DETECTION OF UPPER AIRWAY NARROWING VIA CLASSIFICATION OF LPC COEFFICIENTS: IMPLICATIONS FOR OBSTRUCTIVE SLEEP APNEA DIAGNOSIS

Hisham Alshaer1, 4, Martha Garcia3, M. Hossein Radfar5, Geoffrey R. Fernie2, T. Douglas Bradley1, 6

1The Sleep Research Laboratory, 2 Toronto Rehabilitation Institute, Toronto, Ontario, Canada
3Biomedical Engineering, Faculty of Engineering, Universidad Iberoamericana, Mexico City, Mexico
4Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Ontario, Canada
5Department of Electrical and Computer Engineering, University of Toronto, Ontario, Toronto, Canada
6The Centre for Sleep Medicine and Circadian Biology, University of Toronto, Toronto, Ontario, Canada

ABSTRACT

The similarities between unvoiced speech sounds and turbulent breath sounds were used to detect change in sound characteristics caused by narrowing of the upper airway (UA), similar to that occurring in obstructive sleep apnea (OSA). In 18 awake subjects, UA resistance (RUA), an index of UA narrowing, was measured simultaneously with breath sounds recording. Linear Prediction Coding was applied on turbulent inspiratory sounds drawn from low and high RUA conditions and K-means was used to cluster the resulting coefficients. The resulting 2 clusters were tested for agreement with the underlying RUA status. Distinct clusters were formed when RUA increased relatively high but not in cases with lower rise in RUA (P<0.01 for all indicators.) This is the first work to show the utility of LPC in breath sounds analysis confirmed by an objective indicator or UA narrowing.

Index Terms—sleep apnea, acoustic analysis, turbulent breath sounds, upper airway resistance and narrowing, fluid shift

1. INTRODUCTION

Obstructive sleep apnea (OSA) is a breathing disorder characterized by repetitive cessations of breathing from 5 to 100 times/hour during sleep, each lasting 10-60 seconds, due to narrowing and collapse of the upper airway (UA). This causes episodes of oxygen deprivation and provokes arousals from sleep and consequent sleep fragmentation. As a result, patients suffer from poor sleep quality, daytime sleepiness, and impaired cognitive performance. OSA increases the risk of developing hypertension, heart failure and stroke by 3 to 4 fold compared to subjects without OSA [1, 2]. It is a common disease affecting approximately 7% of adults. Nevertheless, the majority of patients with OSA remain undiagnosed; in one study, it was shown that 93% of women and 82% of men with moderate to severe OSA had not been diagnosed [3].

The current method of choice for diagnosing OSA is overnight polysomnography (PSG) in which the patients have to sleep in a laboratory attached to many monitoring electrodes under the supervision of a technician. PSG is expensive and access to it is limited, resulting in long waiting lists. For this reason, several attempts have been made to devise new methods to diagnose OSA using simple techniques that patients can use independently at home such as the acoustic analysis of respiratory sounds [4, 5]. A strikingly common aspect of almost all available studies on the acoustic analysis for the diagnosis of OSA is the focus on snoring sounds as a basis for differentiating people with and without OSA. Although snoring is a hallmark of OSA, it might not necessarily take place with each apnea and hypopnea. Accordingly, the disease severity might be underestimated if some apneas are missed due to absence of snoring. Therefore, comprehensive acoustic analysis of breath sounds should take into consideration both their snoring and non-snoring components. The later result from turbulence created during the passage of air into and out of the lung through the UA. The degree and character of air turbulence should be influenced by changes in UA caliber and airflow rate.

We pioneered this field by our previous research showing that UA narrowing in OSA is at least partially a consequence of fluid shift from the lower body into the neck. For example, fluid displacement from legs due to application of lower body positive pressure (LBPP) via inflatable trousers narrows the UA and increases UA resistance (RUA) [6], presumably due to accumulation of fluid around the UA. The narrowing of the UA without tissue vibration is analogous to generation of unvoiced fricative sounds in speech production. This notion suggests that the quality of breath sounds will vary according to the degree of narrowing similar to the case of unvoiced frication. A major challenge in this work, however, is detecting objectively the occurrence of UA narrowing.

