Microphone Arrays for Reverberant Spatial Audio Object Capture

Miguel Blanco Galindo1, Philip Jackson2, Philip Coleman3 and Luca Remaggi4

1University of Surrey, Guildford, Surrey GU2 7XH, UK
2p.jackson@surrey.ac.uk 3p.d.coleman@surrey.ac.uk 4l.remaggi@surrey.ac.uk

Introduction to the Reverberant Spatial Audio Object (RSAO)

- Goal: To recreate the reverberation characteristics of a given target room in a standard living room
- RSAO: Method to parameterise Spatial Room Impulse Responses (SRIRs) measured with a microphone array, allowing to edit the reverberation parameters for creative purposes (production) or for improved user experience (reproduction) as well as rendering it for arbitrary reproduction setups

Simulation method

- What is the best microphone array design to capture reverberation parameters? The objective is to evaluate the performance of different array designs to capture RSAO early reflection parameters
- Simulation tool allows to compare many array designs in various spaces without physically building them and the RSAO parameters can be evaluated against a ground-truth reference to assess the arrays’ performance
- Method:
  - Simulate ground-truth reverberation responses at microphone positions and at the center of the array using the Image Source Model (ISM)
  - Extract RSAO parameters from mic array SRIRs and ground-truth parameters from ISM at center of the array and evaluate their errors by reporting the number of correctly detected reflections out of a total of 25 for 2nd-order reflection and among those the Root Mean Square Error (RMSE) of the time of arrival (TOA) and direction of arrival in azimuth (DOA az) and elevation (DOA el)

Microphone array designs

- Seven array geometries were modelled using 48 microphones with two different design criteria: fixed spacing or fixed aperture (below)
  - Linear, Rectangular, Circular, Dual-circular, Spherical, Rigid Circular, Rigid Spherical

Results

- All the array geometries with the two design criteria were evaluated in two rooms (Audio Booth and Vislab). Sensor noise of -50 dB and sensor offset of ±2 mm were also modelled to resemble measurement uncertainty

Conclusions

- Number of detected reflections decreases with larger arrays. However, this effect diminishes in larger rooms where reflections are sparse
- Localization errors were smallest for the circular and dual-circular arrays in azimuth and for the spherical array in elevation re-emphasising its uniform resolution
- Delay-and-sum beamformer proved to be fairly robust in presence of sensor offset, especially for open arrays. Baffled arrays were limited by the additional SNR required to capture sound on the occluded side

Future work

- Evaluate the effect of number of microphones in performance and include cylindrical geometries
- Wavefield Analysis approach to extract early reflections and expand work to late reverberation (diffuse) or source with extinct

Acknowledgements

This work was supported by the EPSRC Programme Grant SIA: Future Spatial Audio for an Immersive Listener Experience at Home (EP/L016059/1) and the BFI as part of the BBC Audio Research Partnership