A Comparative Performance Study of Sound Zoning Methods in a Reflective Environment

Marek Olik¹, Jon Francombe², Philip Coleman¹, Philip JB Jackson¹, Martin Olsen³, Martin Møller³, Russell Mason² and Søren Bech³

¹Centre for Vision, Speech and Signal Processing, University of Surrey
²Institute of Sound Recording, University of Surrey
³Bang & Olufsen A/S

AES 52nd International Conference, Sound Field Control
Paper Session 5: Sound Zones
3rd September 2013
Sound Zone Problem

Bright Zone (A)

Dark Zone (B)
## Motivation

- **Approaches to sound field control**

<table>
<thead>
<tr>
<th>Sound Focusing</th>
<th>Sound Cancelling</th>
<th>Sound Field Synthesis</th>
</tr>
</thead>
</table>
Motivation

• **Approaches to sound field control**

<table>
<thead>
<tr>
<th>Sound Focusing</th>
<th>Sound Cancelling</th>
<th>Sound Field Synthesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delay and Sum</td>
<td>Acoustic Contrast Control</td>
<td>Analytical SFS</td>
</tr>
<tr>
<td>[Veen and Buckley 1988]</td>
<td>[Choi &amp; Kim 2002]</td>
<td>[Wu and Abhayapala 2010]</td>
</tr>
<tr>
<td>Brightness Control</td>
<td>Acoustic Energy Difference Maximisation</td>
<td>Pressure Matching</td>
</tr>
<tr>
<td>[Choi &amp; Kim 2002]</td>
<td>[Shin et al. 2010]</td>
<td>[Poletti 2008]</td>
</tr>
</tbody>
</table>

• There are limited examples of comparative evaluation of these methods [Jacobsen *et al.* 2011, Møller *et al.* 2012, Coleman *et al.* 2013]

• Also, limited examples of studies investigating performance of these methods under reflective conditions [Elliott and Jones 2006, Jacobsen *et al.* 2011]
Scope

• We compare the performance of key sound zone methods implemented in a room
• One representative method from each group: DS, ACC and PM
• Two loudspeaker arrays: linear and circular
Scope

- We compare the performance of key sound zone methods implemented in a room
- One representative method from each group: DS, ACC and PM
- Two loudspeaker arrays: linear and circular
- Evaluation based on physical metrics:
  - Acoustic separation between zones (acoustic contrast)
  - Characteristics of the bright zone sound field (planarity)
- Perceptual evaluation: listening tests based on recordings in the zones to evaluate distraction from the interfering audio
Outline

1. Evaluation metrics and sound zone methods
2. Experimental setup and sound zone reproduction procedure
3. Results based on physical metrics
4. Perceptual evaluation: distraction
5. Summary and further work
Physical Evaluation Metrics

To evaluate acoustic separation between the zones:

- **Acoustic contrast:** ratio between average squared pressures in the bright and dark zones. For equal number of microphones in zone A and B:

\[
C_{AB} = 10 \log_{10} \left( \frac{p_A^H p_A}{p_B^H p_B} \right)
\]
Physical Evaluation Metrics

To evaluate acoustic separation between the zones:
- **Acoustic contrast**: ratio between average squared pressures in the bright and dark zones. For equal number of microphones in zone A and B:

\[
C_{AB} = 10 \log_{10} \left( \frac{p_A^H p_A}{p_B^H p_B} \right)
\]

To evaluate spatial characteristics of bright zone sound field:
- **Planarity** [Jackson *et al.* (2013)]: ratio of energy in the direction of principal plane wave component to total energy in the bright zone

\[
\eta_A = \frac{\sum_i w_i u_i \cdot u_\alpha}{\sum_i w_i}
\]

\[i = 1 \ldots N, \text{ where } N \text{ is the number of angles}\]

\[\alpha = \arg \max_i w_i\]
Sound Focusing: Delay and Sum Beamforming

\[ q_{DS} = [e^{-j\omega \tau_1} \quad e^{-j\omega \tau_2} \quad \ldots \quad e^{-j\omega \tau_L}]^T \]

\( \tau_1, \tau_2, \ldots, \tau_L \) - time delays with respect to the largest distance to bright zone
Control Methods

