital Speech Processing*, Upper Saddle River, Pearson, 2011.

[12] Y. Zigel, A. Tarasiuk, and E. Goldshtein, "Analysis of speech signals among obstructive sleep apnea patients," in *Proc. IEEE 25th Convention of Electrical and Electronics Engineers in Israel*, 2008, pp. 760-764.

[13] M. Kriboy, A. Tarasiuk, and Y. Zigel, "Detection of Obstructive sleep apnea in awake subjects by exploiting body posture effects on the speech signal," in *Proc. 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014, pp. 4224-4227.

[14] M. Kriboy, A. Tarasiuk, and Y. Zigel, "A novel method for obstructive sleep apnea severity estimation using speech signals," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 3606-3610.

[15] D. B. Or, E. Dafna, A. Tarasiuk, and Y. Zigel, "Obstructive sleep apnea severity estimation: Fusion of speech-based systems," in *Proc. 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016, pp. 3207-3210.

[16] R. M. Simply, E. Dafna, and Y. Zigel, "Obstructive Sleep Apnea (OSA) classification using analysis of breathing sounds during speech," in *Proc. 26th European Signal Processing Conference (EUSIPCO)*, Rome, 2018, pp. 1132-1136.

[17] E. Goldshtein, A. Tarasiuk, and Y. Zigel, "Automatic detection of obstructive sleep apnea using speech signals," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 5, pp. 1373-1382, May 2011.

[18] C. A. Kushida, M. R. Littner, T. Morgenthaler, *et al*., "Practice parameters for the indications for polysomnography and related procedures: An update for 2005," *SLEEP*, vol. 28, no. 4, pp. 499-523, 2005.

[19] V. Castronovo, P. Scifo, A. Castellano, M. S. Aloia, *et al*., "White matter integrity in obstructive sleep apnea before and after treatment," *SLEEP*, vol. 37, no. 9, pp. 1465-1475, 2014.

[20] G. D. Krom, "A cepstrum-based technique for determining a harmonics-to-noise ratio in speech signals," *Journal of Speech, Language, and Hearing Research*, vol. 36, no. 2, pp. 254–266, 1993.

[21] L. Rabiner and B. H. Juang, *Fundamentals of Speech Recognition*. *Englewood Cliffs*, NJ: Prentice-Hall., 1993.

[22] P. Boersma, "Praat, a system for doing phonetics by computer," *Glot International*, vol. 5, no. 9-10, pp. 341-345, 2001.

[23] E. H. Norden, S. Zheng, R. L. Steven, C. W. Manli, *et al*., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. Lond. A*., vol. 454, pp. 903–995, 1998.

[24] H. Liu and H. Motoda, *Feature Selection for Knowledge Discovery and Data Mining*, Massachusetts: Springer, 1998.

[25] Y. Luo, S. Xiong, and S. Wang, "A PCA based unsupervised feature selection algorithm," in *Proc. Second International Conference on Genetic and Evolutionary Computing*, 2008, pp. 299-302.

[26] W. Xindong, *et al*., "Top 10 algorithms in data mining," *Knowledge and Information Systems*, vol. 14, pp. 1-37, 2008.

[27] A. J. Smola, P. Bartlett, B. Schölkopf, and D. Schuurmans, *Advances in Large-Margin Classifiers*, MIT Press, 2000.

Kang-Gao Pang was born in Guangdong Province, China, in 1996. He received the master degree from the University of Guangdong university of Technology, in 2021. His research interests include machine learning, digital signal processing, and speech enhancement.

Tai-Chiu Hsung received the B.Eng.(Hons.) and Ph.D. degrees in electronic and information engineering in 1993 and 1998, respectively, from the Hong Kong Polytechnic University, Hong Kong. In 1999, he joined the Hong Kong Polytechnic University as a Research Fellow. Now, he is an Associate Professor in the Department of Computer Science of the Chu Hai College of Higher Education, Hong Kong. He is also an Honorary Associate Professor in the Discipline of Oral and Maxillofacial Surgery, Faculty of Dentistry, The University of Hong Kong. His

research interests include biomedical signal processing, wavelets, tomography, and fast algorithms. He is a member of the IEEE, USA and IET, UK.

Guozhao Liao was born in Guangdong Province, China, in 1995. He is currently pursuing the master's degree with the School of Information Engineering, Guangdong University of Technology, Guangzhou, China. His research interests include time frequency analysis, audio signal processing, pattern recognition, cochlear filter bank design and speech enhancement.

Wing-Kuen Ling (Senior Member, IEEE) received the B.Eng. and M.Phil. degrees from the Hong Kong University of Science and Technology, Hong Kong, in 1997 and 2000, respectively, and the Ph.D. degree from Hong Kong Polytechnic University, Hong Kong, in 2003. He was a Lecturer with King's College London, London, U.K., from 2004 to 2010. In 2010, he joined the University of Lincoln, Lincoln, U.K., as a Principal Lecturer, and was promoted to a Reader in 2011. In 2012, he joined the Guangdong University of Technology, Guangdong, China, as a Hundred People Plan Distinguished Professor. His research interests include optimization theory, time frequency analysis, biomedical signal processing, multimedia signal processing, nonlinear digital signal processing systems, and control theory.

Alex Ka-Wing Law was born in Hong Kong. He is now pursuing M.D.S. degree in the Discipline of Oral and Maxillofacial Surgery, Faculty of Dentistry, The University of Hong Kong. His research interests include Obstructive Sleep apnea, machine learning, digital signal processing, and speech enhancement.

Wing Shan Choi received B.D.S., M.D.S. and Ph.D. degrees at the University of Hong Kong in 2003, 2007, and 2018 respectively. She is a Clinical Assistant Professor in the Discipline of Oral and Maxillofacial Surgery, Faculty of Dentistry, The University of Hong Kong. She is also the deputizing Chief of Service in Queen Mary Hospital. Her research focus is Obstructive Sleep apnea, three-dimensional imaging, computer assisted surgery, reconstructive surgery and orthognathic surgery.