The goal of this work is to detect variations in pure turbulent breath sounds qualities in correlation with change of quantitative index of UA narrowing.

2. METHODOLOGY

2.1. Experimentation and Data Acquisition

2.1.1. Subjects

Data were collected from 18 subjects (4 women, 14 men, age 55.6 ± 10.2, body mass index (BMI) 32.2 ± 8.7, the frequency of apneas and hypopneas per hour of sleep (apnea hypopnea index or AHI) 36.73 ± 20.80. Study protocol was approved by
the research ethics board, and that all subjects provided written informed consent.

2.1.2 Breath Sounds Recordings
Breath sounds were recorded using a cardioid condenser microphone (MX185, Shure®) in front of the subject’s nose and embedded in a full face mask that was strapped to the head as shown in Fig 1. Digitized sound data were transferred to a computer using a USB preamplifier and audio interface (M-Audio, Model Fast Track Pro USB) with a sampling rate (Fs) of 22,050 Hz and resolution of 16 bits. Acquired sound was bandpass-filtered at 20-10,000 Hz.

![Figure 1: Setup of the face mask, microphone, pharyngeal catheters, and the pneumotachometer with a sample waveform of breath sounds.](image)

2.1.3 Lower Body Positive Pressure Application
A pair of deflated medical anti-shock trousers (MAST III-AT; David Clark, Inc.) was wrapped around both legs from the ankles to the upper thighs of supine awake subjects. For the control arm, trousers were left deflated, and for the LBPP arm, trousers were inflated to 40 mmHg to force fluid out of the legs. The subjects were then crossed over to the opposite arm. The duration of each arm lasted 20 minutes. The first five minutes of each arm was a baseline (BL) period, which was used as a reference for the subsequent changes in RUA and breath sounds. Breath sounds and RUA values from the same arm were compared to each other to avoid any possible effect of the change of microphone position during the cross-over.

2.1.4 Measurement of Upper Airway Resistance (RUA)
As an index of UA narrowing we measured RUA. RUA was measured by dividing transpharyngeal pressure (difference between nasopharyngeal and hypopharyngeal pressure measured by two catheters as shown in Fig 1) by simultaneous airflow rate measured by a pneumotachometer attached to the outlet of the facemask given by \( RUA = \frac{\Delta P}{\dot{V}} \) and is expressed in cm.H2O/Leter/second. RUA was calculated at the lowest value of airflow every 30 seconds. Breath sound recordings were synchronized with the pressure and airflow signals in order to correlate sound characteristics with RUA. In this study, we focused solely on the relationship between RUA and breath sounds.

2.2. Data Analysis

2.2.1 Breath Sounds Segmentation and Annotation
Breath sounds included in this analysis are turbulent inspiratory sounds only. Expiratory sounds were excluded to avoid the effect of expired airflow on the microphone. One of the experimenters listened to the breath sounds to exclude snoring and wheezing because they involve tissue vibration and thus are voiced in nature. Two sets of sounds were collected from each experimental arm: one set from the BL and another set at the point at which peak RUA occurred in each of the control and LBPP arms. Each subset of inspiratory sounds was annotated according to the RUA value that accompanied that subset of sounds. Depending on the length of the breathing cycles, 2 to 5 inspirations were selected within each epoch for each RUA value. Breath sounds were displayed and examined in LabVIEW™ (version 9.0 2009).

