University of Huddersfield

3D Microphone Array Comparison: Objective Measurements

Hyunkook Lee and Dale Johnson

h.lee@hud.ac.uk

Applied Psychoacoustics Lab (APL)
University of Huddersfield, UK

JAES paper https://www.aes.org/e-lib/browse.cfm?elib=21536

H. Lee and D. Johnson, "3D Microphone Array Comparison: Objective Measurements"
J. Audio Eng. Soc., vol. 69, no. 11, pp. 871–887, (2021 November).
DOI: https://doi.org/10.17743/jaes.2021.0038

3D Microphone Array Comparison: Objective Measurements

HYUNKOOK LEE,* *AES Fellow,* **AND DALE JOHNSON**
(h.lee@hud.ac.uk) (d.s.johnson2@hud.ac.uk)

Applied Psychoacoustics Laboratory (APL), University of Huddersfield, Huddersfield, United Kingdom

This paper describes a set of objective measurements carried out to compare various types of 3D microphone arrays, comprising OCT-3D, PCMA-3D, 2L-Cube, Decca Cuboid, Eigenmike EM32 (i.e., spherical microphone system), and Hamasaki Square with 0-m and 1-m vertical spacings of the height layer. Objective parameters that were measured comprised interchannel and spectral differences caused by interchannel crosstalk (ICXT), fluctuations of interaural level and time differences (ILD and ITD), interchannel correlation coefficient (ICC), interaural cross-correlation coefficient (IACC), and direct-to-reverberant energy ratio (DRR). These were chosen as potential predictors for perceived differences among the arrays. The measurements of the properties of ICXT and the time-varying ILD and ITD suggest that the arrays would produce substantial perceived differences in tonal quality as well as locatedness. The analyses of ICCs and IACCs indicate that perceived differences among the arrays in spatial impression would be larger horizontally rather than vertically. It is also predicted that the addition of the height channel signals to the base channel ones in reproduction would produce little effect on both source-image spread and listener envelopment, regardless of the array type. Finally, differences between the ear-input signals in DRR were substantially smaller than those observed among microphone signals.

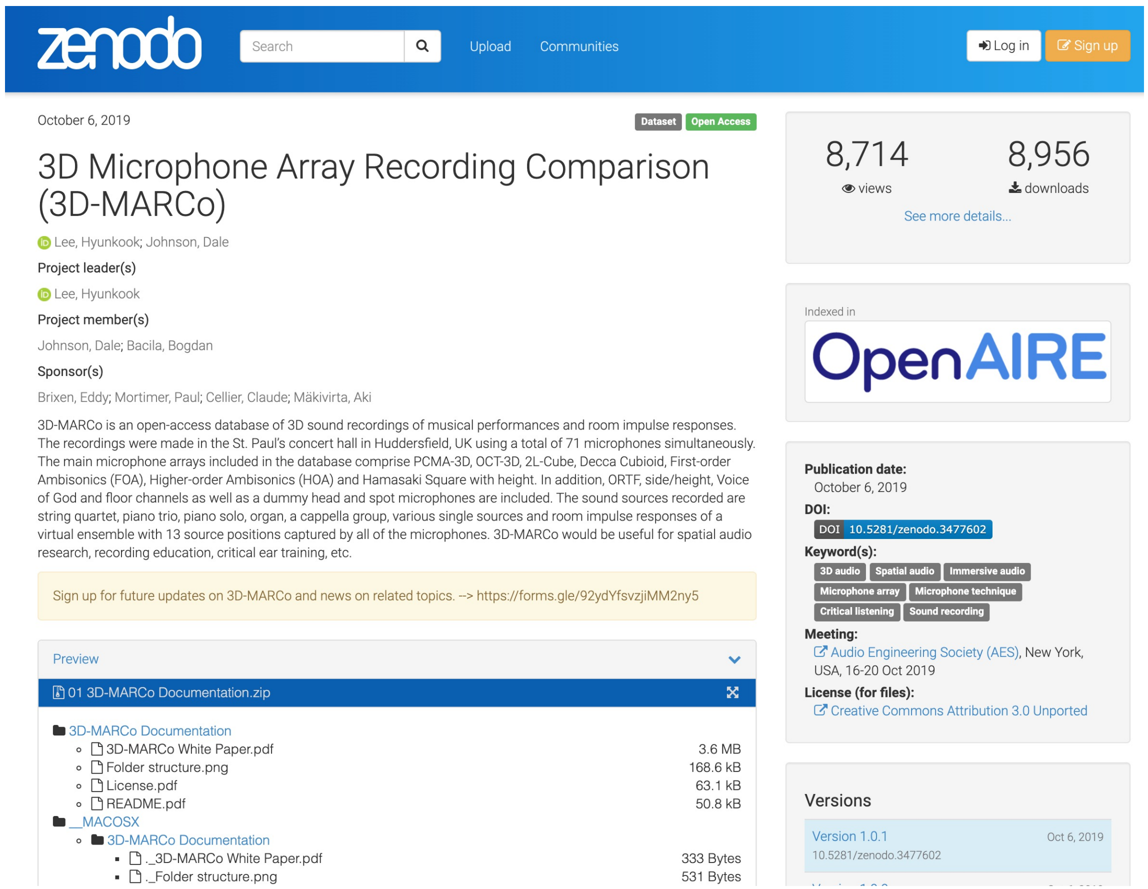
Motivation

- Various different microphone techniques for 3D sound capture have been proposed over the years.
- However, no scientifically rigorous study has been conducted to compare the perceived qualities of the techniques yet.
- More importantly, perceptual differences of different techniques have not been formally elicited yet

  → Attribute scales for evaluating 3D acoustic recordings need to be established.
- As a first step, a set of objective measurements were carried out to gain insights into physical differences among different 3D mic arrays and to form hypotheses for later subjective tests.

3D-MARCo database https://doi.org/10.5281/zenodo.3474285

- The 3D-MARCo database provides a large number of 3D microphone array recordings and impulse responses for research & education.



