UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

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

Marek Olik[1], Jon Francombe[2], Philip Coleman[1], Philip JB Jackson[1], Martin Olsen[3], Martin Møller[3], Russell Mason[2] and Søren Bech[3]

[1]Centre for Vision, Speech and Signal Processing, University of Surrey

[2]Institute of Sound Recording, University of Surrey

[3]Bang & Olufsen A/S

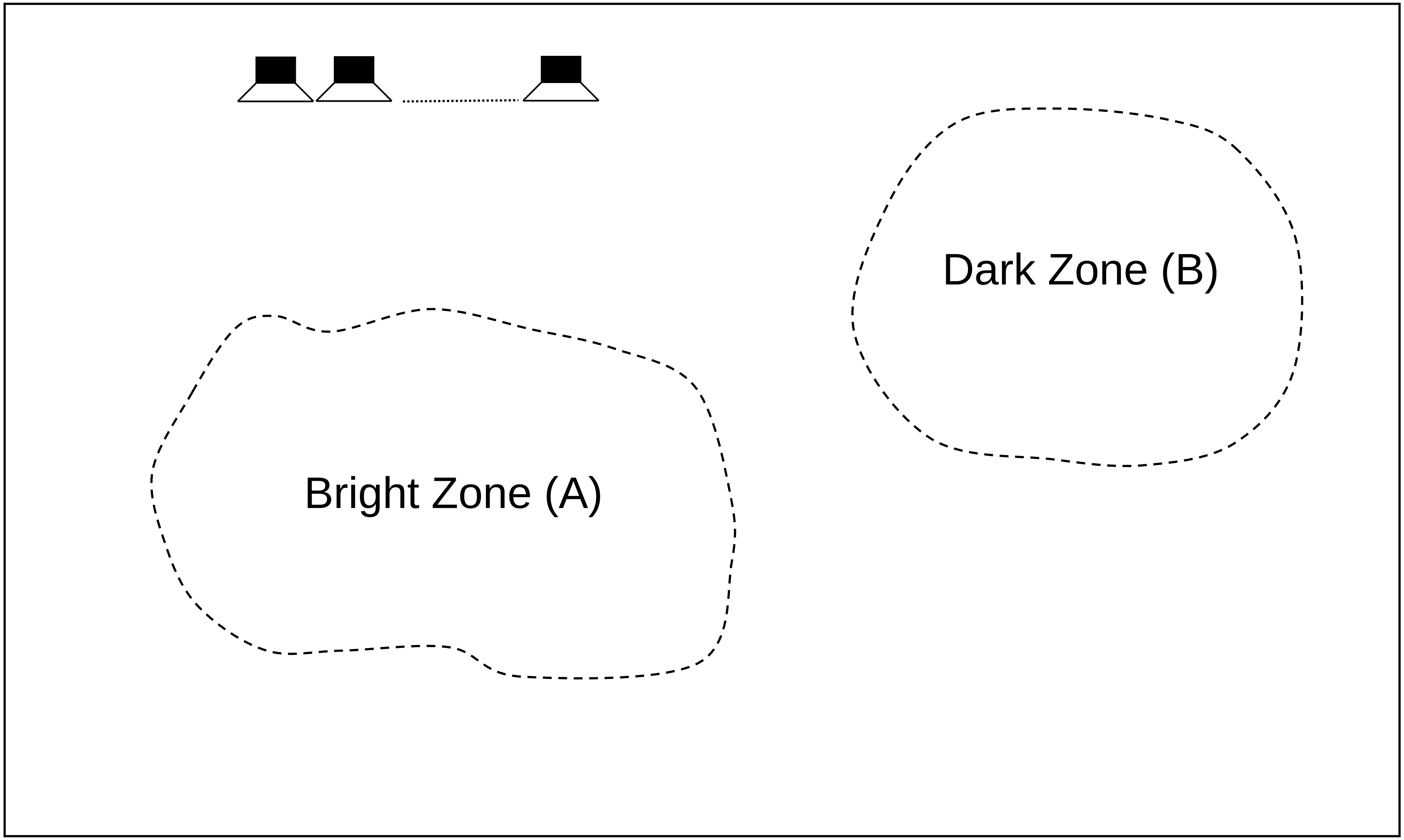
AES 52nd International Conference, Sound Field Control

Paper Session 5: Sound Zones

3rd September 2013

POSZ

Sound Zone Problem



Motivation

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

- **Approaches to sound field control**

Sound Focusing	Sound Cancelling	Sound Field Synthesis
Delay and Sum [Veen and Buckley 1988]	**Acoustic Contrast Control** [Choi & Kim 2002]	**Analytical SFS** [Wu and Abhayapala 2010]
Brightness Control [Choi & Kim 2002]	**Acoustic Energy Difference Maximisation** [Shin *et al*. 2010]	**Pressure Matching** [Poletti 2008]

Motivation

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

- **Approaches to sound field control**

Sound Focusing	Sound Cancelling	Sound Field Synthesis
Delay and Sum [Veen and Buckley 1988]	**Acoustic Contrast Control** [Choi & Kim 2002]	**Analytical SFS** [Wu and Abhayapala 2010]
Brightness Control [Choi & Kim 2002]	**Acoustic Energy Difference Maximisation** [Shin *et al*. 2010]	**Pressure Matching** [Poletti 2008]

- **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

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

- **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

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

- **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

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

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:**

$$C_{AB} = 10\log_{10}\left(\frac{\boldsymbol{p}_A^H \boldsymbol{p}_A}{\boldsymbol{p}_B^H \boldsymbol{p}_B}\right)$$

POSZ

Physical Evaluation Metrics

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

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:**

$$\mathrm{C_{AB}} = 10\log_{10}\left(\frac{\boldsymbol{p}_A^H \boldsymbol{p}_A}{\boldsymbol{p}_B^H \boldsymbol{p}_B}\right)$$

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**

$$\eta_\mathrm{A} = \frac{\sum_i w_i \boldsymbol{u}_i \cdot \boldsymbol{u}_\alpha}{\sum_i w_i}$$

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

$$\alpha = \mathbf{arg\,max}_\mathrm{i} w_i$$

POSZ

Control Methods

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

Sound Focusing: Delay and Sum Beamforming

$$\boldsymbol{q}_{DS} = [e^{-j\omega\tau_1} \quad e^{-j\omega\tau_2} \quad \cdots \quad e^{-j\omega\tau_L}]^T$$

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

Control Methods

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

Sound Focusing: Delay and Sum Beamforming

$$\boldsymbol{q}_{DS} = [e^{-j\omega\tau_1} \quad e^{-j\omega\tau_2} \quad \cdots \quad e^{-j\omega\tau_L}]^T$$

$\tau_1, \tau_2, \dots, \tau_L$ - time delays with respect to the largest distance to bright zone

Sound Cancelling: Acoustic Contrast Control

$$J_{ACC} = \frac{\boldsymbol{p}_A^H \boldsymbol{p}_A}{\boldsymbol{p}_B^H \boldsymbol{p}_B} = \frac{\boldsymbol{q}^H \boldsymbol{G}_A^H \boldsymbol{G}_A \boldsymbol{q}}{\boldsymbol{q}^H \boldsymbol{G}_B^H \boldsymbol{G}_B \boldsymbol{q}}$$

$\boldsymbol{q}_{ACC}$ is proportional to eigenvector of the matrix $[\boldsymbol{G}_B^H \boldsymbol{G}_B]^{-1} \boldsymbol{G}_A^H \boldsymbol{G}_A$ that corresponds to largest eigenvalue. With regularisation: $[\boldsymbol{G}_B^H \boldsymbol{G}_B + \lambda \boldsymbol{I}]^{-1} \boldsymbol{G}_A^H \boldsymbol{G}_A$

Control Methods

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

Sound Focusing: Delay and Sum Beamforming

$$\boldsymbol{q}_{DS} = [e^{-j\omega\tau_1} \quad e^{-j\omega\tau_2} \quad \dots \quad e^{-j\omega\tau_L}]^T$$

$\tau_1, \tau_2, \dots, \tau_L$ - time delays with respect to the largest distance to bright zone

Sound Cancelling: Acoustic Contrast Control

$$J_{ACC} = \frac{\boldsymbol{p}_A^H \boldsymbol{p}_A}{\boldsymbol{p}_B^H \boldsymbol{p}_B} = \frac{\boldsymbol{q}^H \boldsymbol{G}_A^H \boldsymbol{G}_A \boldsymbol{q}}{\boldsymbol{q}^H \boldsymbol{G}_B^H \boldsymbol{G}_B \boldsymbol{q}}$$

$\boldsymbol{q}_{ACC}$ is proportional to eigenvector of the matrix $[\boldsymbol{G}_B^H \boldsymbol{G}_B]^{-1} \boldsymbol{G}_A^H \boldsymbol{G}_A$ that corresponds to largest eigenvalue. With regularisation: $[\boldsymbol{G}_B^H \boldsymbol{G}_B + \lambda \boldsymbol{I}]^{-1} \boldsymbol{G}_A^H \boldsymbol{G}_A$

Sound Field Synthesis: Pressure Matching

$$J_{PM} = \boldsymbol{e}^H \boldsymbol{e} + \lambda \boldsymbol{q}^H \boldsymbol{q}$$

