

Potential and Limits of a High-Density Hemispherical Array of Loudspeakers for Spatial Hearing and Auralization Research

William L. Martens, Denis Cabrera, Luis Miranda, Daniel Jimenez

The Faculty of Architecture, Design and Planning, The University of Sydney, Sydney, Australia
Email: william.martens@sydney.edu.au

Received 16 October 2014

Abstract

A hemispherical array of 196 independent loudspeakers has been constructed for laboratory research on spatial hearing, auralization, sound-field quality and spatial audio signal processing. The array of small loudspeakers is supported by a geodesic frame with a radius is 2.4 m. With this relatively high spatial density of sound sources, spatially quantized rendering of sound fields should be viable for a variety of applications, and this paper considers the potential and limits of this approach in relation to human spatial hearing capacities. The paper also considers the combined use of multiple loudspeakers for the rendering low-frequency components of sound sources in order to extend the low-frequency response of the system.

Keywords

Loudspeaker Array, Sound Field Control, Auralization, Spatial Hearing

1. Introduction

Advances in audio technology have made loudspeaker arrays with many channels increasingly viable in recent years, and such arrays have several applications in acoustics research, such as: investigating spatial hearing, developing spatial audio processing techniques, auralization, sonification, and even measurement applications such as the determination of directional diffusivity. An early instance of a loudspeaker array similar to the one examined in this paper was described in 1965 by Meyer *et al.* [1], with 65 loudspeakers in a hemispherical array, which was used to investigate auditory spatial perception. Similar hemispherical arrays have been constructed for research on a variety of scales in many laboratories—e.g. a 54-loudspeaker array used by Kleiner [2], or smaller arrays such as the 24-loudspeaker array by Hidaka [3]. More recently, larger arrays have become more viable, one example being the 157-loudspeaker array in a small sound-absorptive room (again with no floor loudspeakers) [4] which has been used to examine sound field control and spatial hearing at Tohoku University. A 350-loudspeaker array in IRCAM's variable acoustics performance hall is another related example [5]. These similar large multichannel loudspeaker research-focused installations allow diverse approaches to sound field control, and much recent research in this field has focused on refining techniques such as higher-order Ambisonics, wave field synthesis, vector-based amplitude panning, and others, for reproduction using large loud-

peaker arrays. While these techniques have a great potential, they also have difficult limitations, and so until the loudspeaker count becomes extremely high, decisions on spatial audio processing will inevitably involve pragmatics and compromises.

However, for some applications, a very simple approach to spatial control may be effective using discrete (or spatially quantized) multichannel simulation. Like the other options, this approach has limitations, but also has advantages. One limitation is that the wave-front shape is determined by the distance (and radiation pattern) of each loudspeaker—*i.e.* there is no control over wave-front curvature that might enable the simulation of source distance. This limitation should be tolerable if near-field sound sources are not being simulated, and if the loudspeaker array is sufficiently distant to be considered to be in the far field (at least from a spatial hearing perspective). Another limitation is that the angular resolution of simulated sources can be no finer than that of the loudspeaker array—and this limitation may also be tolerable if the loudspeaker array is sufficiently dense for the application. The advantages of Discrete Multichannel Simulation (DMS) [6] include the minimization of required signal processing (the reproduction system is parsimonious), the avoidance of contrast between virtual sound sources (e.g., panned between two or three loudspeakers) and sources originating from loudspeakers, the creation of a sound field that avoids serious “sweet spot” problems, and the avoidance of problems in high-frequency sound field control which often pose difficulties in Higher Order Ambisonics (HOA) and Wavefield Synthesis (WFS). This paper describes a 196-loudspeaker array, and examines the implications of using it in this mode. It also examines how low-frequency response can be improved with this array by loudspeaker clustering.

2. Description of the Facility

The facility is housed in a room that was originally designed as an artificial sky (with a hemispherical concrete dome ceiling, 6.5 m diameter). Since physical artificial skies no longer offer any advantages over computer-based alternatives, this room has been redeployed for the loudspeaker array. The array is supported by a geodesic frame (3V 5/9 class I icosahedral geometry), with a loudspeaker on every vertex and in the middle of each edge at and above the equator. This positions the 196 loudspeakers approximately 2.15 m from the 5/9 sphere’s centre. The volume behind the geodesic frame is filled with porous sound absorption, so as to create a nearly anechoic environment. Anthony Gallo ‘Micro’ spherical loudspeakers (102 mm diameter) are used, chosen in part to minimize specular acoustic reflections. The loudspeaker array is illustrated in **Figure 1**.