3D-MARCo database

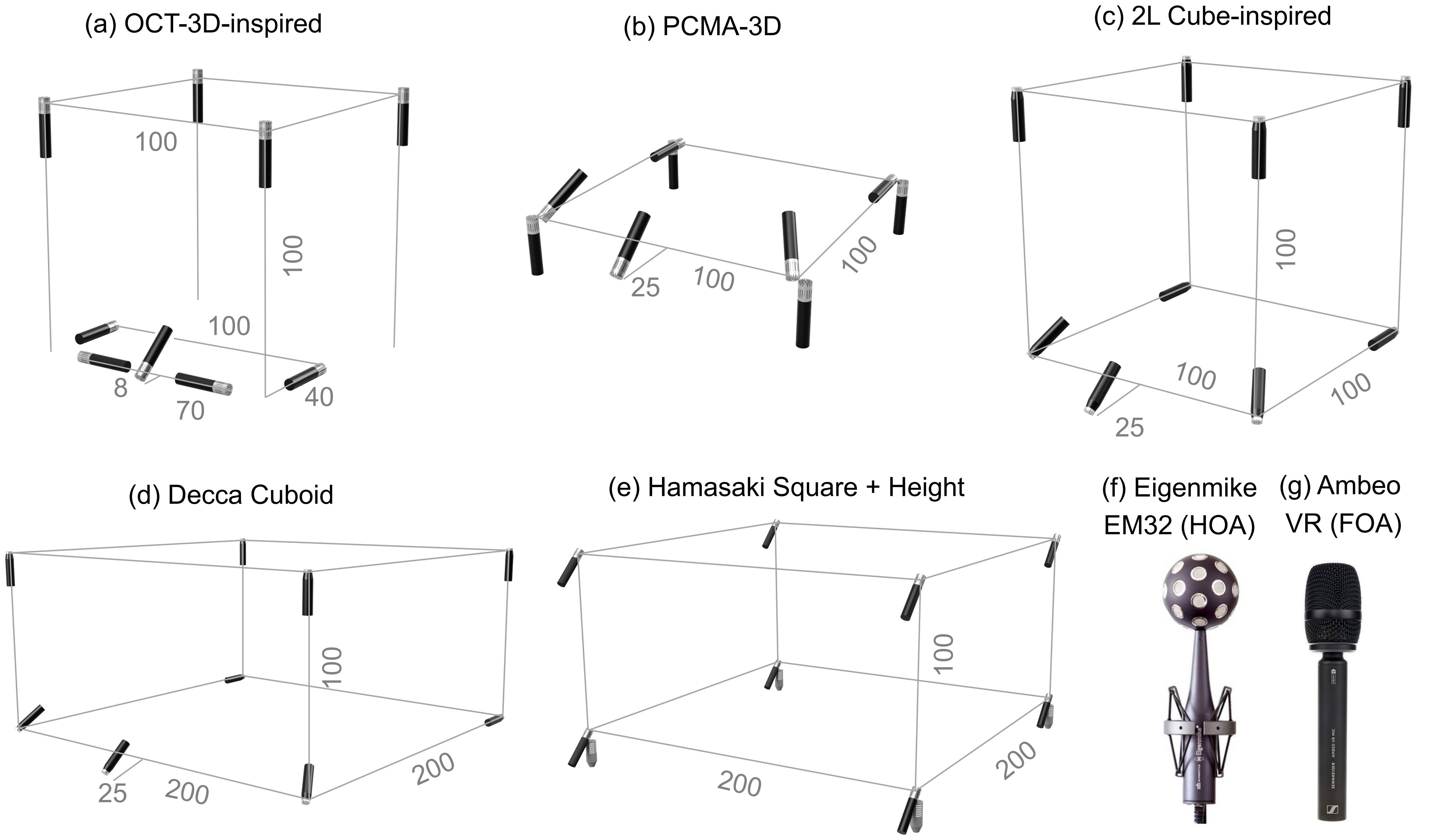
https://doi.org/10.5281/zenodo.3474285

- Record various types of musical performances using multiple microphones of similar tonal characteristics simultaneously.
- DPA d:dicate series were exclusively used for all main arrays.
- A total of 71 microphones were used.

Mic technique	
PCMA-3D	KU100 dummy head
OCT-3D	Side/height mics
2L Cube-inspired	”Voice of God”
Decca Cuboid	Floor mics
Hamasaki Square with height (at 0m and 1m)	Spot mics
Eigenmike EM32 (HOA)	
Ambeo (FOA)	

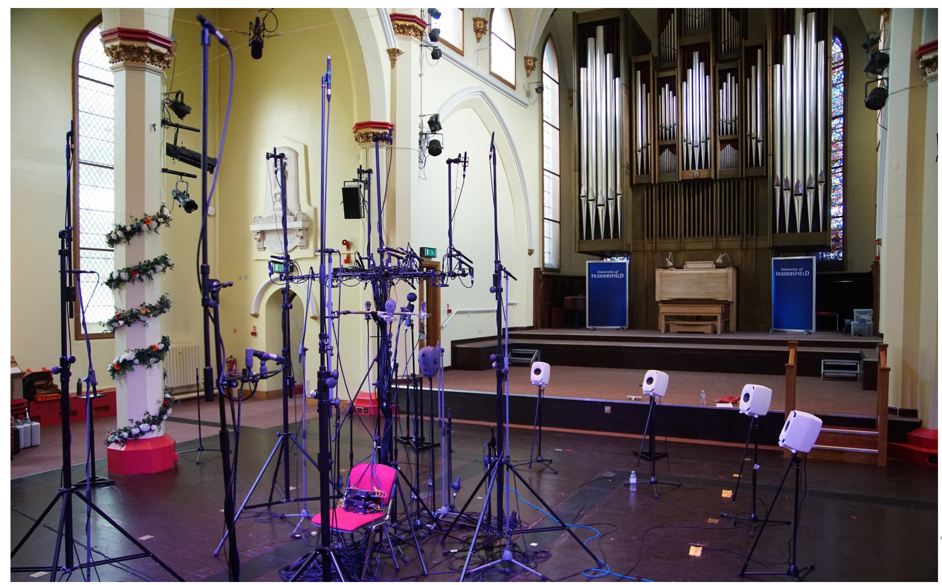
Sound sources
String quartet
Piano trio
Organ
Piano solo
A Cappella
Single sources (Speech, cello, conga, trumpet at 0°, 15°, 30°, 45°, 60°, 75°, 90°)