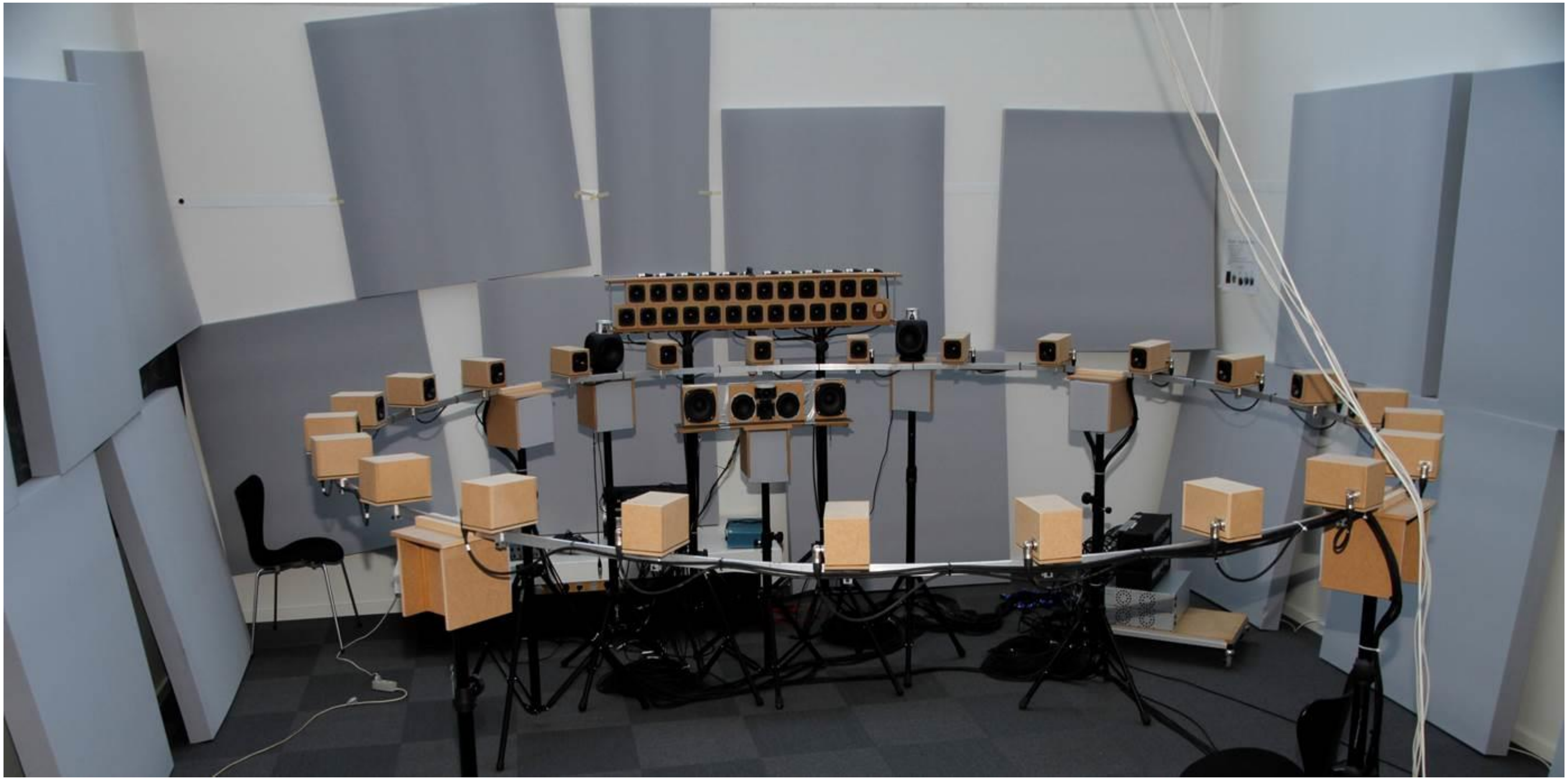
$\boldsymbol{e}$ – error between the desired sound field $\boldsymbol{d}$ and reproduced sound field

$\boldsymbol{q}^H \boldsymbol{q}$ – power constraint

λ – regularisation parameter

$$\boldsymbol{q}_{PM} = (\boldsymbol{G}^H \boldsymbol{G} + \lambda \boldsymbol{I})^{-1} \boldsymbol{G}^H \boldsymbol{d}$$

