IMPACT OF MICROPHONE ARRAY CONFIGURATIONS ON ROBUST INDIRECT 3D ACOUSTIC SOURCE LOCALIZATION

Elizabeth Vargas, Keith Brown, Kartic Subr

ICASSP 2018
1. Record acoustic signals using a microphone array
2. Calculate time difference of arrivals (TDOA)
3. Calculate the position using geometric formulae (optimization)
Locating The Source In 3D

Using a microphone array, we calculate Time Difference of Arrivals (TDOA) before we can estimate the acoustic source location.

**Steered Response Power (SRP)**
- Most likely position amongst a grid of candidate locations
  - ✔ Accurate
  - ✗ Slow
- Probability of the sound source being in this location

**Multilateration**
- Infer the source position via least squares optimization
  - ✔ Fast
  - ✗ Non-convex function, local minima
QUESTION #1: Can localization be accurate and fast at the same time?
Microphone Arrays For Acoustic Source Localization

- 72 microphones
- Sampled at 192 kHz
- Three configurations spanning the same area
QUESTION #2: How does the microphone configuration affect localization accuracy?
Simulated Source Locations

Noise added to Time Difference of Arrivals (TDOA)

\[ \eta \sim \mathcal{N} \left( 0, \frac{\sigma \|s - O\|}{c} \right) \]

Localization Relative Error

\[ \text{error(\%)} = \frac{\|s - \tilde{s}\|}{\|s - O\|} \times 100 \]
Wheel and Spiral Configurations Are More Robust

- Ring
- Wheel
- Spiral
Simulated Source Locations

Noise added to Time Difference of Arrivals (TDOA)

\[ \eta \sim \mathcal{N} \left( 0, \frac{\sigma}{100} \frac{\| s - O \|}{c} \right) \]

Localization Relative Error

\[ \text{error(\%)} = \frac{\| s - \tilde{s} \|}{\| s - O \|} \times 100 \]
Higher Errors Observed With Ring Configuration

100% noise for a 2m by 2m room with 3 different configurations spanning the same area
Using Real Data

We tested both Multilateration and Steered Response Power (SRP)
Short Range Localization

Ring

Wheel

Spiral

A: (2.0, -0.32, 0.5)

Multiilateration

A

SRP
Mid Range Localization

B: (1.5, -0.32, 2.0)

Wheel
Spiral

Multilateration

SRP

Localization Error (%)
Facing The Microphone Array

Ring

Wheel

Spiral

C: (0.0, -0.32, 1.5)

Multilateration

SRP

Localization Error (%)

Localization Error (%)
Mid Range Localization

D: (-1.5, -0.32, 1.0)

Wheel

Spiral

Multilateration

SRP
What Happened To The Speed?

Using a microphone array, we calculate Time Difference of Arrivals (TDOA) before we can estimate the acoustic source location.

**Steered Response Power (SRP)**

- Most likely position amongst a grid of candidate locations
  - ✓ Accurate
  - ✗ Slow

**Multilateration**

- Infer the source position via least squares optimization
  - ✓ Fast
  - ✗ Non-convex function, local minima

Probability of the sound source being in this location
How Many Microphone Pairs To Use?

Array 16 Microphones: 120 Pairs
Array 32 Microphones: 496 Pairs
Array 72 Microphones: 2556 Pairs

72 microphones
Using 2556 Microphone Pairs

<table>
<thead>
<tr>
<th>Signal</th>
<th>Accuracy (%)</th>
<th>Time (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SRP</td>
<td>Multilateration</td>
</tr>
<tr>
<td>Chirp</td>
<td>14.7 (25.9)</td>
<td>12.1 (23.2)</td>
</tr>
<tr>
<td>Gunshot</td>
<td>11.0 (13.3)</td>
<td>6.4 (3.5)</td>
</tr>
<tr>
<td>Dogbark</td>
<td>16.0 (28.5)</td>
<td>48.5 (44.6)</td>
</tr>
<tr>
<td>Speech</td>
<td>13.2 (21.1)</td>
<td>12.9 (22.5)</td>
</tr>
</tbody>
</table>
# Using 100 Microphone Pairs

## Accuracy (%)

<table>
<thead>
<tr>
<th>Signal</th>
<th>SRP</th>
<th>Multilateration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chirp</td>
<td>14.7 (25.9)</td>
<td>14.2 (25.9)</td>
</tr>
<tr>
<td>Gunshot</td>
<td>11.0 (13.3)</td>
<td>9.6 (12.8)</td>
</tr>
<tr>
<td>Dogbark</td>
<td>16.0 (28.5)</td>
<td>58.9 (38.8)</td>
</tr>
<tr>
<td>Speech</td>
<td>13.2 (21.1)</td>
<td>15.2 (23.5)</td>
</tr>
</tbody>
</table>

## Time (minutes)

<table>
<thead>
<tr>
<th>Signal</th>
<th>SRP</th>
<th>Multilateration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chirp</td>
<td>3 (0.2)</td>
<td>0.5 (0.01)</td>
</tr>
<tr>
<td>Gunshot</td>
<td>2.58 (0.2)</td>
<td>0.4 (0.02)</td>
</tr>
<tr>
<td>Dogbark</td>
<td>2.49 (0.1)</td>
<td>0.4 (0.02)</td>
</tr>
<tr>
<td>Speech</td>
<td>2.63 (0.1)</td>
<td>0.4 (0.02)</td>
</tr>
</tbody>
</table>
Conclusions

**QUESTION #1:**
Can localization be accurate and fast at the same time?
Yes it can! Direct optimization yields errors similar to the Steered Response Power (SRP) method with 6 times less computation.

**QUESTION #2:**
How does the microphone configuration affects the localization accuracy?
Circular arrays are the least desirable configuration.