3D Microphone Arrays Included in the session



3D-MARCo database

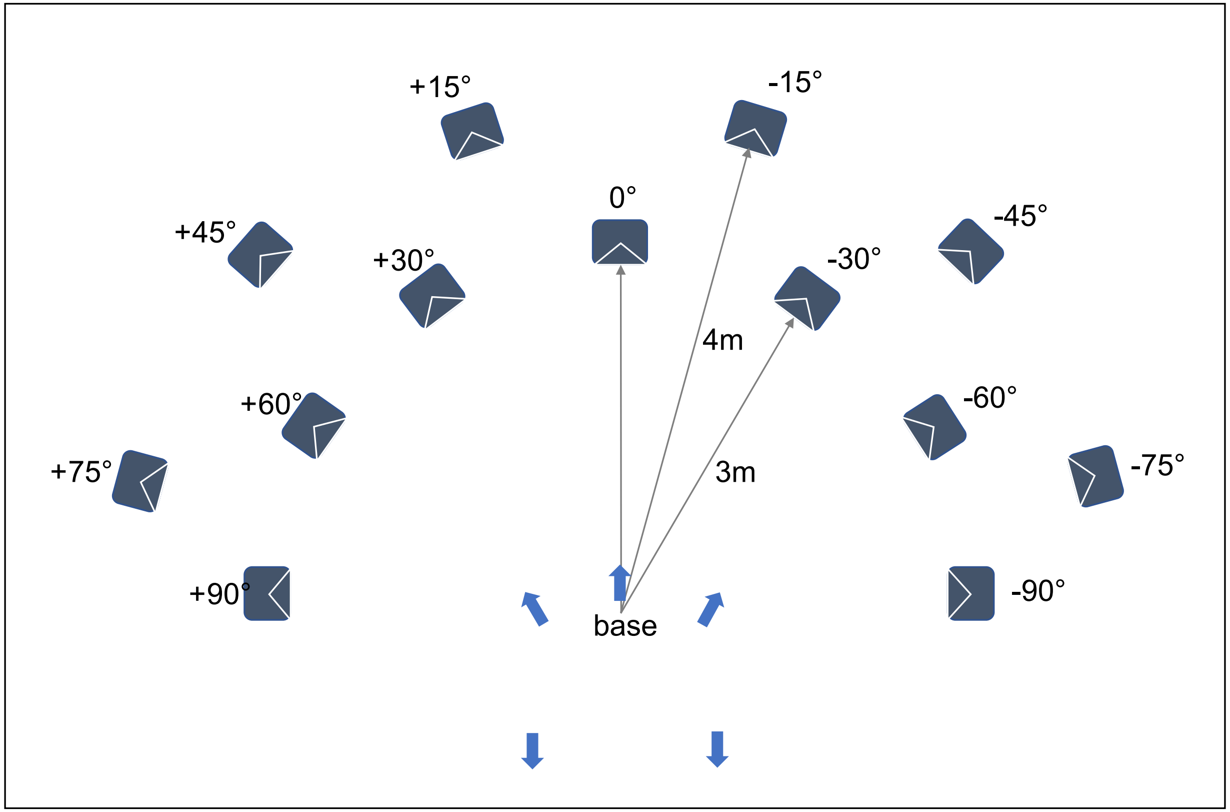
https://doi.org/10.5281/zenodo.3474285



3D-MARCo database

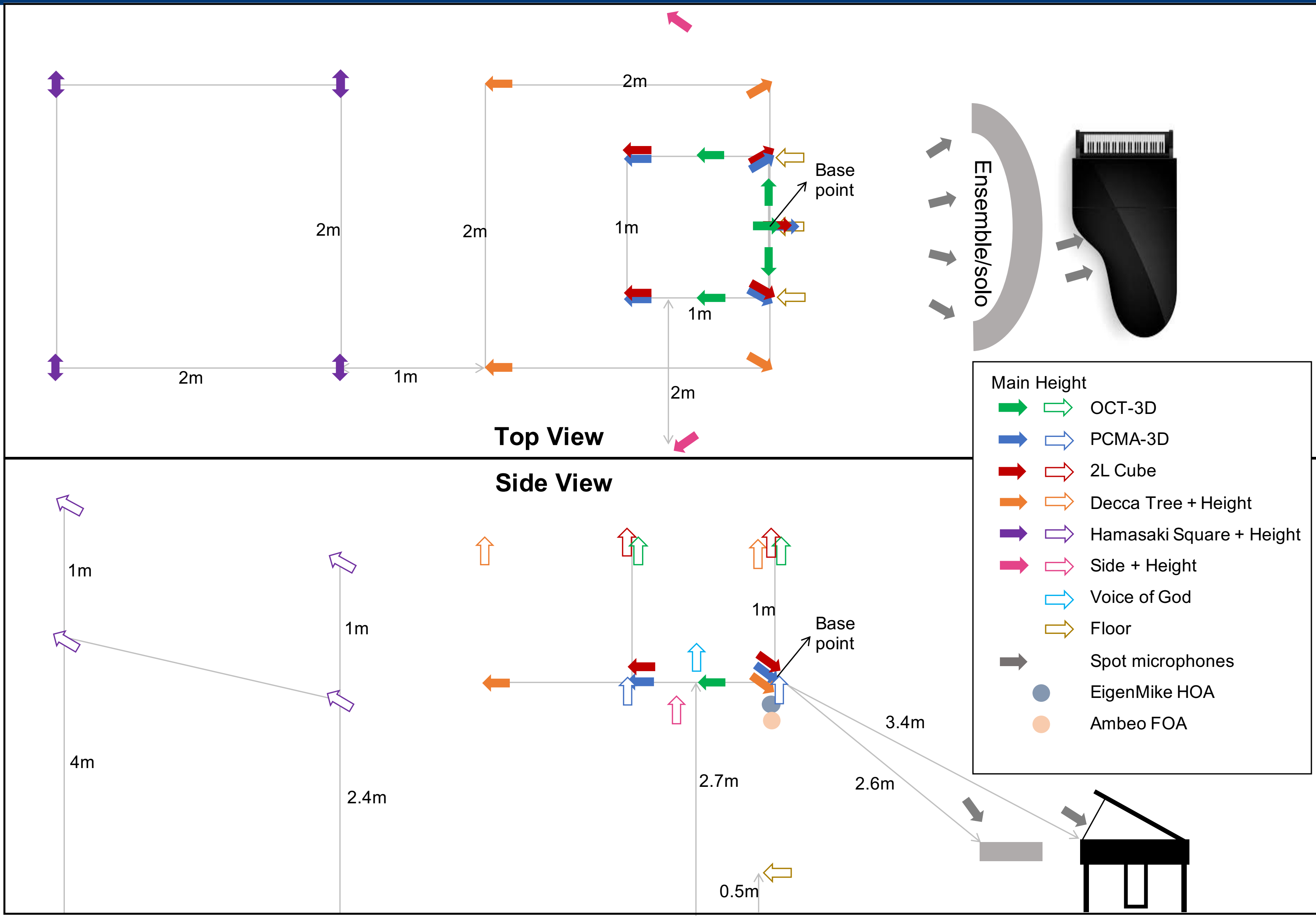
https://doi.org/10.5281/zenodo.3474285

- Room impulse responses at 13 different positions captured by all of the microphones used.
- Loudspeakers used: Genelec 8331A