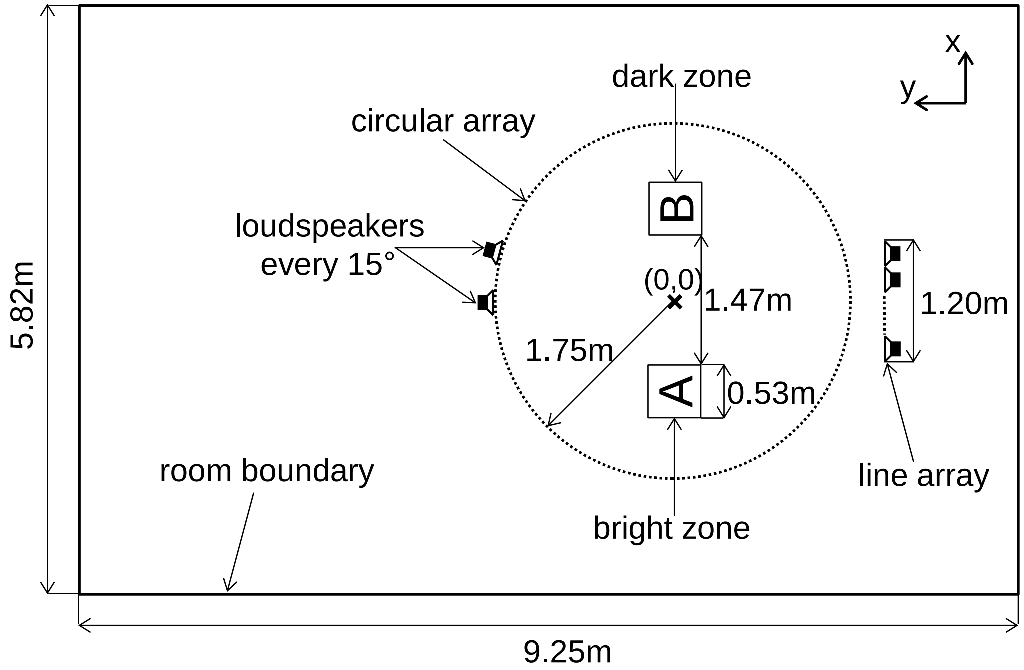
Experimental Setup
Overview



Experimental Setup

Details

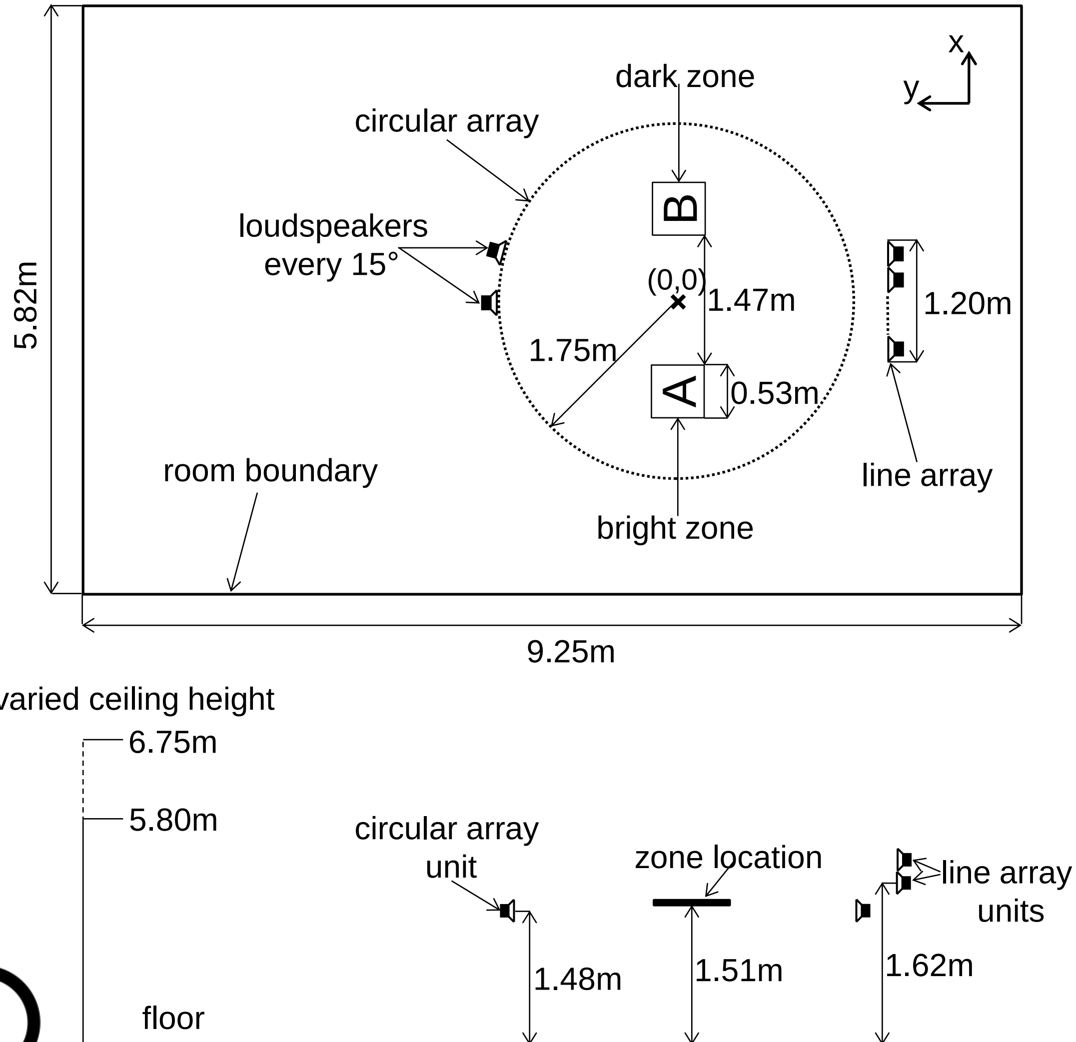
$V \approx 320\ m^3$



Experimental Setup

Details

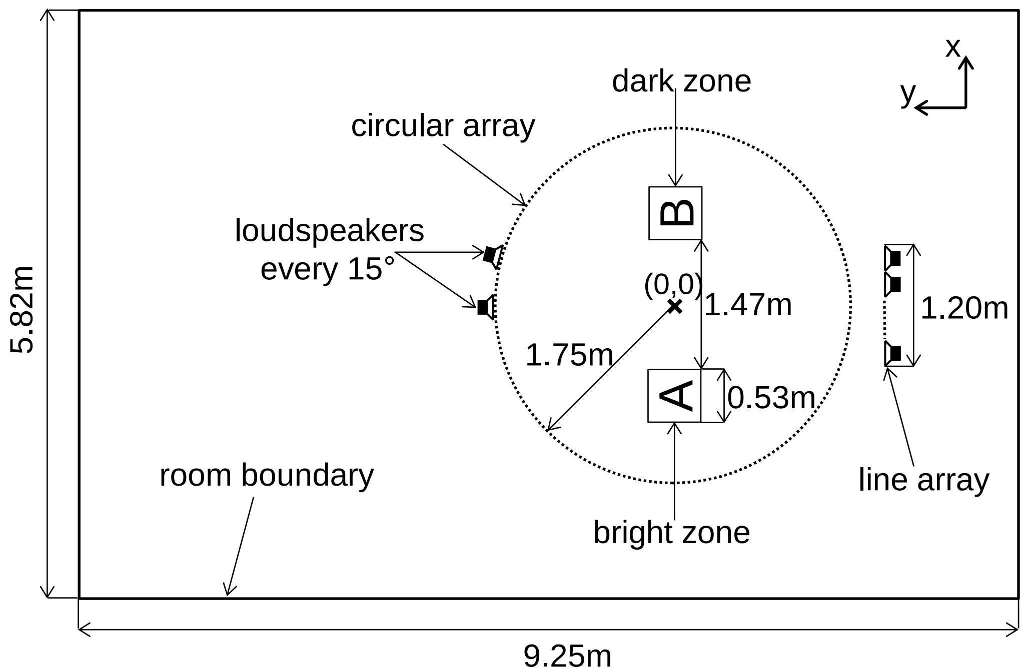
$V \approx 320\ m^3$



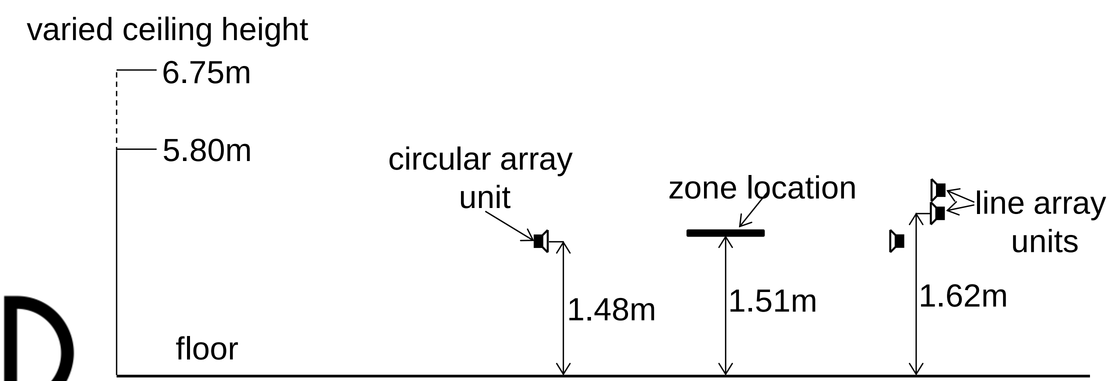
Experimental Setup

Details

$V \approx 320\ m^3$

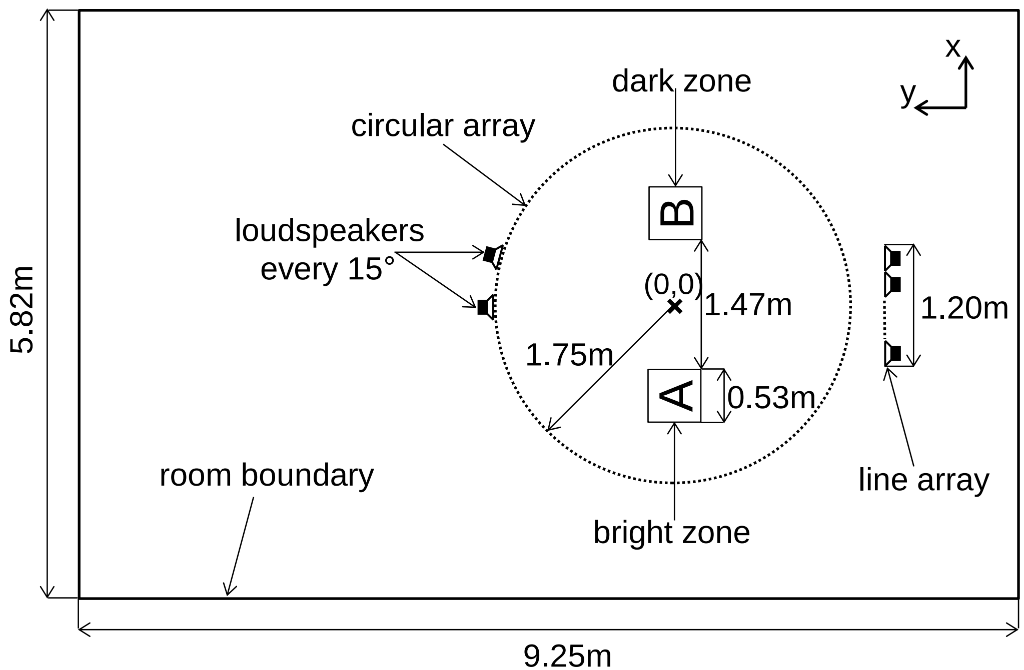


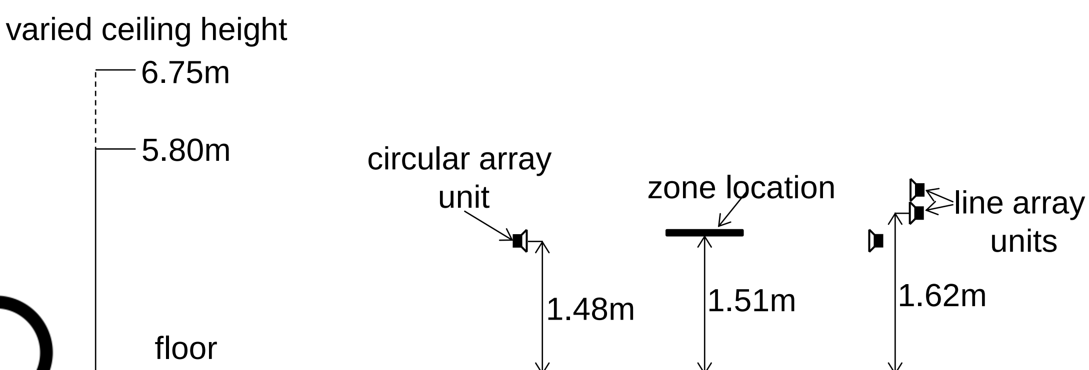
Centre Frequency (Hz)	RT60 (s)
250	0.33
500	0.30
1000	0.28
2000	0.27
4000	0.25