2.2.2 Linear Predictive Coding (LPC) Implementation
LPC has been accepted as one of the best modeling techniques for speech signals in particular unvoiced speech sounds in order to capture the shape of the vocal tract [7, 8]. The LPC model of unvoiced speech sounds assumes a random noise generator as an excitation source. Turbulent breath sounds share this feature with unvoiced speech sounds because both are generated as a result of the passage of air through the UA, whether fully patent or narrowed, but without the occurrence of tissue vibration such as snoring. LPC models the vocal tract, or the upper airway in this context, as an all-pole filter given by:

\[
H(z) = \frac{1}{1 - \sum_{i=1}^{p} a_i z^{-i}}
\]

with an LPC order p=6. Fig 2 demonstrates the similarity between LPC implementation in speech and breath sounds. LPC was applied in the steps:

1- Because breath sounds vary in amplitude due their cyclic nature, they were normalized in amplitude to remove the effect of gain in the LPC model. Signal’s envelop was found by calculating a moving average of the signal’s variance using a 1,100 point (50 ms) window and then normalizing to that envelop as previously described in [9].

2- Pre-emphasis was applied to compensate for the inherent spectral tilt similar to application in speech [8].

3- In order to apply LPC on equal length segment, normalized breath sounds were segmented with a Hamming window of length ~250 ms with a frame rate of 200 ms. This way, we obtained an average of 272 ± 82 vectors of LPC coefficients from the 36 experimental arms, which comprises our training data sets.
2.2.3 Pattern Recognition and Clustering of Breath Sounds

Having our training data set, then the next step is to classify training data into a number of clusters. The aim of this step is to detect the presence of distinct clusters in each of M=36 each derived from an experimental arm as illustrated in Fig. 3 in the following steps:

1- Feature selection: the 6th order LPC coefficients are selected as a feature for our classifier.

2-Clustering algorithm: we implemented the K-means algorithm on M1-36 with a total of 272 ± 82 LPC vectors in each M. Number of clusters was forced into 2 based on the knowledge of the 2 underlying conditions i.e. BL and peak RUA.

3-Finding clustering tendency: to measure the ability of K-means to separate LPC vectors in M based on the underlying RUA status, BL and peak RUA. This is done by calculating the sum of LPC vectors in each of the 2 resulting clusters for each status:

\[ T = \sum_{i=0}^{n} \left( \frac{x_i}{s_i} \right) \in C_j \]  

(2)

which is the sum of the of LPC vectors \( x_i \) in each inspiratory sound segment \( s_i \), where \( n \) is the total number of vectors in \( M \), \( l \) is the number of inspiratory segments in the data set, and \( C_j \) is each of the resulting clusters (\( j=1,2 \)). If that sum shows that 75% or more of sound segments originating from BL aggregate in a distinct and different cluster from those originating from Peak RUA, then each of the 2 clusters is said to be to have high clustering tendency. On the other hand, if the result is below 75% or if BL and Peak RUA sounds do not aggregate in distinct clusters, then this case is said to have low clustering tendency.

4- Calculating the overall classification accuracy: of this method in differentiating between supposedly different sounds. This was achieved by calculating the weighted sum of the percentages of LPC vectors \( x_i \) in each segment \( s_i \) that were classified in \( C_{cor} \):

\[ A = \sum_{i=0}^{n} w_i \cdot \sum_{i=0}^{n} \left( \frac{x_i}{s_i} \right) \in C_{cor} \]  

(3)

where weight \( w_i \) is equal to the number of frames in each inspiration divided by the total number of frames in a single arm. The rest of the variables are similar to equation 2. All acoustic processing techniques were implemented in MATLAB\textsuperscript{TM} (version 7.9.0 R2009b).

2.3. Relation between Sound Properties and RUA

From the aforementioned calculations, inferences were made on the relation between RUA values of BL and Peak RUA on one hand and clustering tendency on the other. The relations were statistically tested using Wilcoxon rank sum test or t-test depending on the data distribution type.

3. RESULTS AND DISCUSSION

Out of 36 experimental arms, 27 showed high clustering tendency (HCT group) and 9 showed low clustering tendency (LCT group). The characteristic of those groups are as displayed in table 1 and Fig 4. In the HCT group, the peak RUA was 14.9±10.2 units, which was significantly higher than that in LCT, 8±3.8 (\( p=0.0041 \)). Similarly, the difference between BL and peak RUA (\( \Delta RUA \)) in HCT group was 11±9.4, which is significantly higher than \( \Delta RUA \) in LCT group, 5.7±3 (\( p=0.0089 \)).