3D-MARCo database

https://doi.org/10.5281/zenodo.3474285

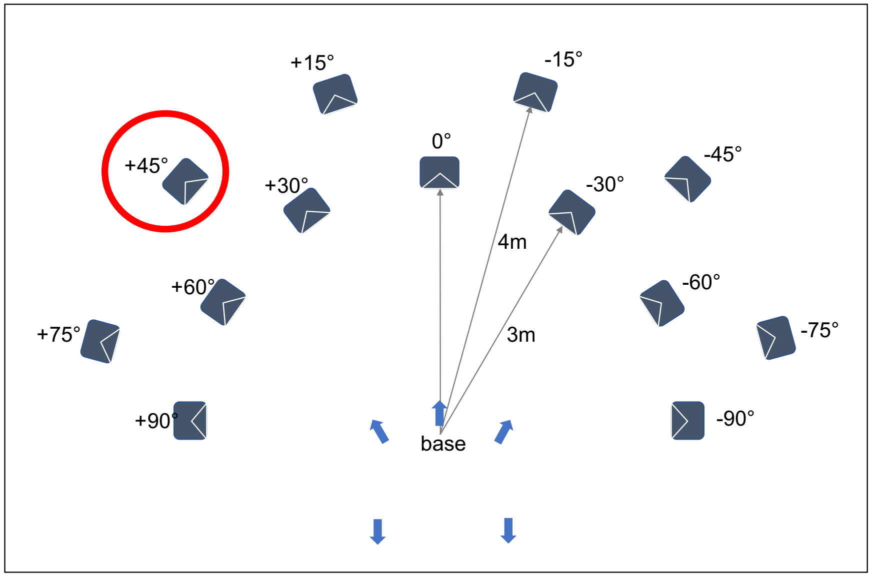


ee 2021

Objective Measurements

Microphone arrays analysed

	Perceptually Motivated		Physically Motivated
Main Array	Horizontally and Verically Spaced (HVS) OCT-3D 2L-Cube Decca Cuboid	Horizontally spaced / Vertically coincident (HSVC) PCMA-3D	Horizontally and Vertically Coincident (HVC) Eigenmike EM-32
Ambience Array	Hamasaki Square (HS) with height layer at 1 m above	Hamasaki Square (HS) with height layer at 0 m	



Using room impulse responses captured for the source at +45°

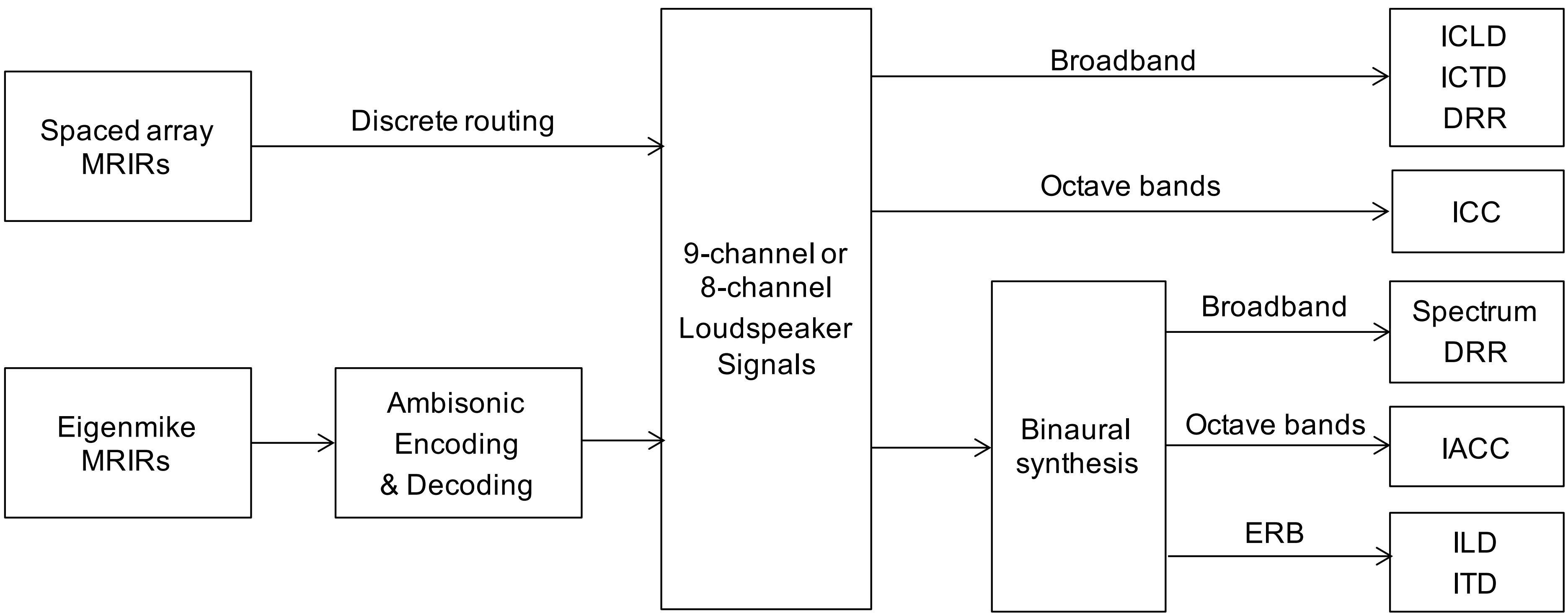
Objective Measurements

Parameters

- The Interchannel level difference (ICLD) and interchannel time difference (ICTD) of interchannel crosstalk (ICXT).
- Temporal fluctuations of interaural level and time differences (ILD and ITD).
- Ear-signal's spectral distortion resulting from the ICXT of the height microphone layer.
- Interchannel correlation coefficient (ICC).
- Interaural cross-correlation coefficient (IACC).
- Direct-to-reverberant energy ratio (DRR).

Objective Measurements

Overall workflow of the objective measurements



Objective Measurements

Channel labels and loudspeaker positions

Channels	Labels	Azimuth (°)	Elevation (°)
Front Left	FL	+30	0
Front Right	FR	–30	0
Front Center	FC	0	0
Rear Left	RL	120	0
Rear Right	RR	–120	0
Front Left height	FLh	+45	+45
Front Right height	FLh	–45	+45
Rear Left height	RLh	+135	+45
Rear Right height	RRh	–135	+45

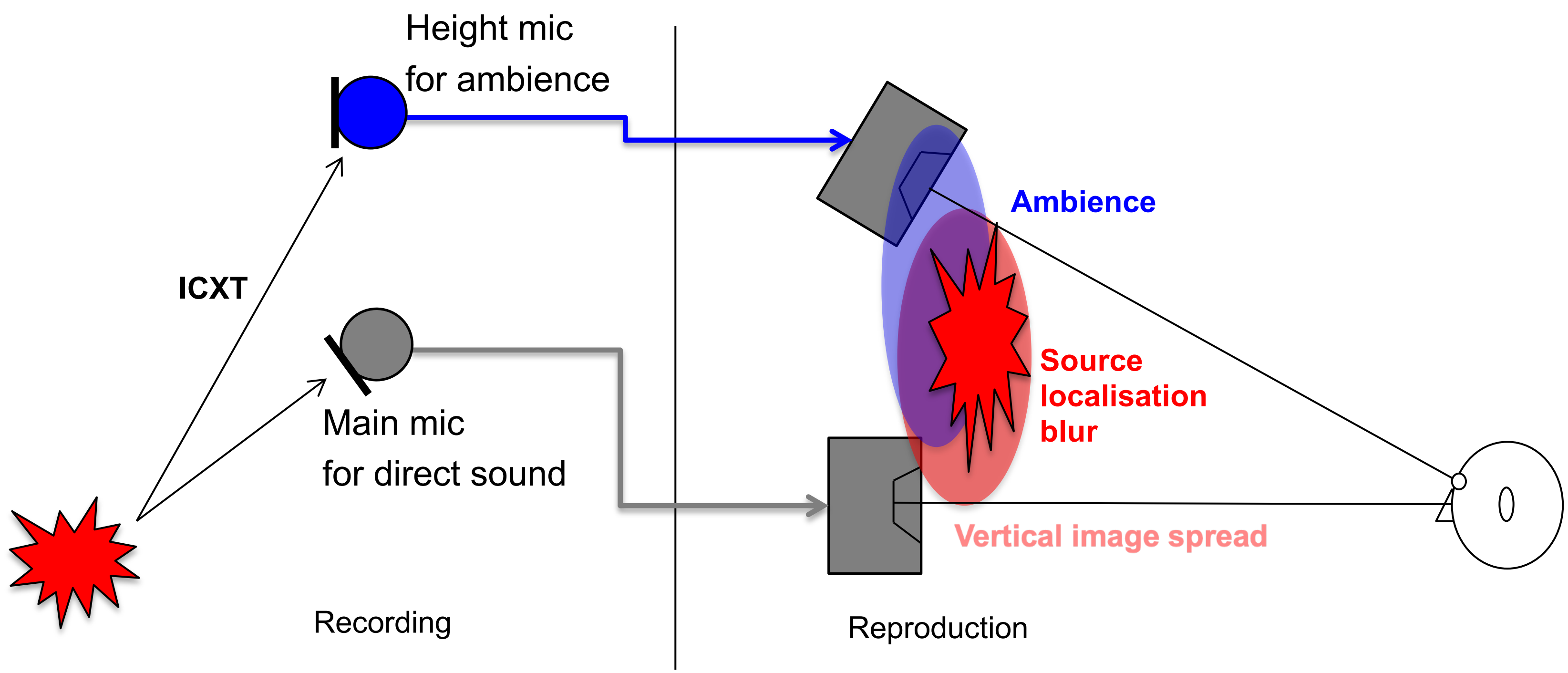
ICLD and ICTD of Interchannel Crosstalk (ICXT)