Experimental Setup
Details

$V \approx 320\ m^3$



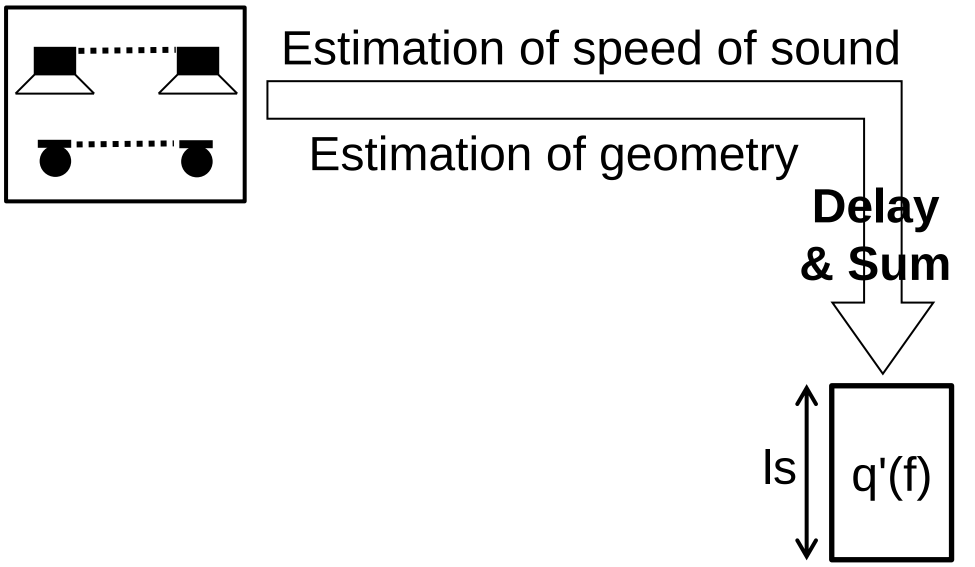


Centre Frequency (Hz)	RT60 (s)
250	0.33
500	0.30
1000	0.28
2000	0.27
4000	0.25

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

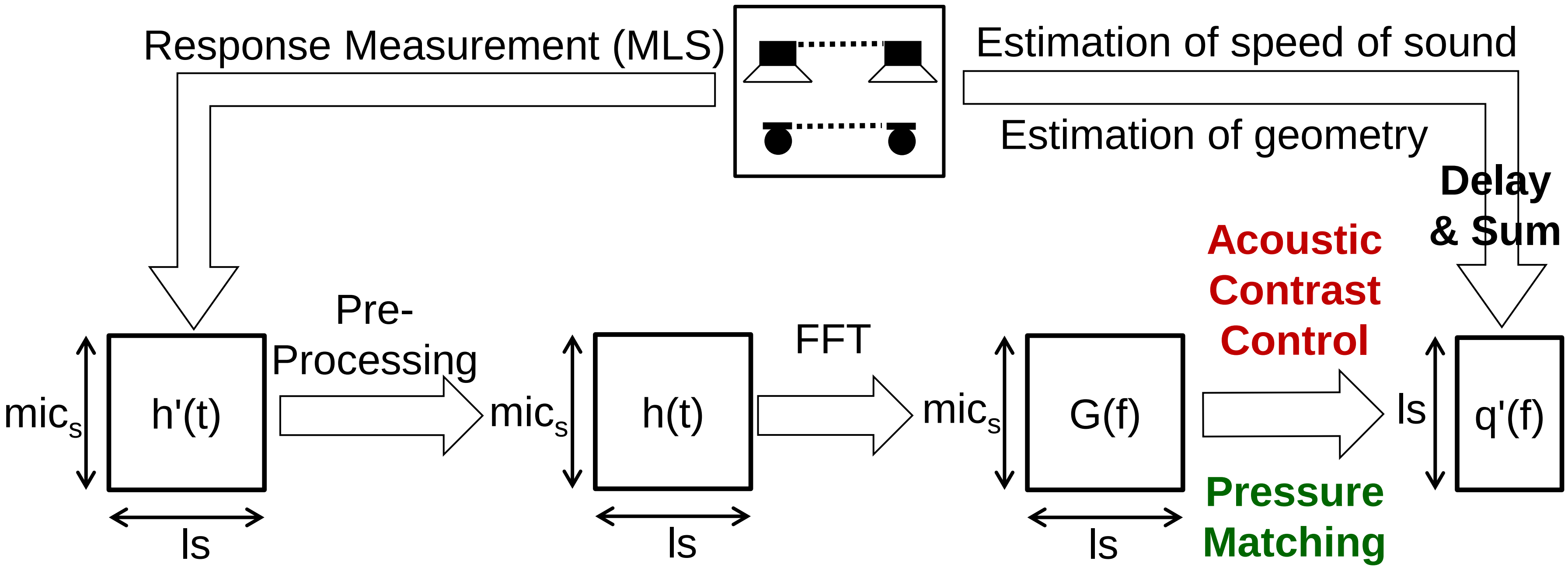
System Diagram

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

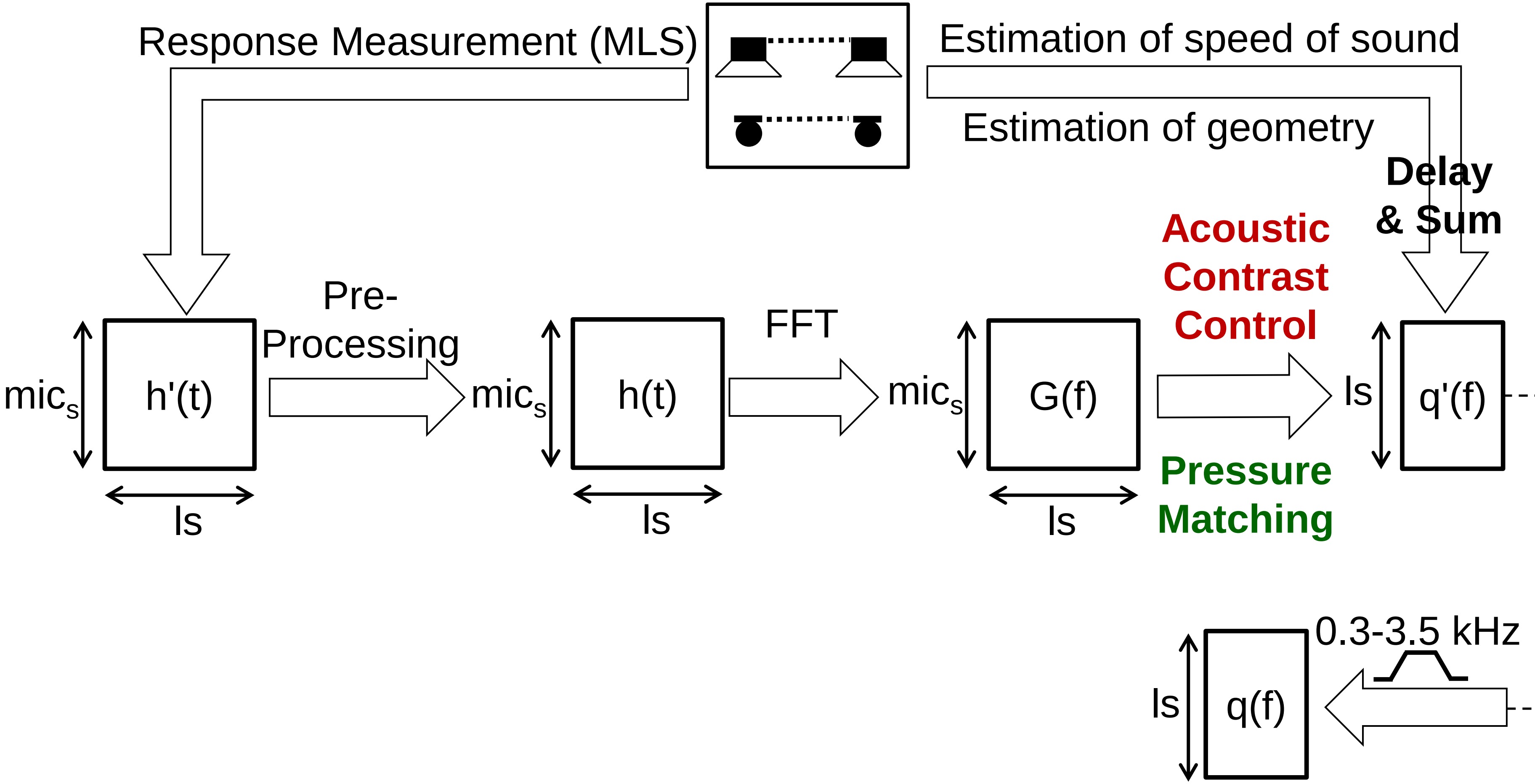


System Diagram

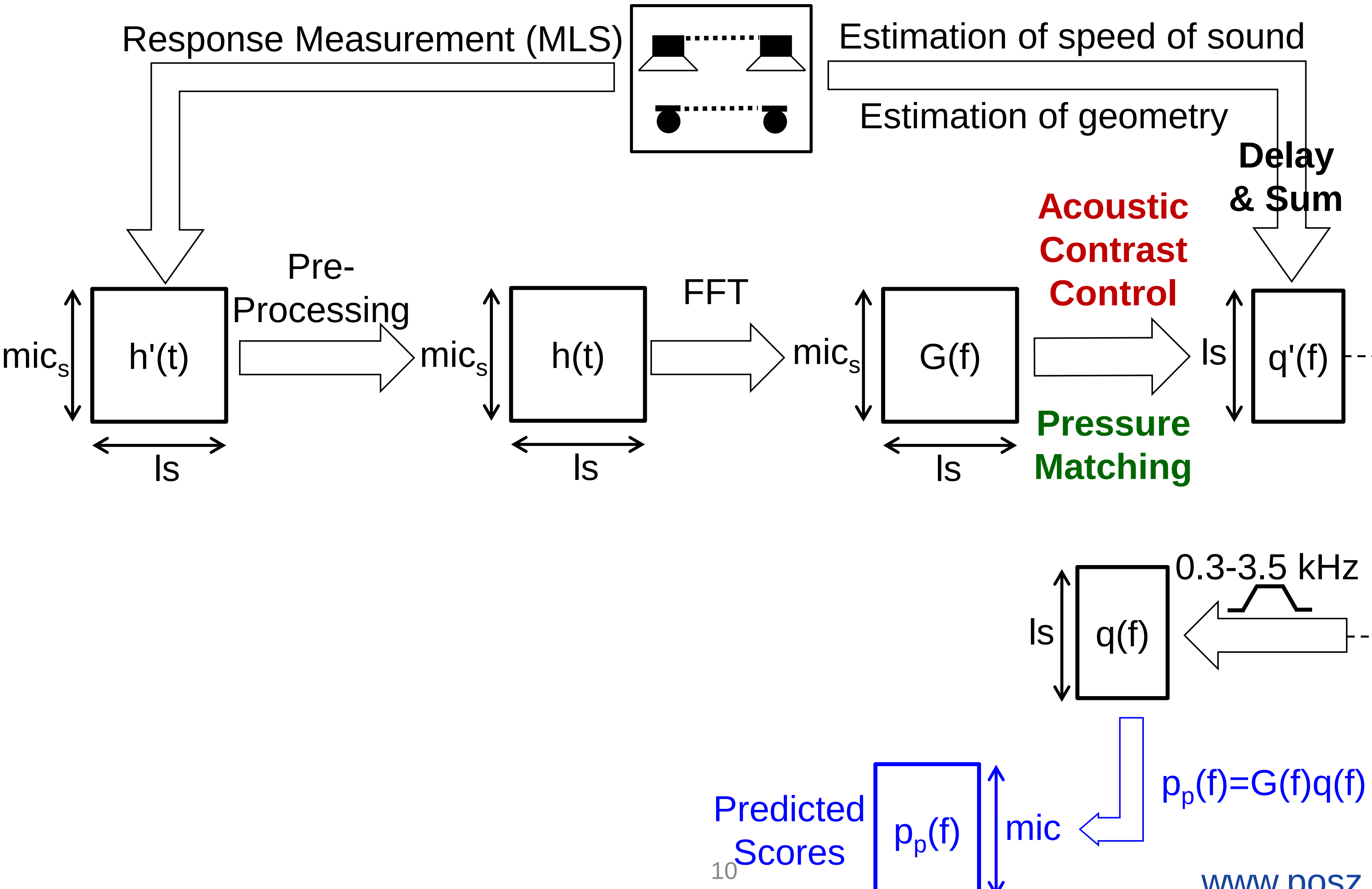
UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing



