Comparison of Cough, Wheeze and Sustained Phonations for Automatic Classification between Healthy subjects and Asthmatic patients

Shivani Yadav, Kausthubha NK, Dipanjan Gope, Uma Maheswari Krishnaswamy, Prasanta Kumar Ghosh

SPIRE LAB
Electrical Engineering,
Indian Institute of Science (IISc), Bangalore, India
Table of Contents

1 Introduction
2 Motivation
3 Dataset
4 Proposed Method
5 Experimental Setup
6 Results
7 Conclusion and Future work
Asthma is an **inflammatory disease** of the airways resulting in a number of symptoms including **obstruction of the airways**, chest discomfort or pain, cough, and wheezes or other peculiar sounds during breathing.

*Figure: Healthy and Asthmatic Bronchioles*  

Asthma is an **inflammatory disease** of the airways resulting in a number of symptoms including **obstruction of the airways**, chest discomfort or pain, cough, and wheezes or other peculiar sounds during breathing.

334 million is global burden of asthma\(^2\).

---

Traditional diagnostic tool: Spirometry

- A lung function test that **measures how much and how fast a person can exhale air**.
- Helps in **diagnosis** and **monitoring** asthma in hospital.
- Drawbacks of spirometry
  1. Very strenuous
  2. More training required
  3. More time required to screen large population

Figure: Spirometry

Introduction

Traditional diagnostic tool: Spirometry

- A lung function test that measures how much and how fast a person can exhale air.
- Helps in diagnosis and monitoring asthma in hospital.
- Drawbacks of spirometry
  1. Very strenuous
  2. More training required
  3. More time required to screen large population

Is there a simpler approach for asthma diagnosis?

Figure: Spirometry

1 http://www.virtualimaging.org/pft.html
# Table of Contents

1. Introduction
2. Motivation
3. Dataset
4. Proposed Method
5. Experimental Setup
6. Results
7. Conclusion and Future work
Voice based analysis

1. Less training required.
2. Less timing required to screen large population.
3. Less strenous.
4. Easy for the aged and children.
Voice based analysis

1. Less training required.
2. Less timing required to screen large population.
3. Less strenous.
4. Easy for the aged and children.

Does human voice encode cues to detect asthma??
Voice based analysis is possible!

- **Key finding**: Acoustics are different between healthy and asthmatic subjects\(^1\).
- **Stimuli used**: Sustained phonations
  - /ɑː/ (as in ‘Father’)
  - /iː/ (as in ‘See’)
  - /ʊː/ (as in ‘Blue’)
  - /oʊ/ (as in ‘Go’)
  - /aɪ/ (as in ’Eye’).

---

Voice based analysis is possible!

- **Key finding:** Acoustics are different between healthy and asthmatic subjects.\(^1\)
- **Stimuli used:** Sustained phonations
  - /ɑ:/ (as in ‘Father’)
  - /iː/ (as in ‘See’)
  - /uː/ (as in ‘Blue’)
  - /ɒʊ/ (as in ‘Go’)
  - /aɪ/ (as in ‘Eye’).

What are the other possible stimuli?

---

Illustration

- Cough
- Wheeze
Illustration

- Cough
- Wheeze

Control

Patient
How well do the different stimuli encode cues to detect asthma?

- Cough
  - Control
  - Patient

- Wheeze
  - Control
  - Patient
How well do the different stimuli encode cues to detect asthma?
■ **35 Patients**- 18 male, 17 female.

■ **36 Controls (healthy)**- 18 male, 18 female.

The age range of controls are between 19-42 years and average age of 24 years. The age range of the patients are 19-78 years with an average age of 43 years.

■ **Stimuli:**

1. Cough
2. Wheeze
3. Sustained phonations: /ɑː/ (as in ‘father’), /iː/ (as in ‘See’), /uː/ (as in ‘Blue’), /eɪ/ (as in ‘Say’), /oʊ/ (as in ‘Go’).