- Parameters and associated perceptual attributes

Parameters	Perceptual attributes
The Interchannel level difference (ICLD) and interchannel time difference (ICTD) of interchannel crosstalk (ICXT).	Horizontal and vertical image locatedness (i.e. ease of localization). Horizontal and vertical image spread.

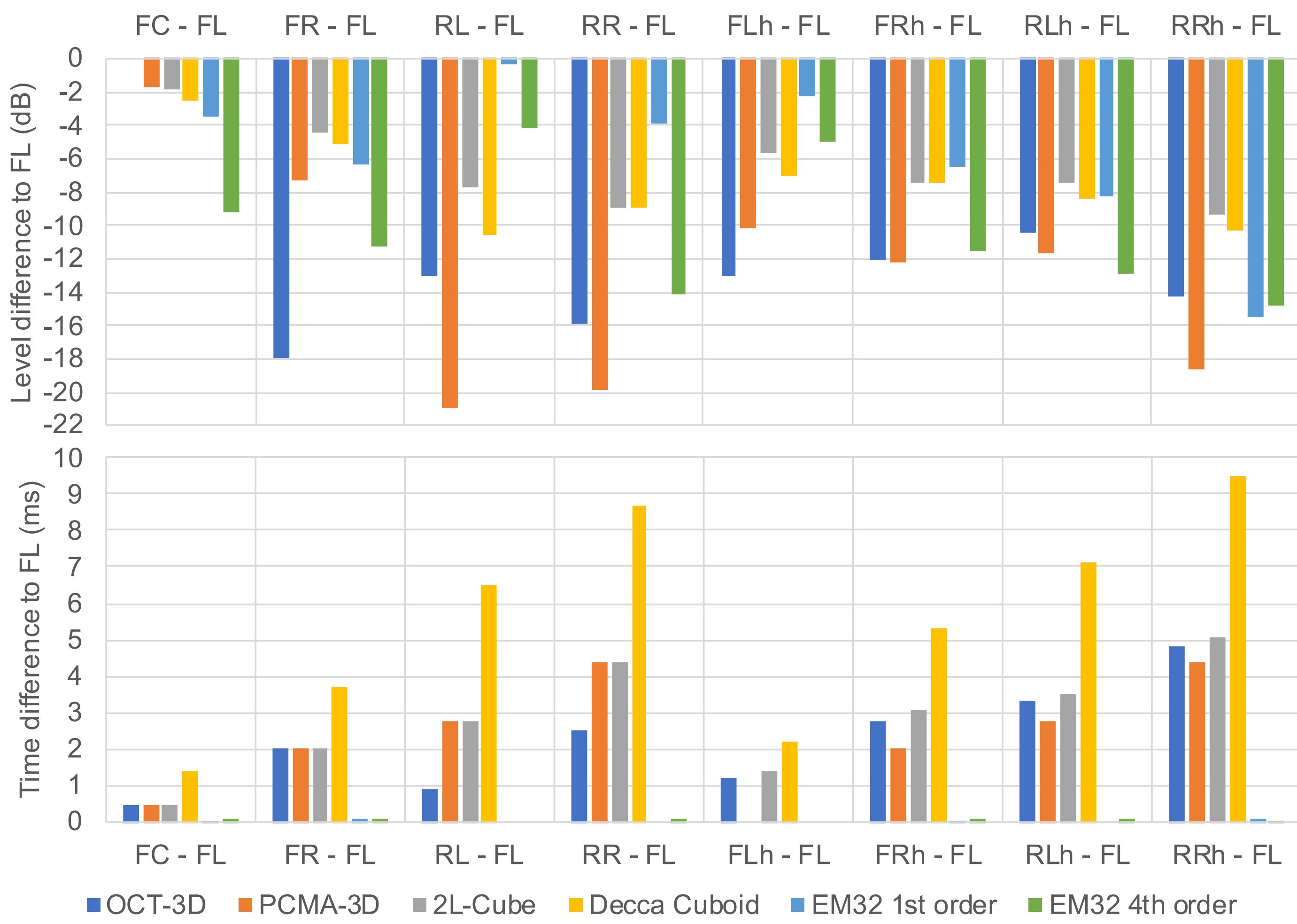
ICLD and ICTD of Interchannel Crosstalk (ICXT)

- Interchannel crosstalk (ICXT)
  - Direct sound captured by other microphones than the ones primarily responsible for source imaging.



ICLD and ICTD of Interchannel Crosstalk (ICXT)

ICLD and ICTD of ICXT to the signal of Front Left channel signal (closest to the source at 45°)



A higher level of ICXT (i.e. a smaller ICLD or a larger ICTD) suggests a stronger interference in source imaging stability (locatedness).