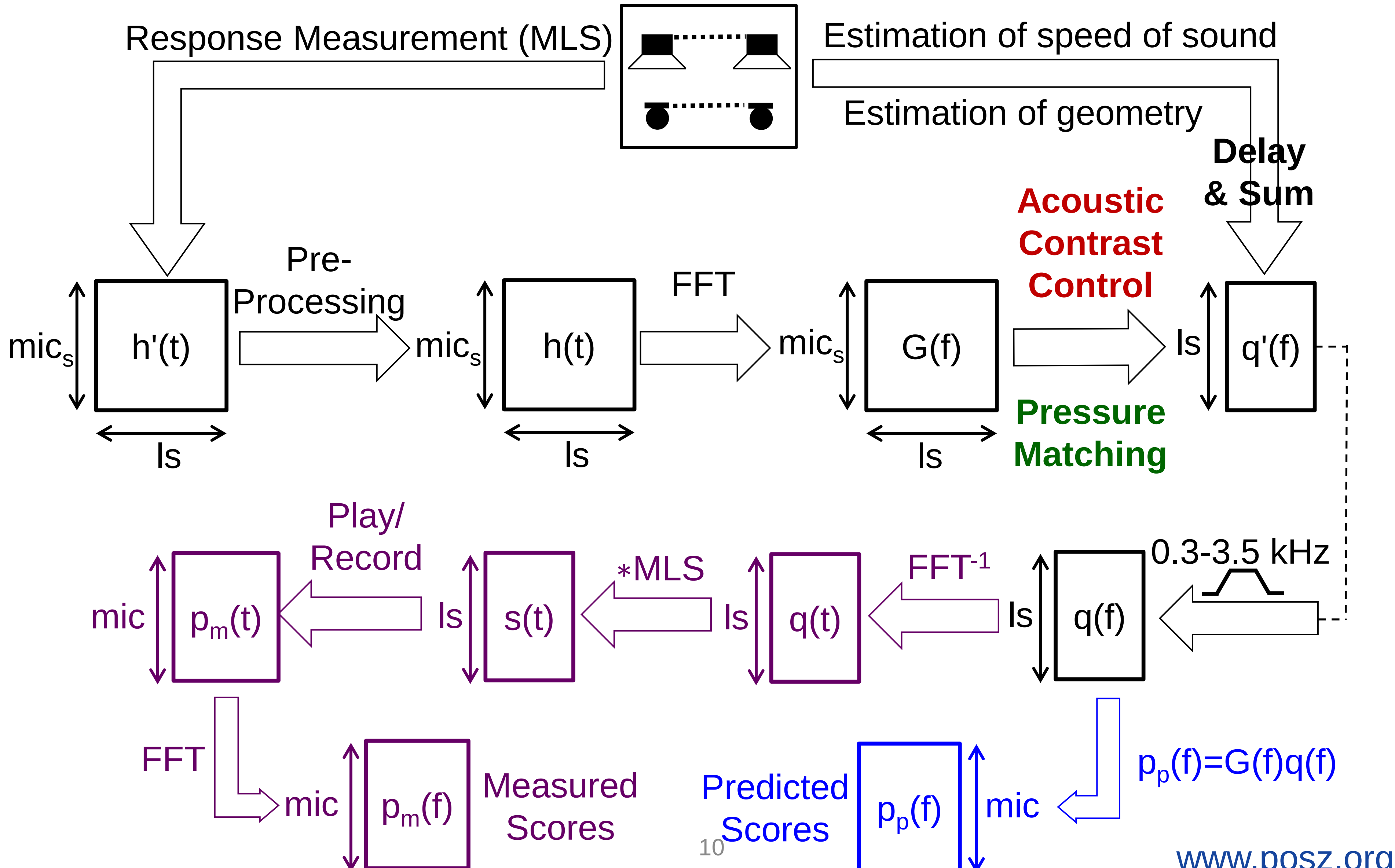
System Diagram



System Diagram

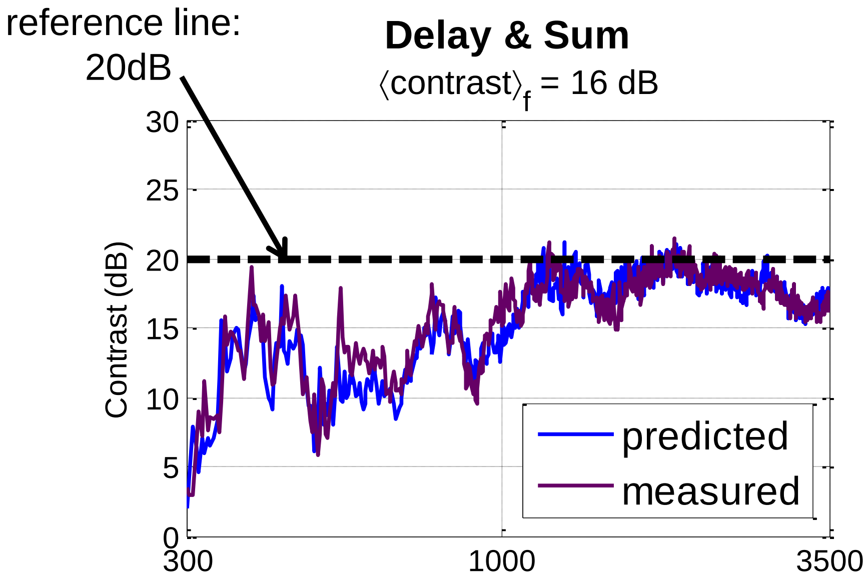


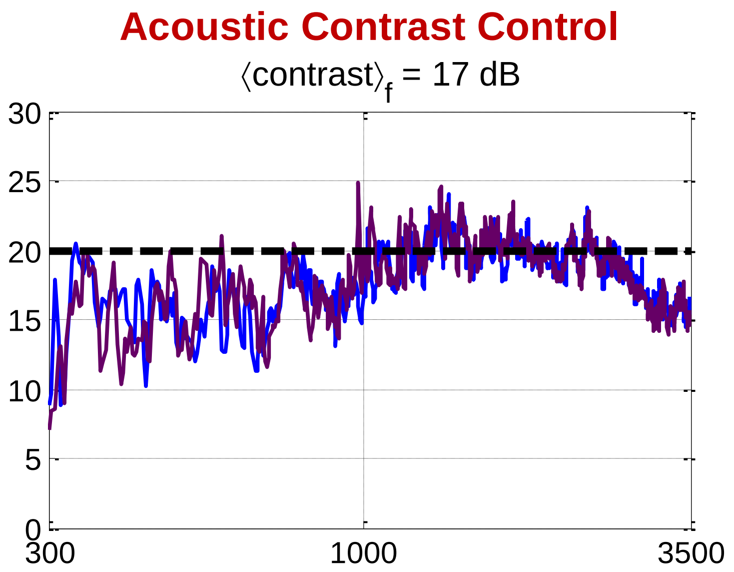
System Diagram



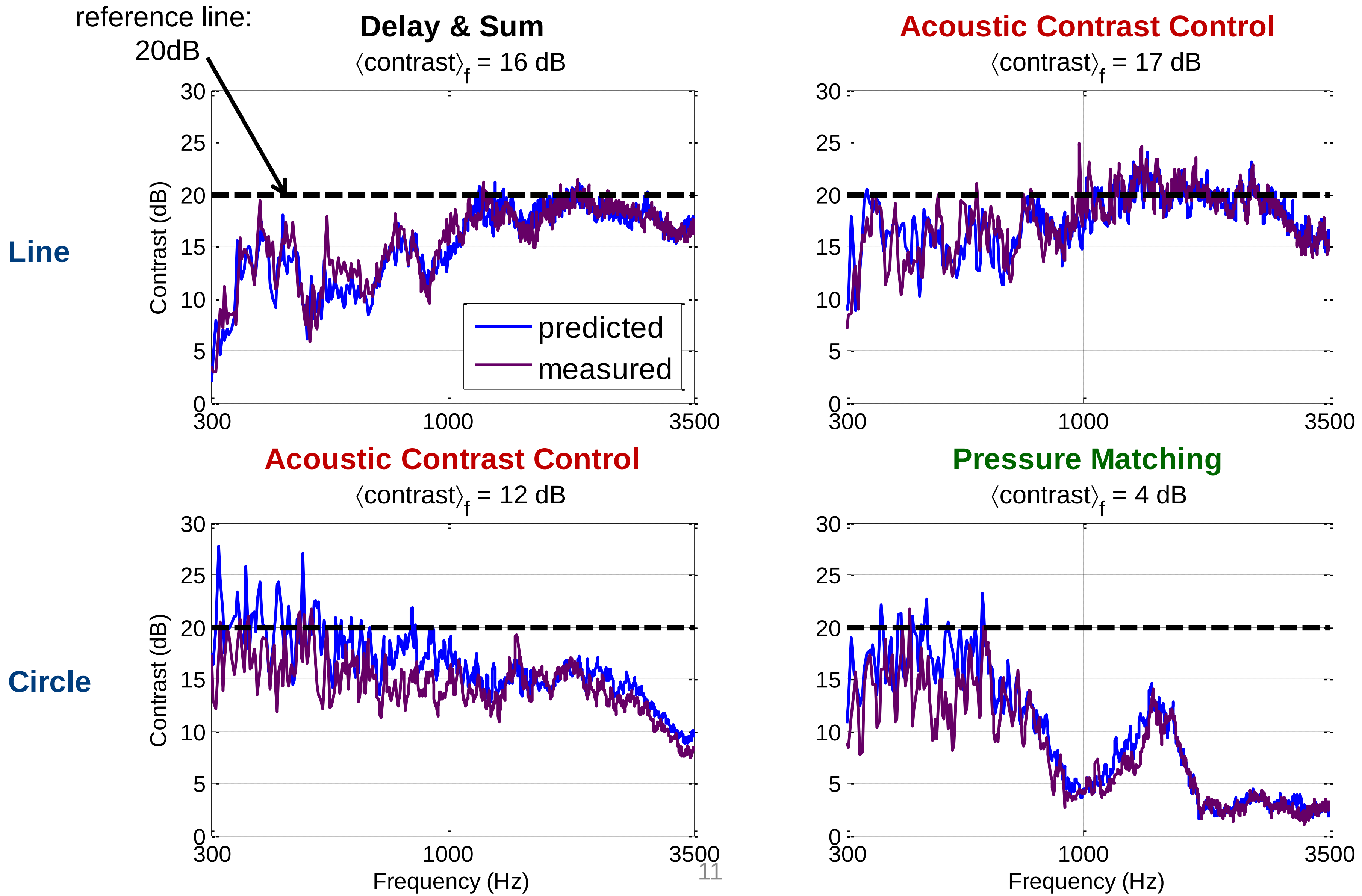
Physical Evaluation: Contrast

Line





Physical Evaluation: Contrast



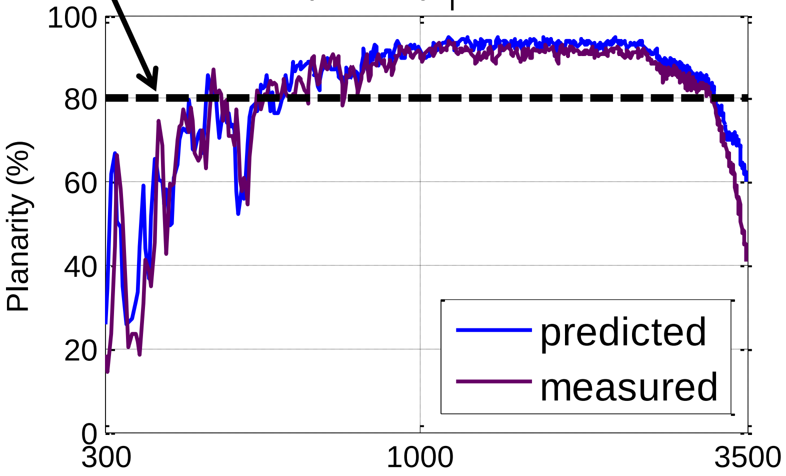
Physical Evaluation: Planarity

Line

reference line: 80%

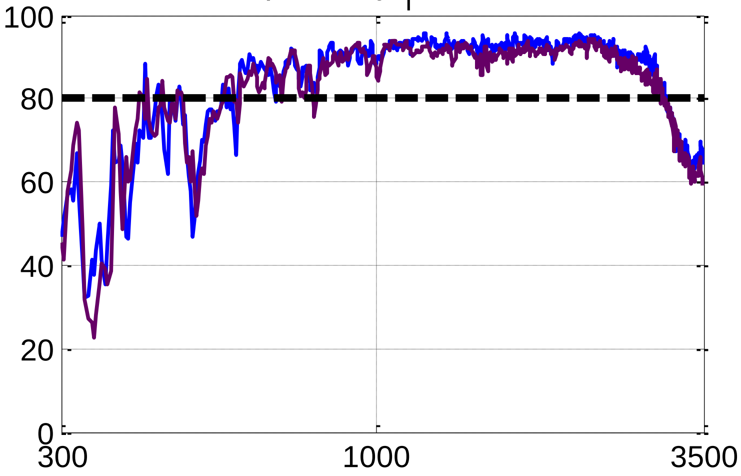
