All Night Analysis of Snoring Signals by Formant Features

Kun Qian, Yuzhuo Fang, Zhiyong Xu*
School of Electronic and Optical Engineering,
Nanjing University of Science & Technology
Nanjing, China
ezyxu@mail.njust.edu.cn

Huijie Xu*
Otolaryngology Department,
Beijing Hospital
Beijing, China
xhj0531@163.com

Abstract—**This paper presents our preliminary research on all night analysis of formant features extracted from snoring signals (SS), which can be used to find changes of anatomical structure in the upper airway (UA) for patients suffering from Obstructive Sleep Apnea-Hypopnea Symdrome (OSAHS). All night analysis of formant features has a significance to reveal the relationship between physiological or anatomical mechanism and SS. It could be potentially helpful for doctors to conduct an accurate and efficient surgical plan.**

Keywords-Obstructive Sleep Apnea-Hypopnea Symdrome (OSAHS), snoring signals (SS), formant features, all night analysis, clustering

I. INTRODUCTION

Obstructive Sleep Apnea-Hypopnea Symdrome (OSAHS) is a serious chronic disorder which has a prevalent community. It will impair the patient's interpersonal circumstance, lead traffic accidents, lower the quality of a patient's life, and even trigger high blood pressure, coronary heart disease, pulmonary heart failure and some other life-threatening diseases [1]. Polysomnogram (PSG) is regarded as the gold standard in medical practice but has the shortcomings of complexity, extravagance, unpleasant user's experience, etc. Therefore, numerous researchers are seeking for a non-invasive method to be an alternative of PSG. Among these ingenuities, the formant analysis based on acoustic analysis shows a promising potential [2]-[3]. Azadeh et al. investigated the formants of SS and grouped it into 7 clusters based on K-means clustering [2]. They found that the first two formant frequencies (F1 and F2) are most characteristic features in distinguishing the breath and snore sounds from each other. Andrew et al. used formant features to classify apneic and benign snores and they achieved a high recognition rate (sensitivity: 88%, specificity: 82%) with threshold value of F1 = 470 Hz via regression analysis [3].

However, it is more demanding for doctors to acquire information about the obstructive site of the upper airway (UA) of OSAHS patients rather than only diagnose this disorder. Studies have also shown that the obstructive site of the UA will experience a constant change during all night for a patient [4]-[5], for instance, from palatal to non-palatal parts of the UA. Hence, all night analysis of SS data should be taken into practice. Hill et al. calculated acoustic crest factor to show the snoring mechanism change in some individuals during a night [4]. Mesquita et al., Raimon et al., and Duckitt et al. studied the all night SS data by time interval, Apnea-Hypopnea Index (AHI), and Hidden Markov Models (HMM), respectively [6]-[8].

This is a preliminary research for our group to embark an unsupervised method to establish the SS database for further study. We utilized K-means approach to group our SS data into clusters based on the physical meaning indicated by formant features. Combining the characterizations of the UA structure variations (which imply the obstructive site changes [5]) from formant features with all night analysis method of an OSAHS patient, we cluster a 5-hour SS data into 8 clusters and selected the main 5 ones to analysis. The main 5 clusters are supposed to represent 5 typical states of the UA structure.

II. MATERIALS AND METHODS

A. SS Data Acquisition and Pre-processing

This study received the approval of local ethics committee. All night SS data of an OSAHS patient were recorded in Beijing Hospital, China (People's Republic of) and we selected the most typical 5-hour part for analysis. The experimental environment is a sleep laboratory. The acquisition system consists of a low-noise microphone (Studio-Project® B3, USA) positioned at roughly 0.5 m above the patient's mouth [9] and a data acquisition card (M-Audio® Fast Track Ultra 8R, USA) configured at a sampling frequency of 16 kHz with 16-bit resolution. Doctors indicated that the cavity shape of the UA of this patient varied during the whole night.

Based on Short-Time-Analysis method, the SS data were framed with a Hamming window of 64ms by 50% overlapping. Besides, an empirical Short-Time-Energy threshold value was set to eliminate the background noise before further analysis. As a result, we obtained 78,126 frames of SS data, namely 78,126 samples of SS for us to analysis. Figure 1 illustrates the threshold for screening snoring episodes.

B. Formant Features Extraction

Linear predictive coding (LPC) is used to model the UA impulse respond function to generate SS, which is depicted as follows [10]:

$$H(z) = \frac{1}{1 - \sum_{k=1}^{p} \alpha_k z^{-k}}$$

(1)

where a_k ($k = 1, 2, ..., p$) is a set of predictor coefficients directly calculated from SS and p is made to be an optimum order [2]:

$$p_{opt} = Fs(kHz) + \gamma \tag{2}$$

where γ can be 4 or 5. As the SS recorded sampling rate is 16 kHz we set the p as 20. Equation (1) is an autoregressive (AR) model and its parameters are determined by Yule-Walker equation and solved by Levinson-Durbin algorithm [10]. We calculated the roots of the AR model and estimated the angles of the complex roots with positive values, which represent the formant frequencies [10]. In this study, the first three formant frequencies (F1, F2 and F3) were selected as the features to describe the structure variations of the UA.