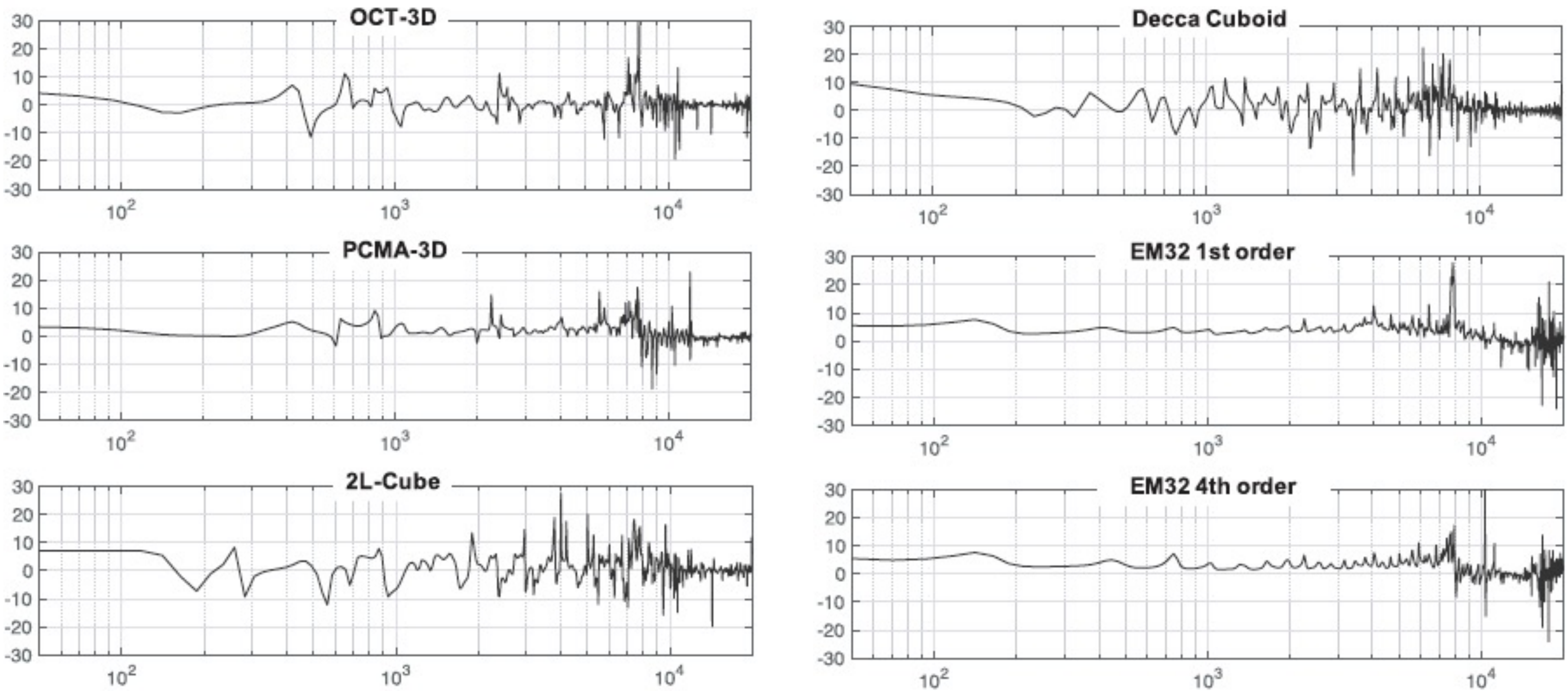
Spectral Distortion due to ICTX

- Parameters and associated perceptual attributes

Parameters	Perceptual attributes
Spectral Distortion due to height channel ICTX	Tonal colouration (comb-filter effect)

Spectral Distortion due to ICTX

- Spectral distortion of the ear signal resulting from adding the height layer to the main layer in reproduction.



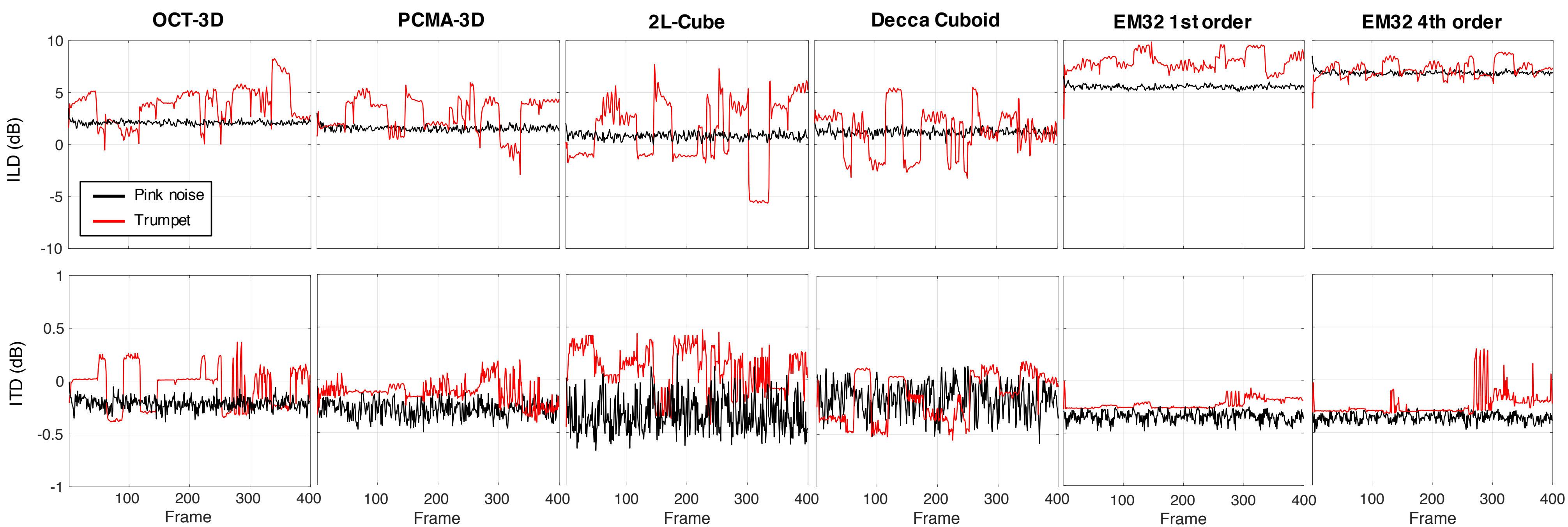
ILD and ITD Fluctuation Over Time

- Parameters and associated perceptual attributes

Parameters	Perceptual attributes
ILD and ITD fluctuations over time	Horizontal image spread (i.e. Apparent source width) Horizontal source movement (at a low fluctuation rate)

ILD and ITD Fluctuation Over Time

- Time-varying ILD and ITD at a low fluctuation rate (<15 Hz) represent source movement (unstable imaging), especially for musical signal with pitch changes.
- A high fluctuation rate suggests the perception of source width (see the result for pink noise).