Delay & Sum

⟨planarity⟩$_f$ = 83 %

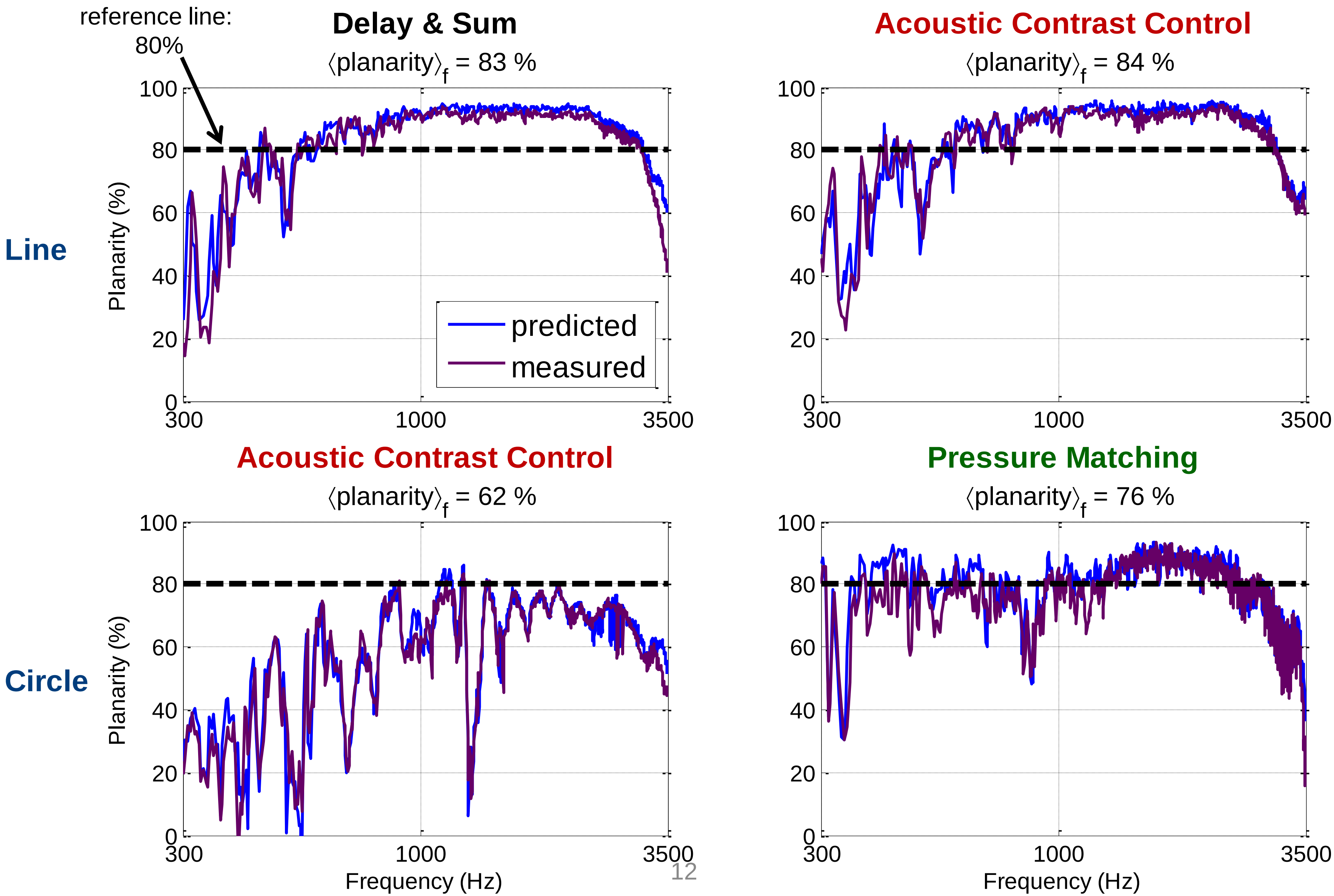


Acoustic Contrast Control

⟨planarity⟩$_f$ = 84 %

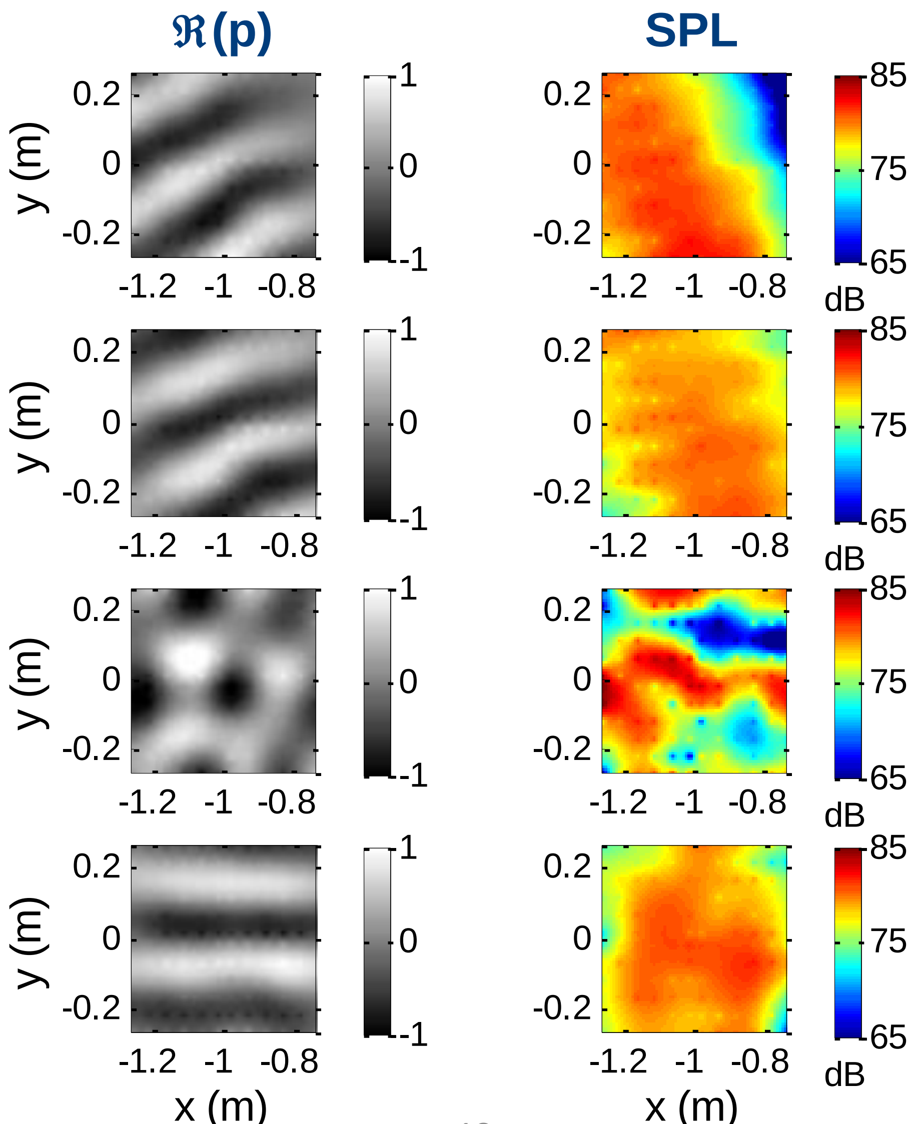


Physical Evaluation: Planarity



Zone A Pressure Maps: 1. 5 kHz
Measured Data

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

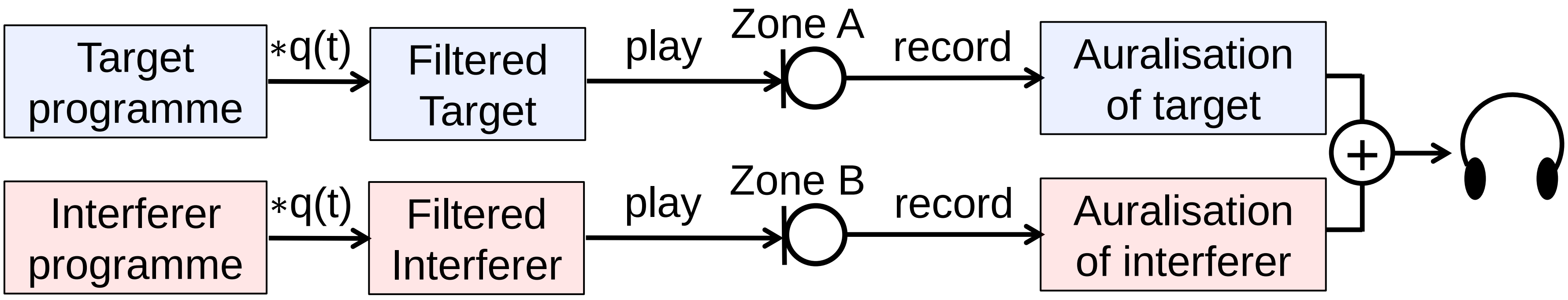


	Planarity (1.5 kHz)
DS (Line)	90%
ACC (Line)	85%
ACC (Circle)	66%
PM (Circle)	89%

