AUTOMATIC SNORE AND BREATHING SOUND CLASSIFICATION BASED ON THE SIGNAL ENVELOPE

Masato Kashihara1, Takahiro Emoto2, Udantha R. Abeyratne3, Ikuji Kawata4, Masatake Akutagawa2, Shinsuke Konaka2, Yohsuke Kinouchi1

1Grad. school of Advanced Tech. and Scie., The Univ. of Tokushima, 2-1 Minamijyousanjima-cho, Tokushima 770-8506, JP
2Institute of Tech. and Scie., The University of Tokushima, 2-1 Minamijyousanjima-cho, Tokushima 770-8506, Japan
3School of Info. Tech. and Electrical Engineering, The University of Queensland, St Lucia, Brisbane QLD 4072, Australia
4Department of otorhinolaryngology, Anan kyoei Hospital, Nakasyo, hanoura-cho, Anan-city, Tokushima 779-1198, Japan
{kashihara,emoto,makutaga,konaka and kinouchi}@ee.tokushima-u.ac.jp
udantha@itee.uq.edu.au

ABSTRACT
Snore sound analysis has been recently proposed for diagnosing obstructive sleep apnea (OSA). The snore sounds were typically recorded simultaneously with full-night polysomnography (PSG) resulting in many hours worth of data. Most of the automated snore-analysis techniques require reliable methods to segment out snore sounds from a recording. In this study we focus on this problem and propose a fully automated method to identify breathing and snore sounds. The method is based on the novel feature ‘positive/negative amplitude ratio (PNAR)’ to measure the shape of sound signal. The classification performance of the proposed method was evaluated using receiver operating characteristic (ROC) analysis and was compared with that of traditional zero crossing rate (ZCR). We show that PNAR-based method clearly outperforms ZCR-based method for snore-breathing classification. In particular, PNAR provided better performance for classifying apneic snore and breathing. The new feature PNAR has the potential to contribute towards developing snore-sound based non-contact technology to diagnosis OSA.

KEY WORDS
Breathing, Snore, Signal shape and Classification.

1. Introduction
Obstructive sleep apnea (OSA) is characterized by recurring obstruction of the upper airways during sleep. It is conducive to frequent arousals and oxygen desaturation. OSA increases the risk for cardiovascular diseases, stroke, and contributes to a decreased quality of life and excessive daytime sleepiness. [1].

The reference standard for diagnosis of OSA is full-night Polysomnography (PSG) which requires physical contact with 15-20 sensors. The data processing requires review by a trained sleep technologist. PSG is inconvenient and expensive.

Snoring is the earliest symptom of OSA, and almost all OSA patients snore [1,2]. However, not all snorers suffer from OSA. In recent years, several researchers have analyzed snoring for diagnosing the obstructive sleep apnea (OSA). These researches have reported significant differences in the snoring between the Non-OSA/OSA groups [3-9]. The snore sounds used in the analysis were recorded simultaneously with PSG in the hospital. Due to the long record of snore related sound, the automatic method is suitable for snore-based OSA diagnosis. Many of these techniques need to accurately identify snore sounds in long overnight sound recordings.

There are several attempts at developing automated technology for snore/breathing classification. These typically used features such as: spectral/signal energy, zero crossing rate (ZCR) and first formant frequency (F1) etc [10,11]. These methods provided accurate classifications (about 90%) of snore and breathing sounds. In this paper our interest is to develop novel, simpler and more accurate features for automatic snore/breathing classification.

In this paper we propose technology based on the envelope of the sound signal. The performance of the proposed method was evaluated via receiver operating characteristic (ROC) analysis and was compared with that of ZCR. To the best of our knowledge, no researcher has thus far reported on the sound classification using the signal shape. The main motivation is to investigate if the signal shape carries useful information on accurate classification of snore and breathing sounds. The new signal shape feature has potential applications in developing non-contact snore-based OSA diagnostic tools. We show that the signal shape carries information, hitherto disregarded, on classification of snore and breathing sounds.

