Array Processing in 3D Sound Fields

Boaz Rafaely

Department of Electrical and Computer Engineering
Ben-Gurion University of the Negev
Beer-Sheva 84105, Israel

3rd Underwater Acoustics Symposium, Tel-Aviv University, 19 June 2014
Content

- Context and motivation
- Environment - sound in enclosed spaces
- Sensing - spherical arrays
- Direction-of-arrival estimation
- Simulation and experimental investigations
Motivation

- **DOA estimation of speakers in a room**
  - Speech communication - beamformer steering
  - Humanoid robots - head steering
  - Video conferencing - camera steering
The environment - room acoustics

- Schroeder frequency \( f_c = 2000 \sqrt{\frac{T_{60}}{V}} \)
  \( \rightarrow \) Diffuse sound field in the audio frequency range
The environment - room acoustics

- Schroeder frequency $f_c = 2000 \sqrt{\frac{T_{60}}{V}}$
  $\rightarrow$ Diffuse sound field in the audio frequency range

- Diffuse field power equation: $\frac{A c \epsilon}{4} + V \frac{d \epsilon}{dt} = W$
  $\rightarrow$ Critical distance $r_c = \sqrt{\frac{A}{16\pi}}$
  $\rightarrow$ Reverberation time $T_{60} = 0.161 \frac{V}{A}$
The environment - room acoustics

- Schroeder frequency $f_c = 2000 \sqrt{\frac{T_{60}}{V}}$
  $\rightarrow$ Diffuse sound field in the audio frequency range

- Diffuse field power equation: $\frac{A\epsilon}{4} + V \frac{d\epsilon}{dt} = W$
  $\rightarrow$ Critical distance $r_c = \sqrt{\frac{A}{16\pi}}$
  $\rightarrow$ Reverberation time $T_{60} = 0.161 \frac{V}{A}$

- Large room: $V = 1000$ m$^3$, $T_{60} = 1$ s, $r_c = 1.8$ m
  $\rightarrow$ At $r > r_c$ significant diffuse field vs. direct sound!
Recording sound in rooms

- Spherical arrays
- Recording of 3D sound
- Elegant formulation: *signal processing with acoustic theory*
Spherical Fourier transform

- Spherical Fourier transform, \( f(\theta, \phi) \in L_2(S^2) \),

\[
\begin{align*}
  f_{nm} &= \int_0^{2\pi} \int_0^{\pi} f(\theta, \phi) [Y_n^m(\theta, \phi)]^* \sin \theta \, d\theta \, d\phi \\
  f(\theta, \phi) &= \sum_{n=0}^{\infty} \sum_{m=-n}^{n} f_{nm} Y_n^m(\theta, \phi)
\end{align*}
\]
Spherical Fourier transform

- Spherical Fourier transform, \( f(\theta, \phi) \in L_2(S^2) \),

\[
f_{nm} = \int_{0}^{2\pi} \int_{0}^{\pi} f(\theta, \phi) \left[ Y_n^m(\theta, \phi) \right]^* \sin \theta \, d\theta \, d\phi
\]

\[
f(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} f_{nm} Y_n^m(\theta, \phi)
\]

- Orthogonality and completeness,

\[
\int_{0}^{2\pi} \int_{0}^{\pi} Y_n^m(\theta, \phi) \left[ Y_{n'}^{m'}(\theta, \phi) \right]^* \sin \theta \, d\theta \, d\phi = \delta_{nn'} \delta_{mm'}
\]

\[
\sum_{n=0}^{\infty} \sum_{m=-n}^{n} Y_n^m(\theta, \phi) \left[ Y_n^m(\theta', \phi') \right]^* = \delta(\cos \theta - \cos \theta') \delta(\phi - \phi')
\]
Spherical harmonics

\[ Y_n^m(\theta, \phi) = \sqrt{\frac{2n + 1}{4\pi}} \frac{(n - m)!}{(n + m)!} P_n^m(\cos \theta) e^{im\phi} \]
The wave equation

- The wave Eq. for \( p(r, \theta, \phi, t) \) can be written as:

\[
\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} p \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} p \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} p - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} p = 0
\]
The wave equation

- The wave Eq. for \( p(r, \theta, \phi, t) \) can be written as:
  \[
  \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} p \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} p \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} p - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} p = 0
  \]

- Assume separation of variables:
  \[ p(r, \theta, \phi, t) = R(r) \Theta(\theta) \Phi(\phi) T(t) \]

- Typical Solutions for \( r, \theta, \phi, t \):
  - \( e^{i\omega t}, \) time.
  - \( Y_n^m(\theta, \phi), \) standing waves in \( \theta \), traveling waves in \( \phi \).
  - \( j_n(kr) \) or \( h_n(kr) \), radial dependence.
Array processing

- Array output as weighted sound pressure:

\[
y(k) = \int_0^{2\pi} \int_0^{\pi} p(k, r, \theta, \phi) w^*(k, \theta, \phi) \sin \theta \, d\theta \, d\phi
\]

- In practice, sampled pressure is available,

\[
y = w^H p
\]

- Beamformer now designed, e.g. MVDR,

\[
\min_w E \left[ |y|^2 \right], \quad s.t. \quad w^H v = 1
\]
Maximum-directivity \( w_{nm}^* = [Y_n^m(\theta_l, \phi_l)]^* / b_n(kr) \)
Direct-path dominance test DOA estimation
Direct-path dominance test DOA estimation