POSZ

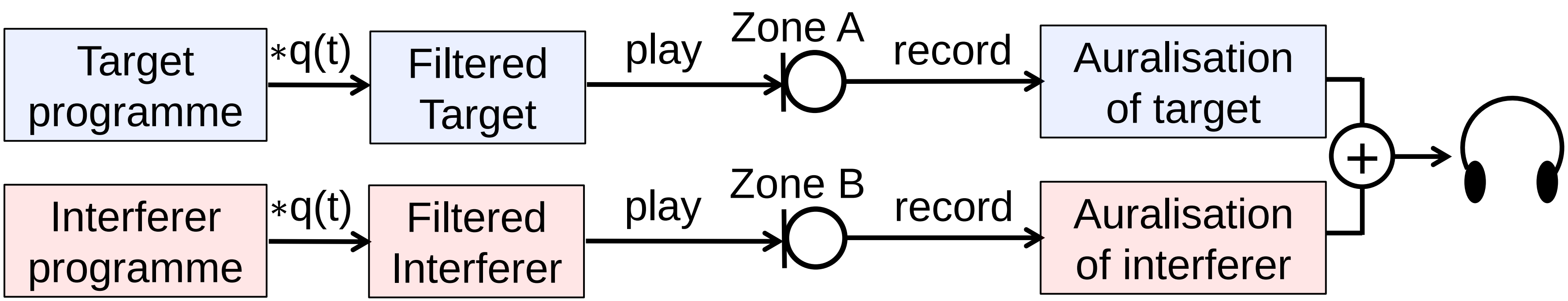
Perceptual Evaluation
Stimulus Creation

- **Listening tests based on recordings in the zones**



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

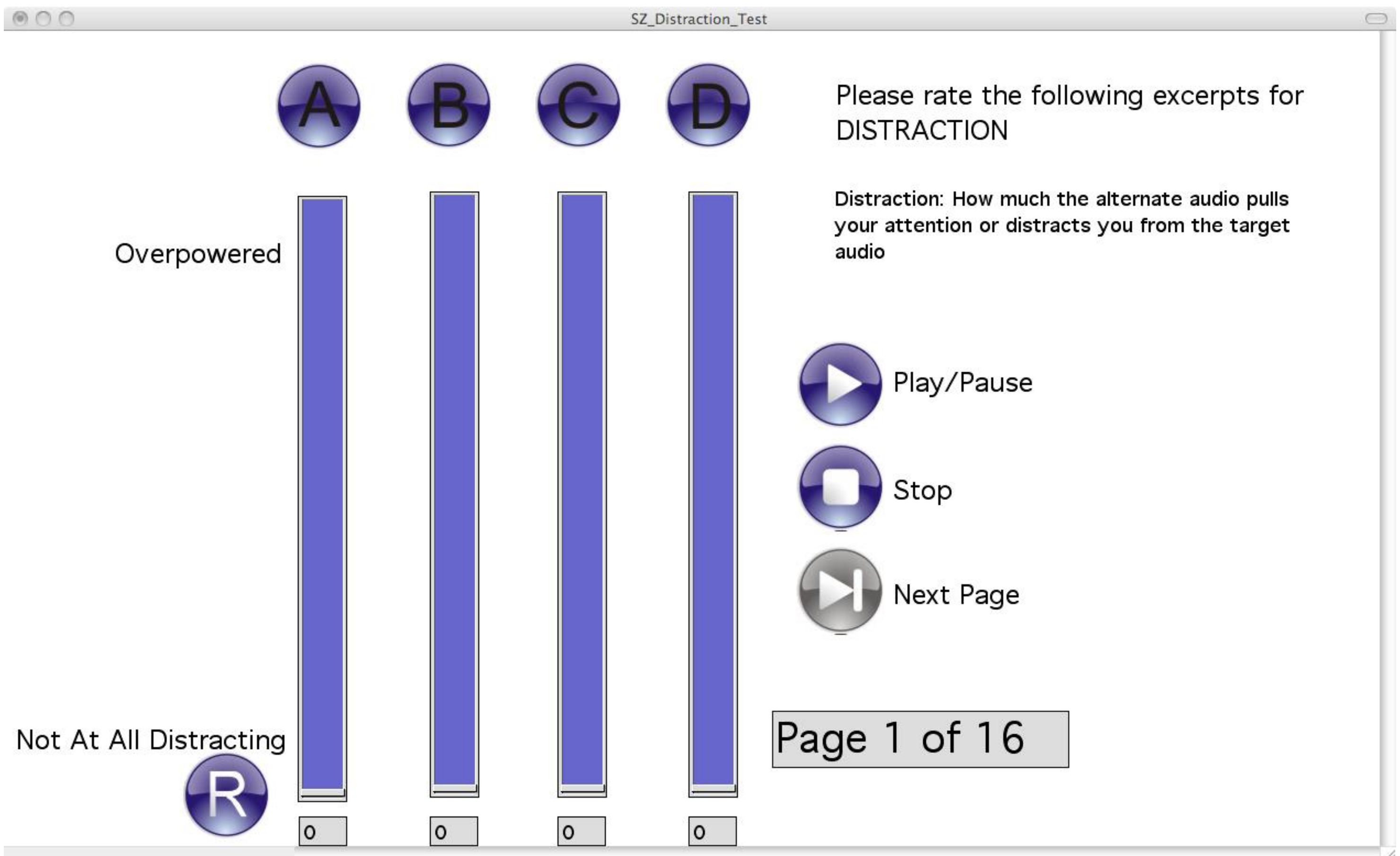
UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

- **24 stimuli (combination of methods and programme items)**
- **Rate 'distraction'** [Francombe *et al.* 2013]

Perceptual Evaluation
Experiment Design

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

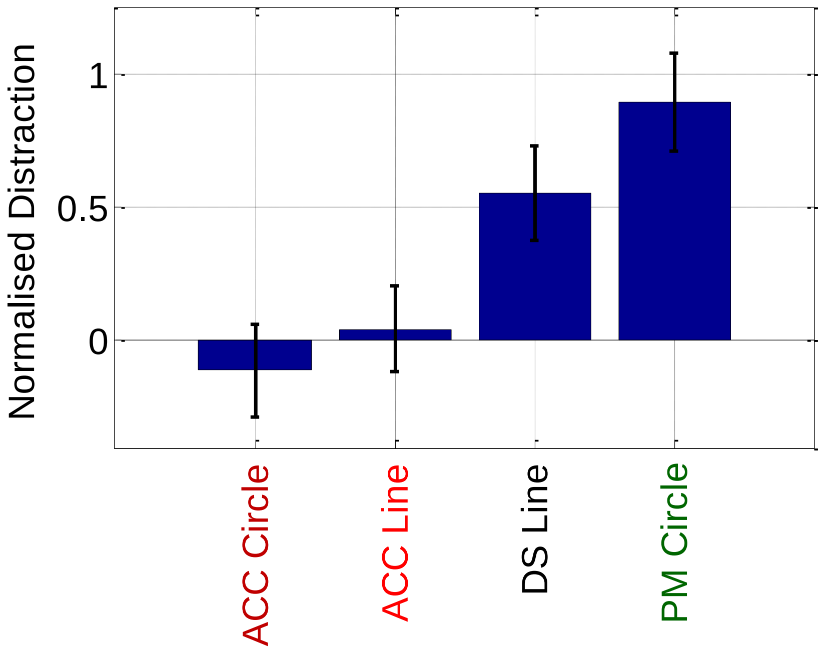
- **24 stimuli (combination of methods and programme items)**
- **Rate 'distraction'** [Francombe *et al.* 2013]



Perceptual Evaluation

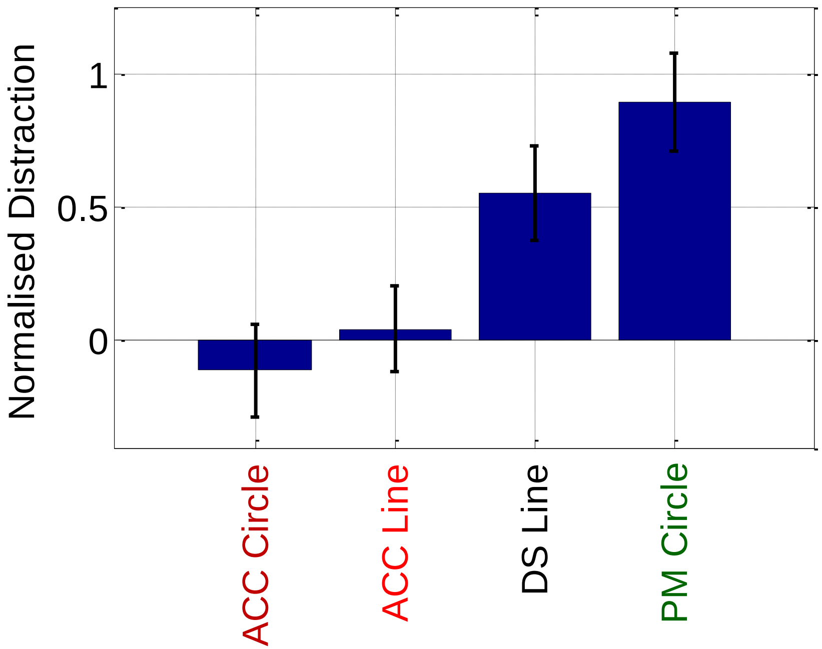
Results

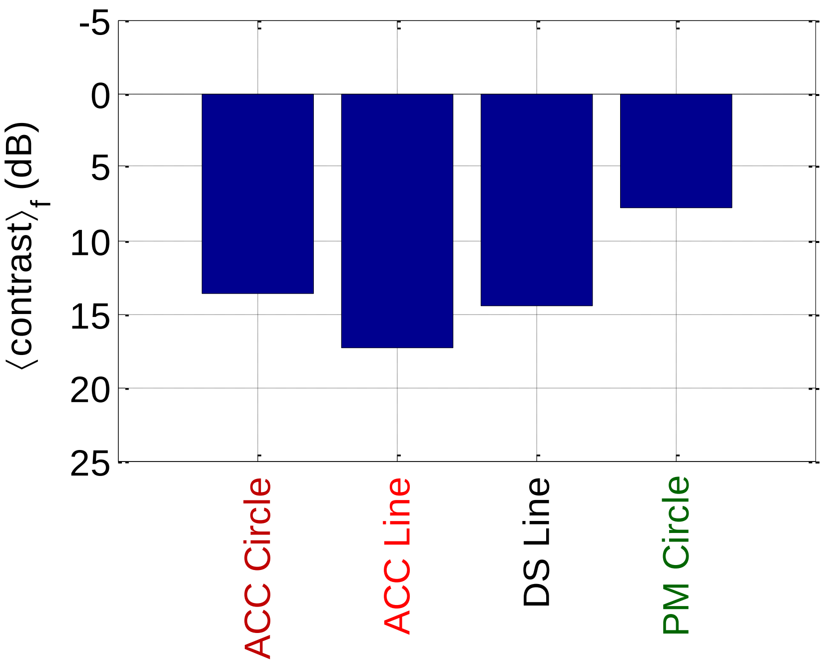
- **ANOVA model – most significant factor: sound zone method**



Perceptual Evaluation
Results

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





- **Other significant factors influenced distraction: programme material and interactions**