Sound Focusing: Delay and Sum Beamforming

\[ q_{DS} = [e^{-j\omega\tau_1} \ e^{-j\omega\tau_2} \ \ldots \ e^{-j\omega\tau_L}]^T \]

\( \tau_1, \tau_2, \ldots, \tau_L \) - time delays with respect to the largest distance to bright zone

Sound Cancelling: Acoustic Contrast Control

\[ J_{ACC} = \frac{p_A^H p_A}{p_B^H p_B} = \frac{q^H G_A^H G_A q}{q^H G_B^H G_B q} \]

\( q_{ACC} \) is proportional to eigenvector of the matrix \([G_B^H G_B]^{-1} G_A^H G_A\) that corresponds to largest eigenvalue. With regularisation: \([G_B^H G_B + \lambda I]^{-1} G_A^H G_A\)
Control Methods

Sound Focusing: Delay and Sum Beamforming

\[ q_{DS} = [e^{-j\omega\tau_1} \ e^{-j\omega\tau_2} \ \ldots \ e^{-j\omega\tau_L}]^T \]

\( \tau_1, \tau_2, \ldots, \tau_L \) - time delays with respect to the largest distance to bright zone

Sound Cancelling: Acoustic Contrast Control

\[ J_{ACC} = \frac{p_A^H p_A}{p_B^H p_B} = \frac{q_A^H G_A^H G_A q}{q_B^H G_B^H G_B q} \]

\( q_{ACC} \) is proportional to eigenvector of the matrix \([G_B^H G_B]^{-1} G_A^H G_A\) that corresponds to largest eigenvalue. With regularisation: \([G_B^H G_B + \lambda I]^{-1} G_A^H G_A\)

Sound Field Synthesis: Pressure Matching

\[ J_{PM} = e^H e + \lambda q^H q \]

\( e \) – error between the desired sound field \( d \) and reproduced sound field

\( q^H q \) – power constraint

\( \lambda \) – regularisation parameter

\[ q_{PM} = (G^H G + \lambda I)^{-1} G^H d \]
Experimental Setup

Overview
Experimental Setup

Details

\[ V \approx 320 \, m^3 \]
Experimental Setup

Details

\[ V \approx 320\ m^3 \]

- **Line array**: units
- **Circular array**: unit
- **Room boundary**: varied ceiling height
- **Loudspeakers**: every 15°
- **Dark zone**: 
- **Bright zone**: 
- **Zone location**: 
- **Line array units**: 

**Dimensions**

- Length: 9.25m
- Width: 5.82m

**Coordinates**

- Origin: (0,0)
- (0, 0.53m)
- (1.47m, 0.53m)
- (1.75m, 0.53m)
- (1.48m, 1.51m)
- (1.62m, 1.51m)
- (1.20m, 0.53m)
Experimental Setup

Details

\[ V \approx 320 \, m^3 \]

<table>
<thead>
<tr>
<th>Centre Frequency (Hz)</th>
<th>RT60 (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>250</td>
<td>0.33</td>
</tr>
<tr>
<td>500</td>
<td>0.30</td>
</tr>
<tr>
<td>1000</td>
<td>0.28</td>
</tr>
<tr>
<td>2000</td>
<td>0.27</td>
</tr>
<tr>
<td>4000</td>
<td>0.25</td>
</tr>
</tbody>
</table>
Experimental Setup

Details

$V \approx 320 \text{ m}^3$

**Table:**

<table>
<thead>
<tr>
<th>Centre Frequency (Hz)</th>
<th>RT60 (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>250</td>
<td>0.33</td>
</tr>
<tr>
<td>500</td>
<td>0.30</td>
</tr>
<tr>
<td>1000</td>
<td>0.28</td>
</tr>
<tr>
<td>2000</td>
<td>0.27</td>
</tr>
<tr>
<td>4000</td>
<td>0.25</td>
</tr>
</tbody>
</table>

- Setup: 72 mics per zone
- Monitoring: 144 mics per zone
- Numerically independent responses included in performance predictions to reduce bias [Akeroyd et al. 2007]
System Diagram