■ Average number of recordings per stimuli per subject is **5**.
Table of Contents

1 Introduction
2 Motivation
3 Dataset
4 Proposed Method
5 Experimental Setup
6 Results
7 Conclusion and Future work
Schematic Diagram

**Proposed Method**

**Training**
- Healthy
  - MFCC Computation
  - Statistics
  - SVM Training
- Asthmatic
  - MFCC Computation
  - Statistics

**Testing**
- Unknown
  - MFCC Computation
  - Statistics
  - SVM Classifier

Decision
Schematic Diagram

**Training**

- Healthy
  - MFCC Computation
  - $F_H$

- Asthmatic
  - MFCC Computation
  - $F_A$

**Mel-frequency cepstral coefficients (MFCC)**

- MFCC dimension
- Time frame index
- $K$
- $N$

- MFCC dimension
- Time frame index
- $K$
- $M$
Proposed Method

Schematic Diagram

Training

Healthy

MFCC Computation ➔ \( F_H \) ➔ Statistics ➔ \( S_H \) ➔ MFCC Statistics matrix

Asthmatic

MFCC Computation ➔ \( F_A \) ➔ Statistics ➔ \( S_A \) ➔ MFCC Statistics matrix

a. Mean
b. Median
c. Mode
d. Root mean square (RMS)
e. Variance
f. Standard Deviation (SD)
Schematic Diagram

**Training**

- Healthy
  - MFCC Computation → $F_H$ → Statistics → $S_H$
- Asthmatic
  - MFCC Computation → $F_A$ → Statistics → $S_A$

**Testing**

- Unknown
  - MFCC Computation → $F_U$ → Statistics → $S_U$ → SVM Classifier → Decision
Proposed Method

Schematic Diagram

Training

Healthy

MFCC Computation → $F_H$ → Statistics → $S_H$ → SVM Training

Asthmatic

MFCC Computation → $F_A$ → Statistics → $S_A$

Testing

Unknown

MFCC Computation → $F_U$ → Statistics → $S_U$ → SVM Classifier → Decision
Table of Contents

1. Introduction
2. Motivation
3. Dataset
4. Proposed Method
5. Experimental Setup
6. Results
7. Conclusion and Future work
Experimental setup

- Sampling rate of: 48 kHz, analysis window: 20ms, shift: 10ms.
- Support vector machine (SVM) used with 4 fold cross-validation setup.
- Evaluation metric
  - Total classification accuracy
    \[ TCA = \frac{TP + TN}{TP + TN + FP + FN} \]
    where, TP (True Positive), TN (True Negatives), FP (False Positives) and FN (False Negatives).
**Key questions addressed**

1. **Which stimuli** (cough, wheeze and sustained phonation) are better for classification?

2. What is the **effect of including** $\Delta$ (*velocity*) and $\Delta\Delta$ (*acceleration*) **coefficients** on the classification results?

3. To **what degree do individual MFCCs** encode asthma related information for the best performing stimuli?

4. **Which among the 6 statistics** are more discriminant in the best performing stimuli for the classification task?
# Table of Contents

1. Introduction
2. Motivation
3. Dataset
4. Proposed Method
5. Experimental Setup
6. Results
7. Conclusion and Future work
Classification Results

Table: Mean (Standard deviation) of Total classification accuracy (TCA%) using static, $\Delta$ (velocity) and $\Delta\Delta$ (acceleration) coefficients.