2. Methods
2.1 Clinical snore/breathing signals
In this paper we use sleep sound recordings acquired with

DOI: 10.2316/P.2013.791-061

46
non-contact instrumentation during a routine overnight PSG. A low noise microphone (Model NT3, RODE, Sydney, Australia) was used for the sleep sound recordings. The nominal distance from the microphone to the mouth of the patient ranges from 40cm to 70cm due to patient movements. A pre-amplifier and A/D converter unit (Mobile-Pre USB, M-Audio, California, USA) was used at a sampling rate of 44.1k samples/s and a bit resolution of 16bits/sample. The sounds recorded during sleep contain not only snoring episodes (SEs) but also sounds such as breathing episodes (BEs), duvet noise, and speech etc.

2.2 A measure to quantify the shape of the signal

We divided the snore recording $x(n)$ into segments of length $M$, with the segment overlap length given by $S$ samples.

The maximum positive amplitude in the $k$th segment of $x(n)$ is computed as

$$P_k = \max[x_k(n)], \quad k = 1, \ldots, K$$

Where $K$ is the total number of segments.

The maximum negative amplitude in the $k$th segment of the signal is also computed as:

$$N_k = \max[-x_k(n)]$$

In this paper, we propose a new measure called the Positive-Negative Amplitude Ratio (PNAR) as given by:

$$PNAR = \text{Var} \left( \frac{P_k - N_k}{P_k + N_k} \right)$$

Where $\text{Var}(\cdot)$ denotes the variance.

In this paper we classify breath and snore sounds based on this measure.

It should be noted that the average value of a snoring episode should be zero, the violation of which leads to a net and progressive accumulation of pressure inside the body. However, over a finite time period within a snore, the signal can be asymmetric about the baseline, with positive going and negative going excursions providing information on the upper airway characteristics. The measure PNAR attempts to capture such information.

We illustrate a typical sample of (a) a snore and (b) a breathing sound in Fig. 1, demonstrating $P_k$ and $N_k$.

2.3 Conventional measure to classify breathing and snore sound

In this study we use the measure ZCR for the comparison of the performance for breathing–snore classification. The average ZCR can be used for distinguishing between voiced and unvoiced speech signals, because unvoiced speech signals have much higher ZCR values than voiced ones under high SNR conditions [14]. The measure ZCR has been used for the breathing and snore sound classification [11].

The measure ZCR was calculated from each segment as,

$$Z_k = \frac{\sum_{m=1}^{M-1} [\text{sgn}(x_k(m + 1)) - \text{sgn}(x_k(m))] - \text{sgn}(x_k(m))]}{2M}$$

where $\text{sgn}(\cdot)$ is ‘1’ for positive arguments and ‘0’ for negative arguments.

The average of $Z_k$ is defined as

$$\text{ZCR} = \frac{1}{K} \sum_{k=1}^{K} Z_k.$$
classified segments, respectively. Note that TN and TP denote the number of correctly classified breathing sounds and the number of correctly classified apneic/non-apneic snore sounds, respectively.

AUC provides a quantitative index of the classification accuracy of the method. The classification varies from 0.5 (random chance) to 1.0 (perfect accuracy).

3. Results

For this work, snore sounds of 10 apneic snorers (9 males; 1 female) and 10 non-apneic snorers (7 males; 3 females) and breathing sounds of 12 apneic/non-apneic snorers (10 males; 2 females) were manually identified in the recordings. Each sound was divided into training and testing data. Training data was used to determine the optimal threshold value by using ROC analysis. Testing data was used to evaluate the sensitivity and specificity of the determined threshold value.

Table 1 gives a detailed description of training and testing data.

We used the following parameters in our experiments: \( M = 28, 55, 110, 220, 440 \) and \( 880 \); \( S = 1 \) for PNAR and \( S = 50\% \) of segment \( M \) for ZCR. The value of \( K \) depends on the actual length of \( x(n) \).

Figure 2 (a) A distribution of the pair (ZCR, PNAR) of apneic snore, non-apneic snore and breathing sounds at case \( M = 28 \) and (b) an extended version for further information. Symbol '+' and 'x' denote apneic, non-apneic snore and breathing, respectively.