- Estimation of speed of sound
- Estimation of geometry
- Delay & Sum

$q'(f)$
System Diagram

Response Measurement (MLS)

Pre-Processing

FFT

Estimation of speed of sound

Estimation of geometry

Delay & Sum

Acoustic Contrast Control

Pressure Matching

Response Measurement (MLS)

Pre-Processing

FFT

Estimation of speed of sound

Estimation of geometry

Delay & Sum

Acoustic Contrast Control

Pressure Matching

\[ h(t) \]

\[ h'(t) \]

\[ G(f) \]

\[ q'(f) \]

\[ g(t) \]

\[ g'(t) \]
System Diagram

Response Measurement (MLS) → Pre-Processing → FFT → G(f) → Estimation of speed of sound

Estimation of geometry

Acoustic Contrast Control

Pressure Matching

Delay & Sum

Estimation of geometry

h'(t) → h(t) → G(f) → q'(f) → q(f)

0.3-3.5 kHz

ls
System Diagram

Response Measurement (MLS)

Pre-Processing

FFT

Estimation of speed of sound

Acoustic Contrast Control

Pressure Matching

Delay & Sum

Estimation of geometry

\[ h'(t) \xrightarrow{\text{ls}} h(t) \xrightarrow{\text{mic}_s} G(f) \xrightarrow{\text{ls}} q'(f) \xrightarrow{\text{ls}} q(f) \xrightarrow{\text{mic}_s} p_p(f) = G(f) q(f) \]

0.3-3.5 kHz
System Diagram

Response Measurement (MLS)

Pre-Processing

FFT

Estimation of speed of sound

Estimation of geometry

Acoustic Contrast Control

Pressure Matching

Delay & Sum

0.3-3.5 kHz

Play/Record

FFT

Measured Scores

Predicted Scores

\[ p_m(t) \]

\[ s(t) \]

\[ q(t) \]

\[ q'(f) \]

\[ p_m(f) \]

\[ p_p(f) = G(f)q(f) \]
Physical Evaluation: Contrast

Reference line: 20 dB

Delay & Sum
\langle\text{contrast}\rangle_f = 16 \text{ dB}

Acoustic Contrast Control
\langle\text{contrast}\rangle_f = 17 \text{ dB}
Physical Evaluation: Contrast

Delay & Sum
\( \langle \text{contrast} \rangle_f = 16 \text{ dB} \)

Acoustic Contrast Control
\( \langle \text{contrast} \rangle_f = 17 \text{ dB} \)

Acoustic Contrast Control
\( \langle \text{contrast} \rangle_f = 12 \text{ dB} \)

Pressure Matching
\( \langle \text{contrast} \rangle_f = 4 \text{ dB} \)

reference line: 20dB
Physical Evaluation: Planarity

Delay & Sum
\(\langle \text{planarity} \rangle_f = 83\%\)

Acoustic Contrast Control
\(\langle \text{planarity} \rangle_f = 84\%\)

Reference line: 80%
Physical Evaluation: Planarity

**Delay & Sum**

\[ \langle \text{planarity} \rangle_f = 83 \% \]

**Acoustic Contrast Control**

\[ \langle \text{planarity} \rangle_f = 84 \% \]

\[ \langle \text{planarity} \rangle_f = 62 \% \]

**Pressure Matching**

\[ \langle \text{planarity} \rangle_f = 76 \% \]

Reference line: 80%
Zone A Pressure Maps: 1.5 kHz

Measured Data

Planarity (1.5 kHz)

DS (Line)

ACC (Line)

ACC (Circle)

PM (Circle)

90%

85%

66%

89%
Perceptual Evaluation
Stimulus Creation

- Listening tests based on recordings in the zones

**Diagram Description**

- **Target programme**
  - $*q(t)$
  - Filtered Target
  - Play in Zone A
  - Record
  - Auralisation of target

- **Interferer programme**
  - $*q(t)$
  - Filtered Interferer
  - Play in Zone B
  - Record
  - Auralisation of interferer

+ Combination

**Output**
Perceptual Evaluation
Stimulus Creation

- Listening tests based on recordings in the zones

- Mono stimuli: we can’t investigate perceptual effect of spatial variation – but we can begin to observe the perceptual effects of contrast and programme material
Perceptual Evaluation
Experiment Design