<table>
<thead>
<tr>
<th>Stimuli</th>
<th>Total classification accuracy (TCA) %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K = 12$</td>
</tr>
<tr>
<td>/a:/</td>
<td>54.7(±8.8)</td>
</tr>
<tr>
<td>cough</td>
<td>77.3(±6.1)</td>
</tr>
<tr>
<td>/i:/</td>
<td>74.3(±3.2)</td>
</tr>
<tr>
<td>/ou/</td>
<td>72.3(±10.7)</td>
</tr>
<tr>
<td>/u:/</td>
<td>74.5(±4.3)</td>
</tr>
<tr>
<td>wheeze</td>
<td>89.8(±8.2)</td>
</tr>
<tr>
<td>/eI/</td>
<td>69.4(±12.8)</td>
</tr>
</tbody>
</table>

1 Which stimuli (cough, wheeze and sustained phonation) are better?
- **Wheeze is the best stimuli** for classification.
- **Sustained /i:/** performs the best among all sustained vowels.
Classification Results

Table: Mean (Standard deviation) of Total classification accuracy (TCA%) using static, ∆ (velocity) and ∆∆ (acceleration) coefficients.

<table>
<thead>
<tr>
<th>Stimuli</th>
<th>Total classification accuracy (TCA) %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K = 12</td>
</tr>
<tr>
<td>/ɑː/</td>
<td>54.7(±8.8)</td>
</tr>
<tr>
<td>cough</td>
<td>77.3(±6.1)</td>
</tr>
<tr>
<td>/ɪː/</td>
<td>74.3(±3.2)</td>
</tr>
<tr>
<td>/ou/</td>
<td>72.3(±10.7)</td>
</tr>
<tr>
<td>/uː/</td>
<td>74.5(±4.3)</td>
</tr>
<tr>
<td>wheeze</td>
<td>89.8(±8.2)</td>
</tr>
<tr>
<td>/ei/</td>
<td>69.4(±12.8)</td>
</tr>
</tbody>
</table>

1. Which stimuli (cough, wheeze and sustained phonation) are better?
   - Wheeze is the best stimuli for classification.
   - Sustained /ɪː/ performs the best among all sustained vowels.

2. What is the effect of including ∆ and ∆∆?
   - Stimulus dependent.
To what degree do individual MFCCs encode asthma related information for the best performing stimuli?

- 3rd, 7th and 12th Static MFCC capture cues of asthma better than ∆ and ∆∆ coefficients.
4 Which among the 6 statistics are more discriminant in best performing stimuli for classification task?

- **Mean, Median, Mode and RMS.**
- Fisher discriminant ratio (FDR) was used.
# Table of Contents

1. Introduction
2. Motivation
3. Dataset
4. Proposed Method
5. Experimental Setup
6. Results
7. Conclusion and Future work
Conclusion and Future work

1. **Wheeze is the best stimuli** for classification and **sustained /ɪ:/** performs the best among all sustained vowels.
1. **Wheeze is the best stimuli** for classification and **sustained /ɪ:/** performs the best among all sustained vowels.

2. **3rd, 7th and 12th MFCC of wheeze** signal encodes the asthma specific signature compared to those for velocity and acceleration coefficients.
Conclusion and Future work

1. **Wheeze is the best stimuli** for classification and **sustained /ɪ:/** performs the best among all sustained vowels.

2. **3rd, 7th and 12th MFCC of wheeze** signal encodes the asthma specific signature compared to those for velocity and acceleration coefficients.

3. **FDR values shown suggests** that the **mean, mode, median and RMS** statistics are relatively more discriminative compared to the variance and SD statistics.
1. *Wheeze is the best stimuli* for classification and **sustained /ɪː/** performs the best among all sustained vowels.

2. **3rd, 7th and 12th MFCC of wheeze** signal encodes the asthma specific signature compared to those for velocity and acceleration coefficients.

3. FDR values shown suggests that the **mean, mode, median and RMS** statistics are relatively more discriminative compared to the variance and SD statistics.

4. As the best performing stimuli is wheeze where there is no voicing, future plan includes investigation of **fricatives** as stimuli for asthma classification task.
THANK YOU
Mel-frequency cepstral coefficients (MFCC)

Speech → Framing and Windowing → FFT → Mel Filter Bank → MFCC Features → DCT → Logarithm
Conclusion and Future work

Support vector machine (SVM)