These results show that the increase in RUA results in change in voice qualities that can be detected with LPC. The overall accuracy of breath sounds classification was 84.7±7.9% vs. 58.6± 5.7% in HCT and LCT respectively (\( P<0.0001 \)). All of those parameters show clearly that LPC coefficients of turbulent breath sounds vary when a rise of RUA takes place above a certain level, but do not when the rise is to lower degree or absent. Since RUA is an indicator of UA narrowing, we suggest that this method can be used to detect UA narrowing.

Figure 2: Proposed analogy of LPC modeling of unvoiced speech sounds and turbulent breath sounds.

Figure 3: Flow chart of the data clustering and analysis algorithm.
Table 1: Summary of $R_{UA}$ values according to the clustering tendency.

<table>
<thead>
<tr>
<th>$R_{UA}$ Status</th>
<th>$H_{CT}$</th>
<th>$L_{CT}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL $R_{UA}$</td>
<td>Average=3.9±1.9 Median = 3.6</td>
<td>Average=2.2±1.3 Median = 2.1</td>
</tr>
<tr>
<td>Peak $R_{UA}$</td>
<td>Average=14.9±10.2 Median = 11.7</td>
<td>Average=8±3.8 Median = 7.6</td>
</tr>
<tr>
<td>$\Delta R_{UA}$</td>
<td>Average=11±9.4 Median = 8.3</td>
<td>Average=5.7±3 Median = 6.3</td>
</tr>
<tr>
<td>A</td>
<td>Average=84.7±7.9 Median = 84.5</td>
<td>Average=58.6 ± 5.7 Median = 58.8</td>
</tr>
</tbody>
</table>

A, overall accuracy (%) give by equation 3.

Figure 4: Box plots of $R_{UA}$ values in the $H_{CT}$ and $L_{CT}$ groups. The horizontal lines represent the median.

We previously showed that fluid displacement by LBPP narrows the UA and increases $R_{UA}$, which is a potential mechanism for OSA [6]. Progressive narrowing of the UA is what results in its collapse and the subsequent apneas. Therefore, detection of the patency status of the UA is an valuable key in diagnosing this disease.

To our knowledge, this is the first study to show the ability of LPC to characterize turbulent breath sounds confirmed by an objective measure of airway narrowing. Snoring sounds have been repeatedly studied for distinguishing people with and without sleep apnea. However, the ability to distinguish totally normal breath sounds, in this study represented by the BL conditions, from those resulting from partial narrowing, in this study represented by peak $R_{UA}$, was not investigated earlier. Snoring in breath sound is analogous to voiced speech. Although snoring is an important sign of OSA, it does not comprise the whole spectrum of breath sounds. For example, snoring might be absent if the cause of apnea originates from the central nervous system, also called central apnea. Also, temporary arousal from sleep could result in increased UA muscle tonus and temporary disappearance of snoring. In such cases, tissue vibration disappears and breath sounds revert to unvoiced sounds resulting from air passage through the airway. Therefore, a comprehensive acoustic diagnosis should take in consideration such sounds and be able to distinguish those that pass through normal vs. narrowed UA.

Future works should focus on aspects needed for clinical implementation of these promising finding. For example, the specific LPC feature that were used by $K$-means to distinguish the different types of breath sounds need to be investigated in each group. Such features can be extracted from the centroids of clusters that were found by $K$-means. Moreover, the relation between those features and the $R_{UA}$ values could be investigated, which could be further developed to non-invasive estimation of airway narrowing. This, in turn, will help creating more accurate tools to diagnose sleep apnea that can be used at home, increase access to diagnosis, and ameliorate the burden of this disease.

4. REFERENCES