Automatic and Unsupervised Snore Sound Extraction From Respiratory Sound Signals

Ali Azarbarzin, Student Member, IEEE, and Zahra M. K. Moussavi*, Member, IEEE

IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 58, NO. 5, MAY 2011

Presenter : Hao-Ting Yen
Advisior : Dr.Chun-Ju Hou
Outline

- Introduction
- Method
- Results
- Discussion & Conclusions
- Reference
Introduction (1/6)

- Snoring is a prevalent disorder affecting 20%–40% of the general population.

- Snoring sound usually disrupts the sleep of the snorer’s bed partner straining the relationship.
Introduction (2/6)

- OSA causes daytime sleepiness, leads to impaired job performance, and increases the risk of accidents.
Introduction (3/6)

- Polysomnography (PSG):
  - Patients need to spend the entire night in a sleep lab.
  - Long waiting list
  - Costly
  - Connecting many wires and electrodes to the patient.
Introduction\(^{(4/6)}\)

- Acoustical analysis
  - Noninvasive
  - Simple to apply
  - Could be integrated into a portable and low-cost bedside devices installed at the patient’s home.
Introduction (5/6)

- Acoustical analysis of snore sounds
  - Assessment of the outcome of surgical treatment
  - Classification of snorers as simple snorer or patients with OSA
The method using hidden Markov models and spectral-based features, segmented into:

- Snoring episodes
- Breathing
- Duvet noise
- Silence

Accuracy of detecting snoring episodes was reported to be 82%–89%
Method

Data Recording

- 30 patients with an average age of 50.6 years (STD = 9.96)

<table>
<thead>
<tr>
<th>Group</th>
<th>Number of Patients</th>
<th>Age</th>
<th>Body mass index</th>
<th>AHI</th>
</tr>
</thead>
<tbody>
<tr>
<td>OSA</td>
<td>23 (7 females)</td>
<td>49.9±10.2</td>
<td>34.1±7.2</td>
<td>26.1±22.9</td>
</tr>
<tr>
<td>Simple Snorers</td>
<td>7 (no females)</td>
<td>53.1±9.3</td>
<td>30.0±3.8</td>
<td>2.3±1.5</td>
</tr>
</tbody>
</table>
Method

- **Data Recording**
  - Two microphones (ECM-77B) 40 Hz–20 kHz
    - Tracheal
    - ambient
  - Biopac (DA100C) amplifiers
    - Gain: 200
    - Bandpass filter: 0.5–5 kHz
Method

- Data Recording
  - NI9217 data-acquisition module
    - Sample rate: 10240 Hz
  - Software: LabView
  - 5665 snoring segments
    (around 15 min, mostly when the patient was snoring)
Method

• Signal Analysis

- Modified V-Box Control Chart
- Feature Extraction
- Unsupervised Classification Using FCM
Method

- Modified V-Box Control Chart

\[ B(L, H, n, Y_n) = [n - L, n] \times [-H, H] \]

- \( Y_n \): last observation
- \( n \): time point
- \( L > 1 \): the width of the V-Box (number of past observations to be taken into account)
- \( 2H \): the height of the V-Box
Method

- Modified V-Box Control Chart

  - current time point $\geq L + 1$, $B(L, H, n, Y_n)$, $b_{LH}(n)$, is counted.
  
  - $b_{LH}(n) \leq L \times \theta$, mark a change in the process. The parameter $0 < \theta < 1$ should be chosen such that the fraction of observations located in the box $B(L, H, n, Y_n)$ equals $L \times \theta$.

---

Fig. 1. Typical 20 s of a tracheal sound recording including breathing and snoring episodes and the segmentation result using the V-Box algorithm. At this stage, all potential snore sound segments are identified. $E_i$ denotes episode $i$, where $i$ is 1, 2, 3, . . . . For example, in this 20-s interval, there are six episodes.
Method

- **Feature Extraction**
- First, the short-time Fourier transform (STFT) of each episode was calculated using 50-ms windows with 50% overlap between successive windows.
- calculate the elements of feature vector $\varepsilon_j$, in which $i$th element of the feature vector was computed as

$$
\varepsilon_i^j = \frac{\sum_{n=1}^{N_j} \sum_{f=500(i-1)}^{500i} |\hat{S}_j(n, f)|^2}{\sum_{n=1}^{N_j} \sum_{f=0}^{5000} |\hat{S}_j(n, f)|^2}, \quad i = 1, \ldots, 10
$$

or $\varepsilon_j$, in which $i$th element of the feature vector was computed as
Method

- Feature Extraction

\[
COV = \frac{1}{N-1} \sum_{j=1}^{N} (\varepsilon^j - \bar{\varepsilon})(\varepsilon^j - \bar{\varepsilon})^T
\]

\[
\hat{\varepsilon^j} = W^T \varepsilon^j.
\]

Fig. 2. Clustered feature vector using the FCM method. This 2-D feature vector was obtained using a projection from 10-D feature space onto a 2-D feature space by PCA. The original 10-D feature vector was obtained by calculating 500-Hz energy band distribution from a segmented sound signal. The big rectangle is a zoomed-in version of the snoring cluster located around horizontal axes.
Method