- 24 stimuli (combination of methods and programme items)
- Rate ‘distraction’ [Francombe et al. 2013]
Perceptual Evaluation
Experiment Design

• 24 stimuli (combination of methods and programme items)
• Rate ‘distraction’ [Francombe et al. 2013]
Perceptual Evaluation

Results

- ANOVA model – most significant factor: sound zone method
Perceptual Evaluation

Results

• **ANOVA model** – most significant factor: sound zone method

• **Other significant factors** influenced distraction: programme material and interactions
Summary and Further Work

• A comparative study of sound zone methods implemented in a room
• Methods: Delay and Sum, Acoustic Contrast and Pressure Matching
• Arrays: linear and circular (each 24 elements)
Summary and Further Work

• A comparative study of sound zone methods implemented in a room
• **Methods:** Delay and Sum, Acoustic Contrast and Pressure Matching
• **Arrays:** linear and circular (each 24 elements)

• **Contrast:** highest overall contrast from ACC; DS and PM were more sensitive to limitations of the array geometries
• **Planarity:** line array - high planarity; PM improved planarity at low frequencies; ACC circle – inhomogeneous bright zone, lowest planarity score

• Listening tests indicated that contrast may be perceptually important
Summary and Further Work

- A comparative study of sound zone methods implemented in a room
- **Methods:** Delay and Sum, Acoustic Contrast and Pressure Matching
- **Arrays:** linear and circular (each 24 elements)

- **Contrast:** highest overall contrast from ACC; DS and PM were more sensitive to limitations of the array geometries
- **Planarity:** line array - high planarity; PM improved planarity at low frequencies; ACC circle – inhomogeneous bright zone, lowest planarity score

- Listening tests indicated that contrast may be perceptually important

**Further work:**
- In situ listening tests – to evaluate effect of spatial differences
- Evaluation of sound quality in the bright zone (signal to distortion ratio, basic audio quality)
References

Appendix A
Pre-Processing: Impulse Smoothing

\[ n = \binom{N}{N-1} \ldots 2 1 \]

\[ l_n = \frac{f_s i}{f_n} \]

- \( l_n \): number of bins before \( f_n \)
- \( f_n \): cut-off freq. of LPF in \( n \)-th band
- \( f_s \): sampling frequency

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( i )</td>
<td>10</td>
</tr>
<tr>
<td>( f_s )</td>
<td>48 kHz</td>
</tr>
<tr>
<td>( N )</td>
<td>141</td>
</tr>
<tr>
<td>( f_1 )</td>
<td>8 Hz</td>
</tr>
<tr>
<td>( f_N )</td>
<td>2.376 kHz</td>
</tr>
<tr>
<td>Freq. spacing</td>
<td>1/12 Oct.</td>
</tr>
<tr>
<td>Slope</td>
<td>80 dB/Dec.</td>
</tr>
</tbody>
</table>
Appendix B
Contrast: Two Recording Points

reference line: 20dB

Delay & Sum
\( \langle \text{contrast} \rangle_f = 15 \text{ dB} \)

Acoustic Contrast Control
\( \langle \text{contrast} \rangle_f = 18 \text{ dB} \)

Acoustic Contrast Control
\( \langle \text{contrast} \rangle_f = 14 \text{ dB} \)

Pressure Matching
\( \langle \text{contrast} \rangle_f = 4 \text{ dB} \)

Line
Circle
Appendix C
Perceptual Evaluation: Details of Experiment Design

- **24 stimuli**
  - 2 target programmes: sports commentary, pop music
  - 3 interferer programmes: male speech, pop music, classical music
  - 4 sound zone methods (DS Line, ACC Line & Circle, PM Circle)

- **Rate ‘distraction’** [Francombe *et al.* 2013]
  - How much the alternate audio pulls your attention or distracts you from the target audio

- **Multiple stimulus presentation**
  - 3 test items per page (target programme and SZ method held constant)

- **Each page also had a reference**
  - Just the target audio with no interference
  - Hidden reference included should be scored at 0