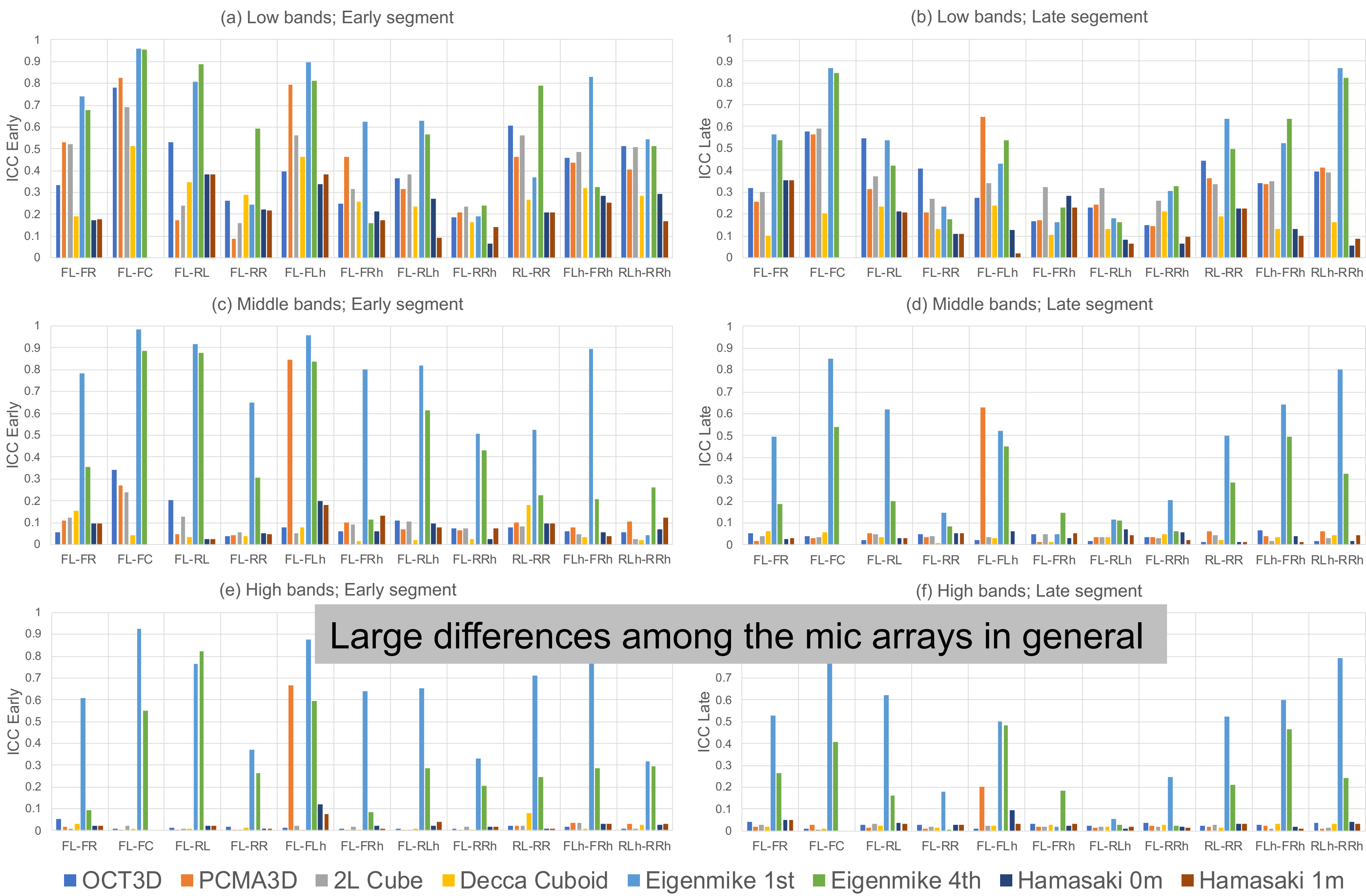


Interchannel Correlation Coefficient (ICC)

- Parameters and associated perceptual attributes

Parameters	Perceptual attributes
Interchannel Correlation Coefficient (ICC)	Horizontal and vertical image spread Listener envelopment Size of the listening area

Interchannel Correlation Coefficient (ICC)



Large differences among the mic arrays in general

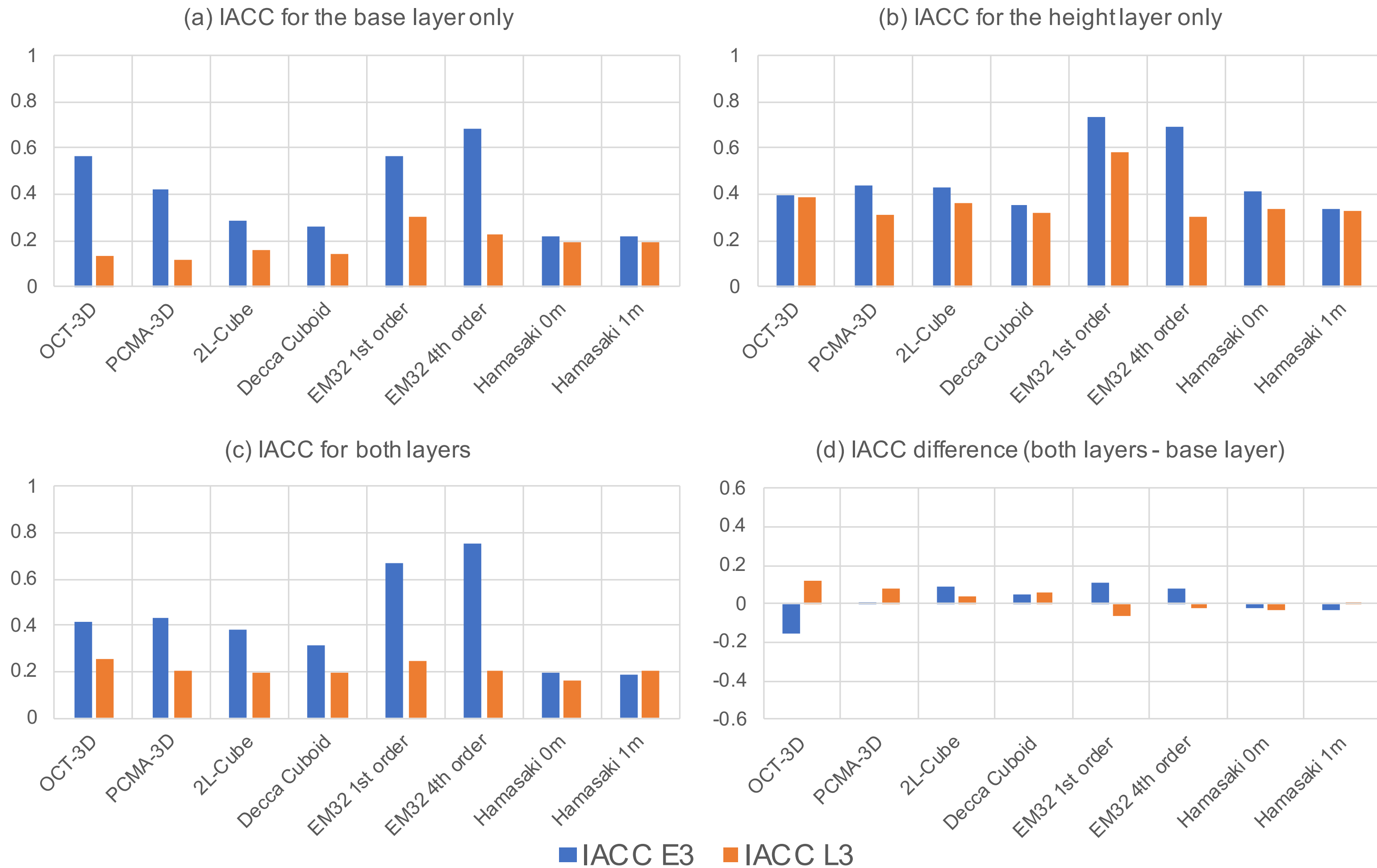
Interaural Cross-Correlation Coefficient (IACC)

- Parameters and associated perceptual attributes

Parameters	Perceptual attributes
Interchannel Cross-Correlation Coefficient (IACC)	Horizontal and vertical image spread (Early segment) Listener envelopment (Late segment)

Interaural Cross-Correlation Coefficient (IACC)

- The addition of the height layer does not vary IACC considerably.