Table 1 Training and testing database

<table>
<thead>
<tr>
<th>Type</th>
<th>Age</th>
<th>AH1</th>
<th>Data size (train; test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apneic snore</td>
<td>63.8±13.2</td>
<td>35.4±19.3</td>
<td>2116; 1058</td>
</tr>
<tr>
<td>Non-apneic snore</td>
<td>63.4±13.7</td>
<td>39±2.7</td>
<td>3473; 1736</td>
</tr>
<tr>
<td>Breathing</td>
<td>62.0±15.1</td>
<td>22.1±23.5</td>
<td>1521; 761</td>
</tr>
</tbody>
</table>

Fig. 2 shows a distribution of the pair (ZCR, PNAR) of apneic snore, non-apneic snore and breathing sounds at case \( M = 28 \). According to Fig. 2, both ZCR and PNAR can successfully distinguish breathing from snore sounds. Mostly the high ZCR indicates breathing and low ZCR indicates snore sound. However we find that less frequently, apneic snores have high ZCR. Even in that case, we can distinguish snore and breathing with the help of PNAR.

The classification performance was compared for apneic snore-breathing and non-apneic snore-breathing. The influence of segments of length \( M \) on the measures ZCR and PNAR, in classifications of training data are shown in Table 2. The classification results on testing data are also given in Table 3.

Table 2

<table>
<thead>
<tr>
<th>Classification</th>
<th>ZCR</th>
<th>PNAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>0.935 0.928 0.979 0.950 0.955</td>
<td>0.935 0.929 0.974 0.956 0.943</td>
</tr>
<tr>
<td>55</td>
<td>0.935 0.925 0.962 0.959 0.936</td>
<td>0.935 0.927 0.951 0.932 0.926</td>
</tr>
<tr>
<td>110</td>
<td>0.935 0.920 0.899 0.891 0.843</td>
<td>0.935 0.912 0.778 0.733 0.751</td>
</tr>
<tr>
<td>220</td>
<td>0.935 0.912 0.898 0.890 0.842</td>
<td>0.935 0.902 0.776 0.723 0.755</td>
</tr>
<tr>
<td>440</td>
<td>0.935 0.902 0.897 0.890 0.842</td>
<td>0.935 0.893 0.776 0.723 0.755</td>
</tr>
</tbody>
</table>

According to these tables, when \( M \) was small, PNAR-based method clearly outperformed ZCR-based method. No variation in the performance of ZCR was found by increasing \( M \). However PNAR provided a poor performance when \( M \) was large. This is due to the loss of information in pitch period, which is derived from periodic pulse train. On the other hand, when \( M \) was small, PNAR had better performance compared to ZCR. In particular, this trend was confirmed for apneic snore-breathing classification case. Based on our observations, ZCR indicated poor performance under the following condition: apneic snores contain enough high-frequency spectral components [15]. Even under this condition,
PNAR can be used with high accuracy as it can still capture the feature of the quasi-periodic pulse train. These results clearly indicate that PNAR is superior to ZCR when $M$ was small.

4. Discussion and Conclusion

We showed that the proposed measure PNAR can accurately distinguish between snore and breathing, because snores have higher PNAR than breathing. The PNAR-based method clearly outperformed ZCR-based method when a small segment was employed for the analysis.

The proposed method has following advantages over conventional techniques: (1) in particular, the method is superior to ZCR-based method for classifying apneic snore and breathing, (2) the method is simpler than existing features and is easier to implement on portable computational devices, (3) the method is sensitive to snore episodes, which contain enough high-frequency spectral components.

The ZCR may fail to separate apneic snores from breathing. However PNAR is very sensitive to apneic snores and is free from that problem. This feature should help improve the performance of snore-based OSA diagnosis technology.

Acknowledgements

Authors would like to acknowledge the following organizations for partially supporting this work: Australian Research Council (Grant #: DP120100141, Japan Society for the Promotion of Science (Grant #: 24700498).

References