C. Clustering Method

K-means clustering is a commonly used method to group unsupervised data into K different clusters [11]. As the doctors indicated that there were probably more than 3 structure variations of the UA of this patient, we set K = 8 to cluster formant features and selected the principle 5 states for further study. To avoid the possibility of local optimum, we randomly started the initial cluster centroid positions with 10 replicates.
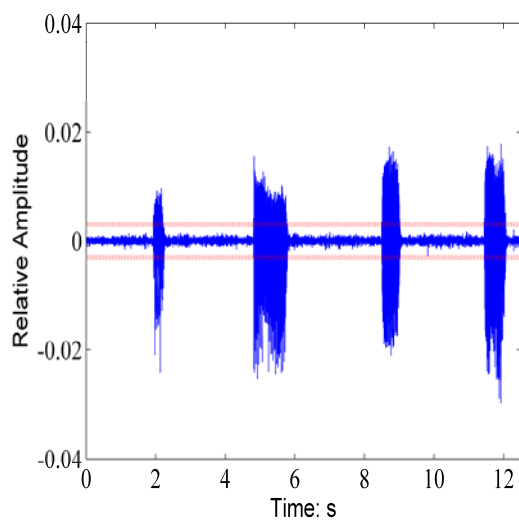

Figure 1. The threshold to eliminate background noise.

III. RESULTS AND DISCUSSIONS

Formant frequencies, also known as the resonance frequencies, are good indicators to reflect the generation of sounds by an acoustic filter: sound energy is attenuated at some certain frequencies while maximal at some others. UA can be regarded as a variable acoustic filter, which generates the SS by vibration of soft tissues due to turbulent airflow through a narrow oropharynx. We can understand from the physical mechanism of snoring that with a narrower cavity structure, namely a more severe extent of obstructive, the airflow through it will generate a higher acoustic frequency (formant) in resonance.

The main 5 states are presented in a 3-D feature space (ie., F1, F2, F3) in Figure 2. It is clear that each state has a significant distinction with other states. The distributions of formant features of all clusters are illustrated by box plots in Figure 3 to Figure 5, respectively. We can see that the other states (which possess a relatively less proportion in all SS data) have much more outliers than the main 5 states therefore these conditions should be excluded from further analysis. The box plot, from the bottom to the top, is the state1, state2, state3, state4, state5 and other states, respectively.

F1 is directly associated with the high and low position of tongue, which achieves a lower value when the tongue is in a higher position [5]. In Figure 3 it is clearly observed that state5 has a lower F1 than the other four states, which implies there will be a larger possibility for state5 to experience an obstruction of the oropharynx.
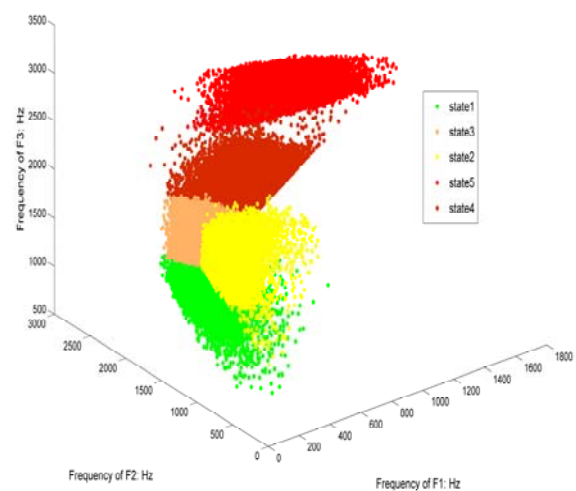

Figure 2. The 3-D feature space of formants.

Figure 4 tells us the distribution of F2 in all states. F2 has a relationship with the front and back positions of the tongue relative to its neutral positions [3]. Studies have indicated that F2 will be at lower frequencies when the tongue is in a more back position [5]. In particular, state1 has the lowest F2 values, which might have a correlation with the problem of the tongue base.