- Unsupervised Classification Using FCM
Results
Results

Fig. 3. Successful removal of noise and breathing episode after clustering. Note that snore episode $E_6$ was not detected, and therefore, we have a false alarm in this case.
Results

\[
\text{Accuracy} = \frac{TP}{TP + FN} \times 100
\]

\[
\text{PPV} = \frac{TP}{TP + FP} \times 100
\]
Results

**TABLE II**

CLASSIFICATION RESULTS AND TOTAL NUMBER OF BREATHS FOR SAMPLE CASE WITH 171 SNORING EPISODES

<table>
<thead>
<tr>
<th>Data Set</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th># of Breaths</th>
<th># of noises</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample case</td>
<td>169</td>
<td>16</td>
<td>2</td>
<td>324</td>
<td>2</td>
</tr>
</tbody>
</table>
Results

TABLE III
CLASSIFICATION RESULTS FOR TRACHEAL MICROPHONE

<table>
<thead>
<tr>
<th>Data Set</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>Accuracy (%)</th>
<th>PPV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Snorers</td>
<td>5588</td>
<td>304</td>
<td>77</td>
<td>98.6</td>
<td>94.8</td>
</tr>
<tr>
<td>and OSA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OSA</td>
<td>3816</td>
<td>203</td>
<td>48</td>
<td>98.8</td>
<td>94.9</td>
</tr>
<tr>
<td>Simple Snorers</td>
<td>1772</td>
<td>101</td>
<td>29</td>
<td>98.4</td>
<td>94.6</td>
</tr>
</tbody>
</table>

TABLE IV
CLASSIFICATION RESULTS FOR AMBIENT MICROPHONE

<table>
<thead>
<tr>
<th>Data Set</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>Accuracy (%)</th>
<th>PPV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Snorers</td>
<td>5275</td>
<td>223</td>
<td>390</td>
<td>93.1</td>
<td>95.9</td>
</tr>
<tr>
<td>and OSA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OSA</td>
<td>3569</td>
<td>104</td>
<td>327</td>
<td>91.6</td>
<td>97.2</td>
</tr>
<tr>
<td>Simple Snorers</td>
<td>1706</td>
<td>119</td>
<td>63</td>
<td>96.4</td>
<td>93.5</td>
</tr>
</tbody>
</table>
### Results

#### TABLE V
**Classification Results for Different Values of Overlap**
\( L = 50 \text{ ms and } \theta = 0.95 \)

<table>
<thead>
<tr>
<th>Overlap(%)</th>
<th>Accuracy (%)</th>
<th>PPV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>97.4</td>
<td>94.4</td>
</tr>
<tr>
<td>70</td>
<td>97.5</td>
<td>93.8</td>
</tr>
<tr>
<td>80</td>
<td>98.6</td>
<td>94.8</td>
</tr>
</tbody>
</table>

#### TABLE VI
**Classification Results for Different Values of \( L \)**
\( \text{Overlap} = 80\% \text{ and } \theta = 0.95 \)

<table>
<thead>
<tr>
<th>L(ms)</th>
<th>Accuracy (%)</th>
<th>PPV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>98</td>
<td>93.3</td>
</tr>
<tr>
<td>50</td>
<td>98.6</td>
<td>94.8</td>
</tr>
<tr>
<td>100</td>
<td>97.5</td>
<td>94.5</td>
</tr>
</tbody>
</table>
Results
## Results

### Table VII

**Classification Results for Different Values of $\theta$**

*(Overlap = 80% and $L = 50$ ms)*

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>Accuracy (%)</th>
<th>PPV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7</td>
<td>88.3</td>
<td>96.9</td>
</tr>
<tr>
<td>0.75</td>
<td>92.3</td>
<td>96.6</td>
</tr>
<tr>
<td>0.8</td>
<td>94.7</td>
<td>95.8</td>
</tr>
<tr>
<td>0.85</td>
<td>95.8</td>
<td>95.7</td>
</tr>
<tr>
<td>0.9</td>
<td>96.6</td>
<td>95</td>
</tr>
<tr>
<td>0.95</td>
<td>98.6</td>
<td>94.8</td>
</tr>
<tr>
<td>0.98</td>
<td>97.6</td>
<td>92.5</td>
</tr>
</tbody>
</table>
Discussion & Conclusion

- The main advantages of the proposed algorithm in this paper are its high accuracy, robustness and insensitivity to the severity of AHI unsupervised, automatic operation, and low computational cost.

- Classification of data recorded by an ambient microphone would reduce to classify snore versus silence, while it is snore versus breath in case of data recorded by a tracheal microphone.
Discussion & Conclusion

- An automatic, reliable, fast and accurate snore extraction method is essential for diagnosis and treatment of different snore-related disorders, such as OSA.

- Investigate the relationship between body position and snore intensity, and identify the person’s best sleeping position in which the least number of snoring episodes occur.
References


References


References


References


References


Thanks you for your attention