Microphone Input signals

SFT  STFT

P WD spectrum

Time averaging

Frequency averaging

Direct-path dominance test

Info. Fusion and DOA estimation
Direct-path dominance test DOA estimation

1. Short-time Fourier transform (STFT)
   - SFT

2. Pseudo-Wigner Distribution (PWD) spectrum
   - Time averaging

3. Frequency averaging

4. Direct-path dominance test

5. Info. Fusion and DOA estimation
Direct-path dominance test DOA estimation

(Spatial) Spherical Fourier transform

X → SFT → PWD spectrum → Time averaging → Frequency averaging → Direct-path dominance test → Info. Fusion and DOA estimation → Ω
Direct-path dominance test DOA estimation

1. X
2. SFT
3. STFT
4. PWD spectrum
5. Time averaging
6. Frequency averaging
7. Direct-path dominance test
8. Info. Fusion and DOA estimation
9. Ω

Spatial spectrum matrix in the plane-wave domain
Direct-path dominance test DOA estimation

1. SFT
2. STFT
3. PWD spectrum
4. Time averaging
5. Frequency averaging
6. Direct-path dominance test
7. Info. Fusion and DOA estimation
Direct-path dominance test DOA estimation

\[ x \]

- SFT
- STFT

\[ \text{PWR spectrum} \]

\[ \text{Time averaging} \]

\[ \text{Frequency averaging} \]

\[ \text{Direct-path dominance test} \]

\[ \text{Info. Fusion and DOA estimation} \]

\[ \Omega \]

Over adjacent frequencies to unfold coherent signals (room reflections)
Direct-path dominance test DOA estimation

- SFT
- STFT
- PWD spectrum
- Time averaging
- Frequency averaging
- Direct-path dominance test
- Info. Fusion and DOA estimation
- Rank ONE spectral matrices
Direct-path dominance test DOA estimation

- SFT
- STFT
- PWD spectrum
- Time averaging
- Frequency averaging
- Direct-path dominance test
- Info. Fusion and DOA estimation

Fuse local (time, frequency) estimates
Summary of equations

- Array equation (space)
  \[ p(k) = Y(\Omega)B(kr)Vs(k) + n(k) \]

- Array equation (SH)
  \[ a_{nm}(k) = Y^H(\Psi)s(k) + \tilde{n}(k) \]

- In time-frequency
  \[ a_{nm}(\tau, \nu) = Y^H(\Psi)s(\tau, \nu) + \tilde{n}(\tau, \nu) \]

- Spatial spectrum, time/frequency smoothing
  \[ \tilde{R}_a(\tau, \nu) = Y^H(\Psi)\tilde{R}_s(\tau, \nu)V(\Psi) + \tilde{R}_n(\nu) \]

- Direct-path dominance test
  \[ (\tau, \nu) : \text{erank} \left( \tilde{R}_a(\tau, \nu) \right) = 1 \]

- MUSIC spectrum
  \[ S_{MUSIC}(\Omega) = \frac{1}{y^T(\Omega)(I - UsU_s^H)y^*(\Omega)} \]

- Clustering and DOA estimation
Simulation study

Simulated room $6 \times 4 \times 3$ m, 3 or 4 speakers 1.5 m away, 32 microphones, varying $T_{60}$, SNR=25 dB

Nadiri and Rafaely, Localization of Multiple Speakers under High Reverberation using a Spherical Microphone Array and the Direct-Path Dominance Test, IEEE TASLP, minor revision.
Simulation study

Simulated room $6 \times 4 \times 3$ m, 3 or 4 speakers 1.5 m away, 32 microphones, varying $T_{60}$, SNR=25 dB

Nadiri and Rafaely, Localization of Multiple Speakers under High Reverberation using a Spherical Microphone Array and the Direct-Path Dominance Test, IEEE TASLP, minor revision.
Experimental study

- Real room $8.7 \times 5.9 \times 3.2$ m, 3 speakers 2 m away, 32 microphones, $T_{60} = 1.3$ s
- Recording of the speakers

<table>
<thead>
<tr>
<th>Table: DOA estimation error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Speaker #1</td>
</tr>
<tr>
<td>Speaker #2</td>
</tr>
<tr>
<td>Speaker #3</td>
</tr>
<tr>
<td>RMSE</td>
</tr>
</tbody>
</table>

Nadiri and Rafaely, Localization of Multiple Speakers under High Reverberation using a Spherical Microphone Array and the Direct-Path Dominance Test, IEEE TASLP, minor revision.
Experimental study

- Spatial spectrum: Broadband MUSIC

Nadiri and Rafaely, Localization of Multiple Speakers under High Reverberation using a Spherical Microphone Array and the Direct-Path Dominance Test, IEEE TASLP, minor revision.
Experimental study

- Spatial spectrum: Coherence test

Nadiri and Rafaely, Localization of Multiple Speakers under High Reverberation using a Spherical Microphone Array and the Direct-Path Dominance Test, IEEE TASLP, minor revision.
Experimental study

- Spatial spectrum: Direct-path dominance test 1

Nadiri and Rafaely, Localization of Multiple Speakers under High Reverberation using a Spherical Microphone Array and the Direct-Path Dominance Test, IEEE TASLP, minor revision.
Experimental study

- Spatial spectrum: Direct-path dominance test 2

Nadiri and Rafaely, Localization of Multiple Speakers under High Reverberation using a Spherical Microphone Array and the Direct-Path Dominance Test, IEEE TASLP, minor revision.
Conclusion

- Spherical microphone array - 3D sound capture and processing
- Spherical harmonics - separation of space and frequency
- Leading to accurate DOA estimation of multiple sources under high reverberation
- Application to underwater acoustic signal processing?