Smart Audio Sensor on Anomaly Respiration Detection using FLAC features

Jiaxing YE  
Computer Science Dept. Univ. of Tsukuba, Japan

Takumi Kobayashi  
AIST, Japan

Tetsuya Higuchi  
AIST, Japan  
Computer Science Dept. University of Tsukuba, Japan
Content overview

◊ Research background
◊ The proposed approach
   The FLAC feature extraction method
   The anomaly detection framework
◊ Experimental validations
◊ The generalized application of proposed scheme
◊ Conclusions
Research background

The facts of respiration symptoms from World Health Org. (WHO)

1. Many kinds of respiration diseases, worldwide spread, e.g. 235 million people suffering from asthma.
2. Chronic disease or cannot be cured, by well treatment works for higher life quality.
3. Early and effect diagnosis is of crucial importance
Respiration symptom manual diagnosis:

- Hard for early symptoms (neglect)
- Error in symptom sampling
- Time-consuming (healthy + symptom)

Computerized Symptom targeting
Research background

Goal: supporting respiration symptom early diagnosis

Respiration sounds: healthy (Common)+ symptom (Seldom)

Anomaly detection by the proposed system

Discard the healthy sections

Normal parts

Abnormal sections recording: suspected symptoms

Further diagnosis by doctor
Research background

Requirements for computerized respiration diagnosis
1. Suitable for each individual with variation adaptation
2. High detection performance, especially for early diagnosis
3. High efficiency for real-time operation and hopefully low cost for wide applications.

Algorithm interpretation:
1. Unsupervised manner for different individual
2. Online scheme for self-adaptation
3. Algorithm effectiveness and efficiency
The proposed approach

The FLAC feature extraction method

Acoustic feature extraction:

Structured sounds: (30 years more)

Speech — formantic structure

Music — harmonic structure

unstructured sounds: (new field, 2005–)

Variably composed, not comply with pre-defined model: respirations
The proposed approach
The FLAC feature extraction method

Acoustic signals:

Structured sounds: (30 years more)

Unstructured sounds: (new, 2005 --)

Features:

- MFCC (Mel-frequency cepstral coefficients)
- LPC (Linear Predictive Code), Subband Energy
- ZCR (Zero-Cross Rate), Pitch, Spectrogram...

Utilizing the new feature for respiration analysis (FLAC)

IEEE Sensor Application Symposium, Brescia, Italy, Feb. 8th, 2012
The proposed approach

The FLAC feature

☆ 1. Taking advantage of complex spectrogram, not rely on the magnitude only.

☆ 2. Extracting the time-frequency domain dynamics features for representing unstructured sounds.

\[ a + bi = A \cdot e^{i\phi} \]

Phase

Magnitude
The proposed approach
The FLAC feature extraction method

Recent advancements on using phase information


Do magnitude + phase!

The proposed approach
The FLAC feature extraction method

1.STFT (Short-time Fourier Transform): obtaining the magnitude spectrogram

Extracting dynamics on t-f plane by local descriptor


Image processing on (complex) spectrogram
The proposed approach
The FLAC feature extraction method

- Details in FLAC feature extraction

  a. Short-time Fourier Transform

  \[ a + bi = A \cdot e^{j\phi} \]

  - Phase + magnitude

  - Time-frequency dynamics

  - FLAC mask patterns

  b. FLAC features

  - Temporal Dynamics

  - Phase

  - Magnitude

  - Frequency (d-dim.)

  - Mask patterns

  - Frequency dynamics

  - FLAC features

  - Inner Domain Dynamic Features

  - Cross Domain Dynamic Features

- The proposed approach
- The FLAC feature extraction method

IEEE Sensor Application Symposium, Brescia, Italy, Feb. 8th, 2012
The proposed approach

The anomaly detection framework

Principle components, within-class

Minor components, between classes

\[ S_{\text{Abnormal}} \]

\[ S_{\text{Norm}} \]

\[ S = S_{\text{Norm}} \oplus S_{\text{Abnormal}} \]

\[ d_0 : \text{Abnormality index} \]

(Principle Component Analysis)

PCA based signal decomposition

Anomalies:

Noise

Interference

Outliers

The proposed approach
The anomaly detection framework

Detection framework
Experimental validation

Dataset Introduction: stethoscope recordings at 44.1 kHz with 16bit quantization.

<table>
<thead>
<tr>
<th>Healthy respiration sounds</th>
<th>Symptom respiration sounds</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Class1 Normal</strong></td>
<td><strong>Adventitious symptoms</strong></td>
</tr>
<tr>
<td></td>
<td>Fine crackle, Bronchiectasis, Bronchial stenosis,</td>
</tr>
<tr>
<td></td>
<td>Asthma, Tracheobronchial stenosis, Pulmonary fibrosis</td>
</tr>
<tr>
<td></td>
<td>(Interstitial pneumonia), Coarse crackle, Rhonchus,</td>
</tr>
<tr>
<td></td>
<td>Adult respiration distress syndrome (ARDS), Wheeze,</td>
</tr>
<tr>
<td></td>
<td>Pneumonia, Expectoration, Congestive heart failure,</td>
</tr>
<tr>
<td></td>
<td>Pulmonary edema, stethocatharsis sedimentation</td>
</tr>
<tr>
<td><strong>Symptom respiration sound data</strong></td>
<td><strong>Breath symptoms</strong></td>
</tr>
<tr>
<td></td>
<td>Spontaneous pneumothorax, Retention of pleural effusion,</td>
</tr>
<tr>
<td></td>
<td>Atelectasis, Tracheal stenosis, Hemopneumothorax,</td>
</tr>
<tr>
<td></td>
<td>Enhanced respiration</td>
</tr>
<tr>
<td><strong>Class2 Abnormal</strong></td>
<td></td>
</tr>
</tbody>
</table>
Experimental validation

Parameter setup

Frequency location
\[ f < f_{\text{max}} \]

Fourier analysis window length
\[ L_{\text{Fourier}} = \text{Cycle of respiration} \]
Experimental validation

Modeling normal respirations

Data Input

MFCC

Complex spectrogram

FLAC

IEEE Sensor Application Symposium, Brescia, Italy, Feb. 8\textsuperscript{th}, 2012
Experimental validation

Modeling normal respirations: Base line wandering effect
Experimental validation

Anomaly detection in respiration sound

Case 1:

<table>
<thead>
<tr>
<th>Result (%)</th>
<th>Normal clips</th>
<th>Symptom clips</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Positive</td>
<td>13/13 = 100</td>
<td>25/27 = 92.59</td>
<td>38/40 = 95</td>
</tr>
<tr>
<td>Detection ratio</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Efficiency**: Proposed approach costs 21ms for detecting 1 second respiration by using Intel Core i5 580M with 4GB memory.

Case 2: symptoms among normal respirations
Generalized application: foreground acoustic event detection

Acoustic environment modeling for acoustic event detection

Acoustic event in noisy factory environment

20s Laughing

1m06s Bell ringing
Conclusions

Unsupervised learning for individuals respiration character

Adaptation to the variations in respiration in an online manner

High anomaly respiration detection performance

High efficiency (FFT with Local Auto-correlation)
Thank you very much for your attention!

Questions