Summary and Further Work

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

- **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

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

- **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**

POSZ

Summary and Further Work

University of Surrey
Centre for Vision, Speech and Signal Processing

- **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

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

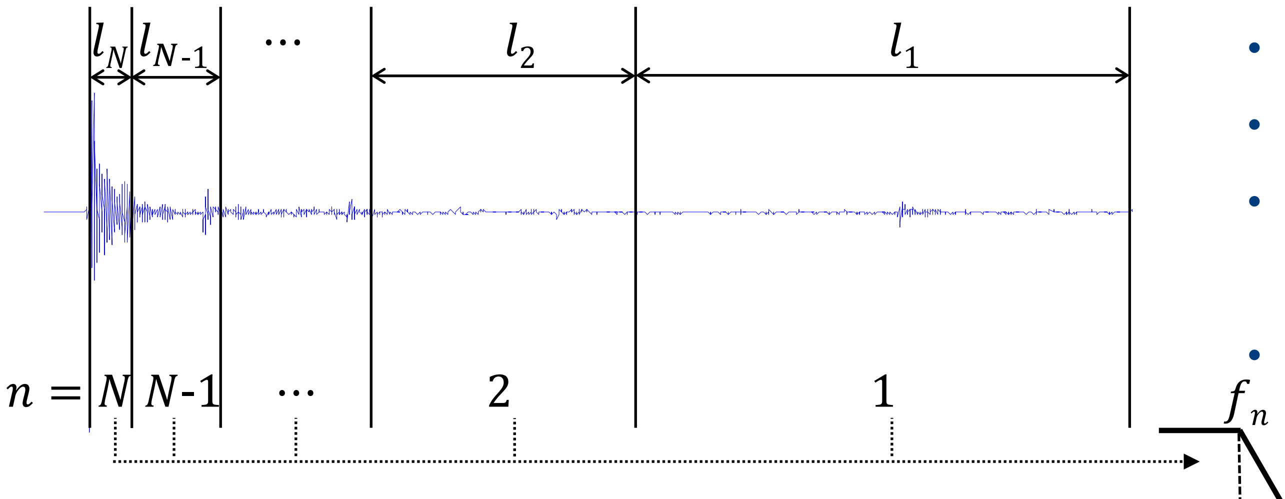
- M.A. Akeroyd, J. Chambers, D. Bullock, A.R. Palmer, A.Q. Summerfield, P.A. Nelson and S. Gatehouse, "The binaural performance of a cross-talk cancellation system with matched or mismatched setup and playback", J. Acoust. Soc. Am. 121, 1056-1069 (2007).
- J. Choi and Y. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources", J. Acoust. Soc. Am. 111, 1695–1700 (2002).
- P. Coleman, P.J.B. Jackson, M. Olik, M. Olsen, M. Møller and J. Pedersen, "The influence of regularization on anechoic performance and robustness of sound zone methods", 21st International Congress of Acoustics (2013).
- S. Elliott and M. Jones, "An active headrest for personal audio", J. Acoust. Soc. Am. 124, 2702-2709 (2006).
- J. Francombe, R. Mason, M. Dewhirst and Søren Bech, "Modelling listener distraction resulting from audio-on-audio interference", 21st International Congress of Acoustics (2013).
- P.J.B. Jackson, F. Jacobsen, P. Coleman and J. Pedersen, "Sound field planarity characterized by superdirective beamforming", 21st International Congress of Acoustics (2013).
- F. Jacobsen, M. Olsen, M. Møller and F. Agerkvist, "A comparison of two strategies for generating sound zones in a room", 18th International Congress of Sound and Vibration (2011).
- M. Møller, M. Olsen and F. Jacobsen, "A hybrid method combining synthesis of a sound field and control of acoustic contrast", AES 132nd Convention (2012).
- M. Poletti, "An investigation of 2-D multizone surround sound systems", AES 125th Convention (2008).
- M. Shin, S.Q. Lee, F.M. Fazi, P.A. Nelson, D. Kim, S. Wang, K.-H. Park and J. Seo, "Maximization of acoustic energy difference between two spaces", J. Acoust. Soc. Am. 128, 121-131 (2010).
- B.D. Van Veen and K.M. Buckley, "Beamforming: a versatile approach to spatial filtering", IEEE ASSP Magazine 5, 4-24 (1988).
- Y.J. Wu and T.D. Abhayapala, "Simultaneous soundfield reproduction at multipole spatial regions", AES 128th Convention (2010).

POSZ

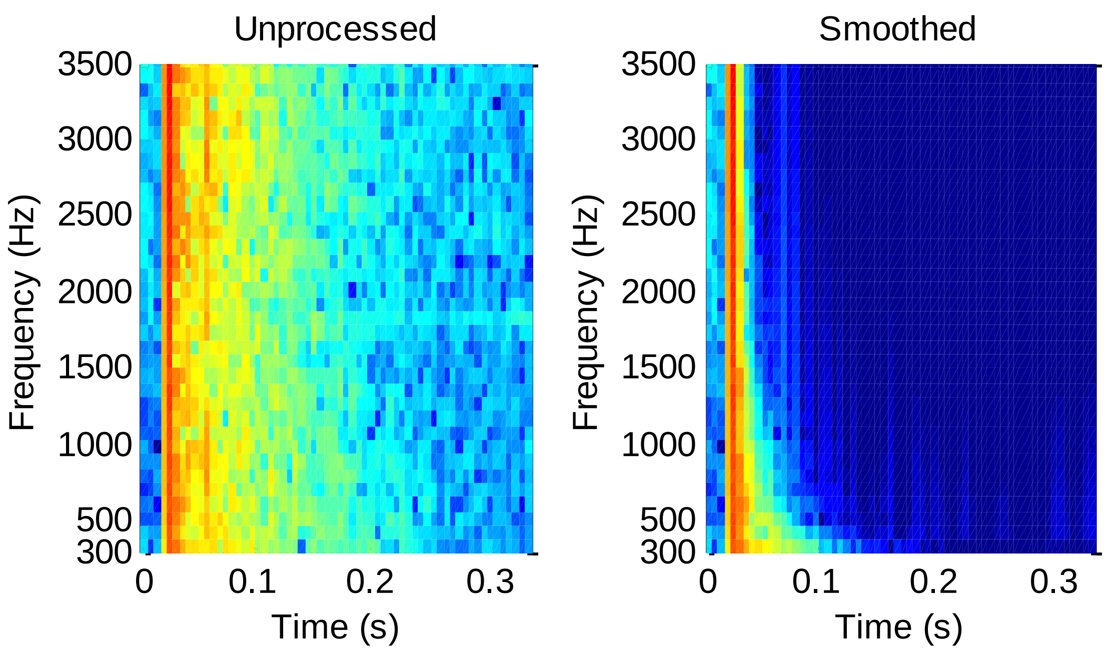
Appendix A

Pre-Processing: Impulse Smoothing

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing



- $l_n = f_s i / f_n$
- i - number of bins before f_n
- f_n- cut-off freq. of LPF in n-th band
- f_s - sampling frequency

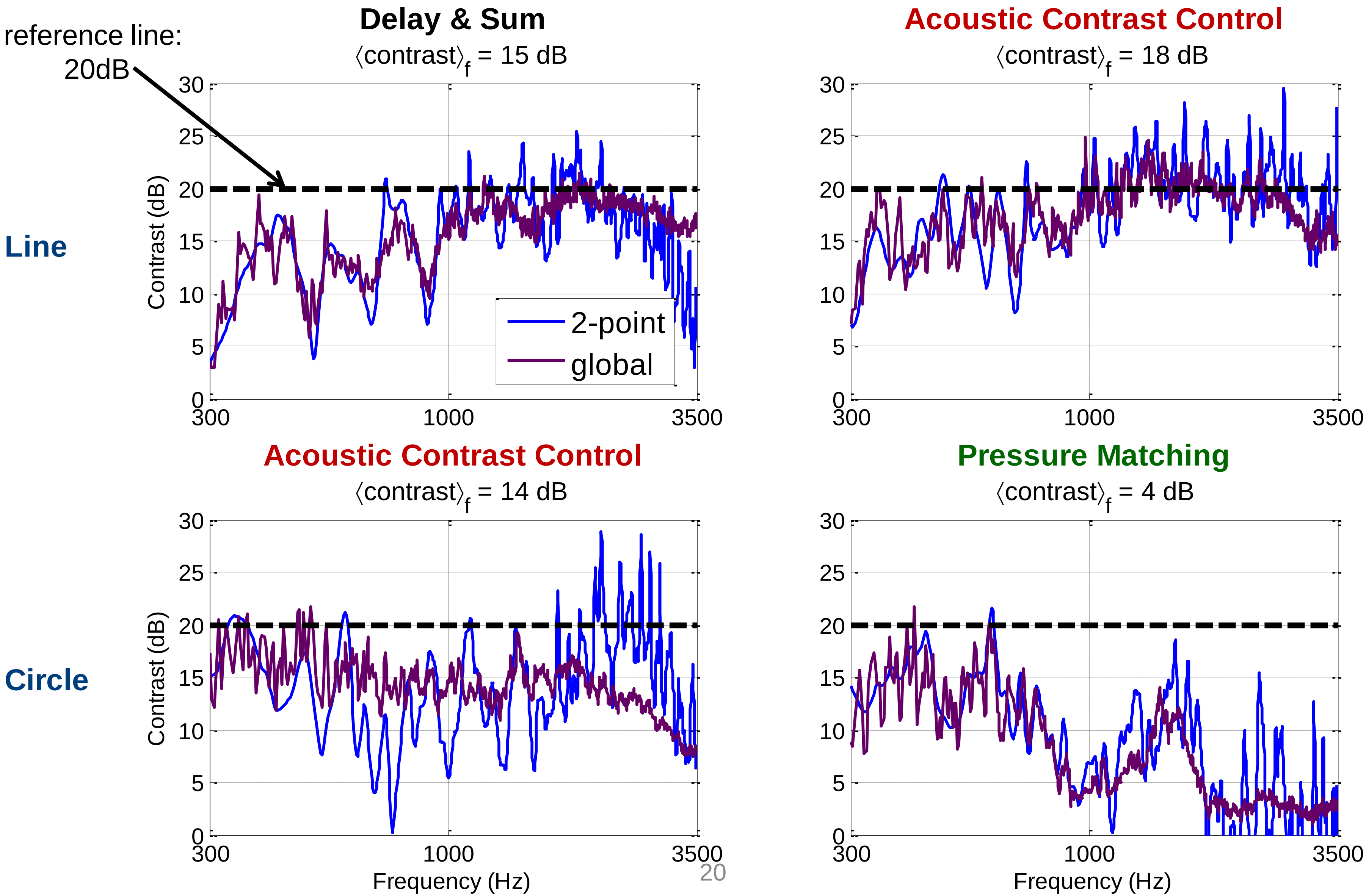


Parameter	Value
i	10
f_s	48 kHz
N	141
f_1	8 Hz
f_N	2.376 kHz
Freq. spacing	1/12 Oct.
Slope	80 dB/Dec.

Appendix B

Contrast: Two Recording Points

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing



Appendix C

Perceptual Evaluation: Details of Experiment Design

UNIVERSITY OF SURREY
Centre for Vision, Speech and Signal Processing

- **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

POSZ