Conventionally, F3 is correlated with the lip-rounding [3], which seems to be more vulnerable to individual characteristics compared with the other two formant features. However, the reflection of the extent of obstruct of the UA has a consistent trend in F3 with F1 and F2 (see Figure 5). We must note that the potential of F3 to indicate whether the SS data is inspiration or expiration should not be ignored. From a respiratory mechanism point of view, inspiration and expiration have different impacts on snoring generation [5].
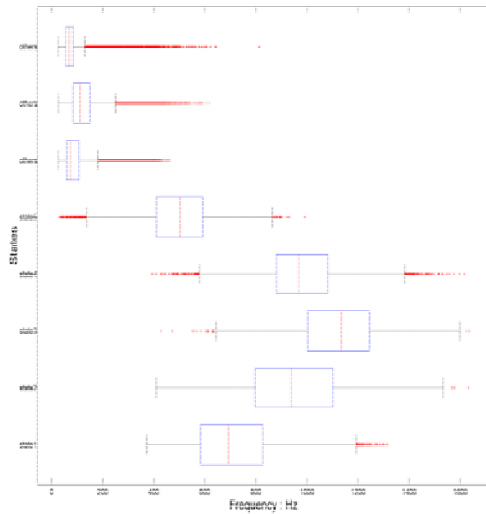
Figure 3. The box plots of F1 for all states (the edges of the box are the 25th and 75th percentiles, the line inside the box is the median, whiskers indicate full range, and outliers are plotted individually).
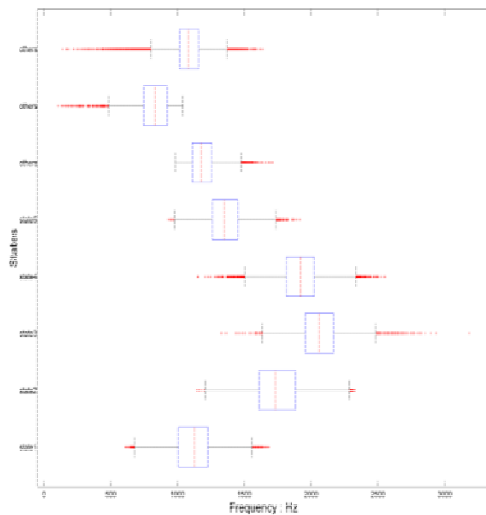

Figure 4. The box plots of F2 for all states.

Figure 5. The box plots of F3 for all states.

Figure 6. The time and frequency domain distribution of 5 typical samples for each main state.

Figure 6 illustrates the features in time and frequency domains of typical samples selected from each main state, respectively. We can clearly see that compared with the other three states, state1 and state3 have relatively regular waveforms in time domain. Specifically, state1, which has been considered as an obstruction of the tongue base, has a periodic time waveforms. In the frequency domain, we can see that the frequency band below 3000Hz is the major area for SS energy distribution. We notice that with more irregular waveforms in time domain, more components appear in higher frequency band in frequency domain.
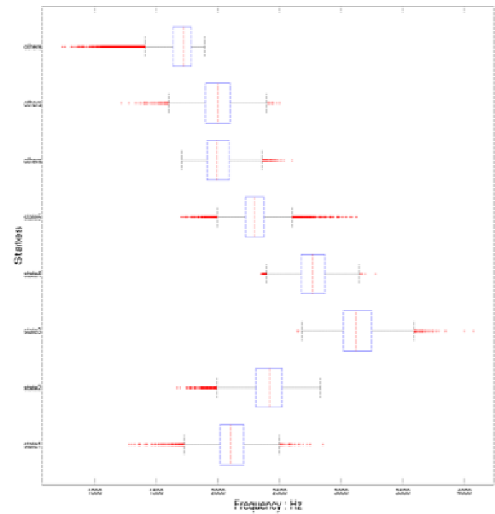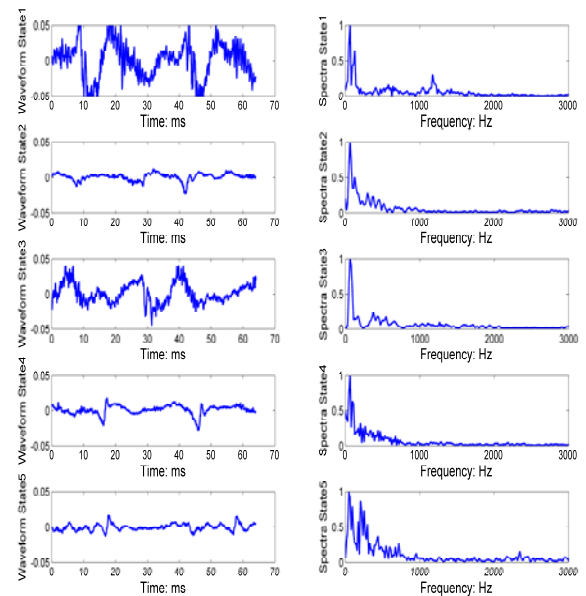
Table 1 illustrates the information about the percentage of each main state in both every hour interval and the whole 5-hour SS Data of this patient. We can see the proportion of each state experiences a slight fluctuation at different durations in the night but always keeps a relative constant ratio of all. The overall data shows that state1 composes the maximum proportion (22.74%) whereas state4 possesses the minimum (10.49%). So it can be infered this OSAHS patient has a possibility for obstructive site of the tongue base.