Lee 2021

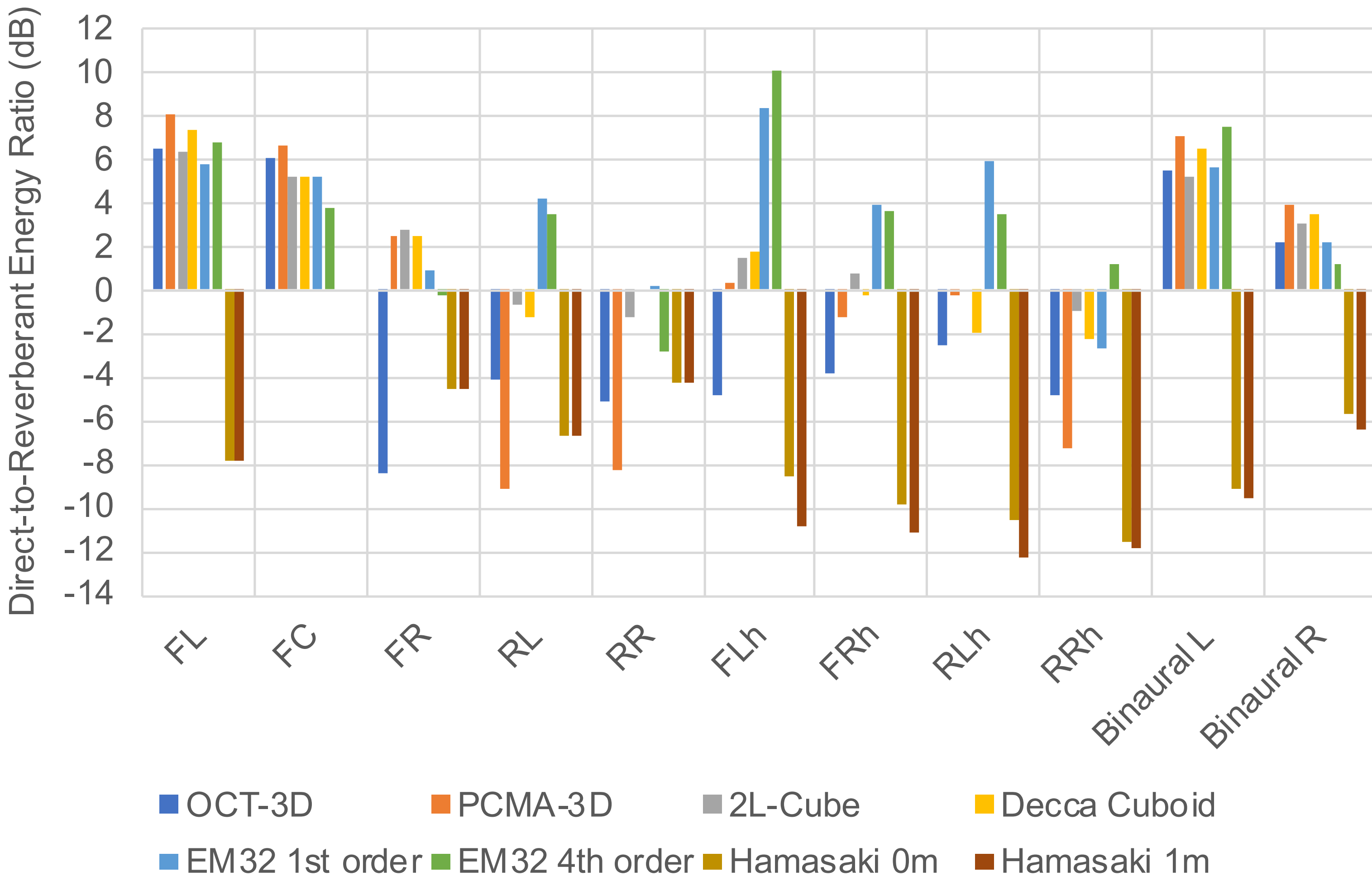
Direct-to-Reverberant Energy Ratio (DRR)

- Parameters and associated perceptual attributes

Parameters	Perceptual attributes
Direct-to-Reverberant Energy Ratio (DRR)	Auditory distance

Direct-to-Reverberant Energy Ratio (DRR)

- A lower DRR suggests a potentially greater depth/distance perception.



Discussions

- There were substantial differences among the investigated microphone arrays in the amount of both horizontal and vertical ICXT.
  - This was found to be associated to the differences in the amount of spectral distortion in the ear signal as well as in the magnitudes of ILD and ITD fluctuations over time.
  - From this, it is expected that the arrays would have audible differences in perceived timbral characteristics as well as the locatedness and spread of phantom image.

Discussions

- The differences in horizontal ICC between the early segments of the main layer impulse responses were large.
  - It is hypothesized that the arrays would have considerable differences in the perceived magnitudes of apparent source width (ASW) and the size of listening area.
- The differences in vertical ICC were considerable, but may not cause large perceptual differences in vertical image spread based on the findings from previous studies on vertical decorrelation effect.

Discussions

- The analysis of IACC suggests that the addition of the height layer to the base layer in reproduction would have little effect on ASW and LEV regardless of the array type, even though the two layers might have audible differences in those attributes when they are reproduced independently.

- The differences between the microphone arrays in the DRRs of ear-input signals resulting from the virtual nine-channel loudspeaker reproduction were around or below the just noticeable difference of perceived auditory distance (i.e., around 4 dB), even though the DRRs of individual microphone signals had considerably larger differences among the arrays.
  - Would the perceived source distance be determined by the channel-signal DRRs or ear-signal DRRs?

Discussions

Future works

- Verbal elicitation of perceptual differences among the microphone arrays to establish a set of attribute scales that will then be used for grading.
- Subjective ratings and comparisons against the objective results.
- Develop a statistical model for 3D acoustic recording quality evaluation.

Contact: h.lee@hud.ac.uk

APL Website: www.hud.ac.uk/apl

APL Facebook page:
https://www.facebook.com/applied.psychoacoustics.lab

APL YouTube channel:
https://www.youtube.com/channel/UCOPJ-N-b9dh1tE4-Xi36P2g