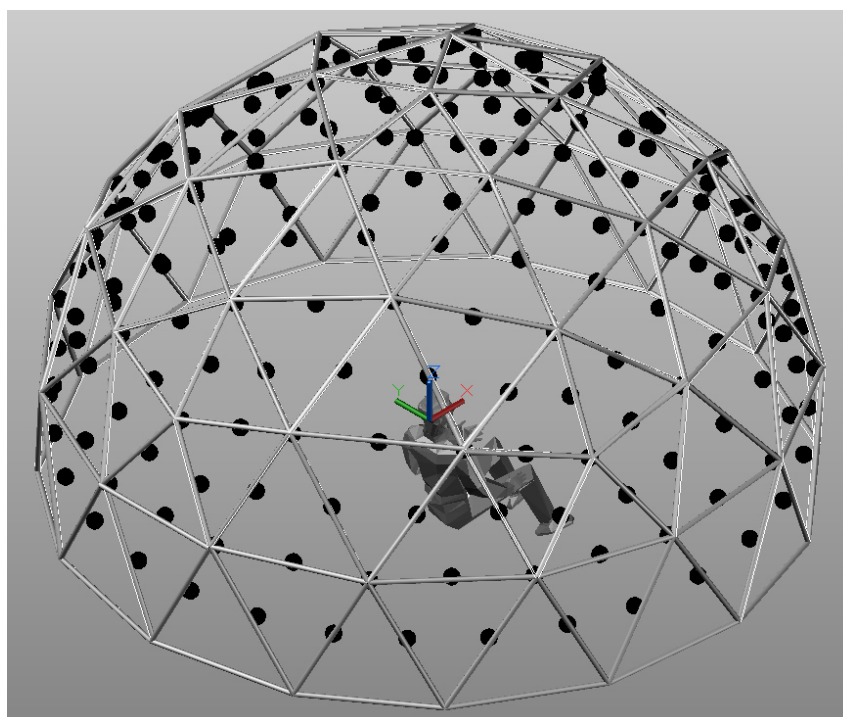


Figure 1. Arrangement of the 196 loudspeakers on the geodesic support frame.

The spatial arrangement of loudspeakers in this array is derived from the geodesic geometry, rather than being based on an auditory polar coordinate system or a strictly regular distribution (not that even distribution is possible beyond Platonic solids). There are 30 loudspeakers on the equator (0° elevation) at approximately 12° increments. In fact, both of the lowest two rings of loudspeakers, those nearest ear level, divide the circumference of the spherical array up into 30 azimuthal sections each; however, the azimuth angles of these sets of loudspeakers are offset by 6-degree to one another, so that if the 10-degree modulation in elevation is ignored, the loudspeakers in these two rings taken together cover 60 azimuthal sections, presenting sound sources at 6-degree intervals in azimuth. These two rings of speakers define a regular triangular mesh with inter-speaker distances for adjacent speakers averaging to 0.47 m.

The third and fourth lowest rings of speakers are positioned nearly 20 degrees above ear level, and are spaced irregularly; however these sets of respectively 10 and 20 loudspeakers are interspersed so as to present sound from azimuth angles at 12-degree intervals. In contrast to the first two pair of rings, this pair of rings has quite close positioning in elevation (each removed from 19.5 degrees of elevation by approximately 1 degree). These two rings of speakers define an irregular triangular mesh with inter-speaker distances for adjacent speakers again averaging to 0.48 m, as is expected for a geodesic structure (with all points roughly equidistant). This value is also just slightly greater than what would be expected were all 30 loudspeakers positioned at a single elevation angle and spaced in azimuth at 12-degree intervals, which for the sphere radius of 2.17 m would give an inter-speaker distance for adjacent speakers of 0.46 m. The fifth, sixth and seventh lowest rings of speakers also are placed irregularly, but these sets of respectively 5, 10, and 15 loudspeakers again are interspersed so as to present sound from azimuth angles at 12-degree intervals. These three rings are all positioned close to 28.7 degrees above ear level.

Now if the relatively constant distance between adjacent loudspeakers is considered, an opportunity for improved low-frequency reproduction presents itself, since the combination of loudspeaker signals allows greater sound reproduction levels to be achieved. The anechoic response of the employed “Micro” spherical loudspeakers falls off rapidly as the frequency of the applied signal is decreased below 250 Hz. However, since these small loudspeakers are positioned with relatively high spatial density in the hemispherical array, it was proposed to use multiple adjacent loudspeakers to achieve a better low-frequency reproduction level (avoiding the distortion that inevitably would result from attempting to drive a single loudspeaker at too high amplitude at lower frequencies). To illustrate the potential for success of this approach, a simulation was executed using local subsets of loudspeakers in triangular and hexagonal configurations, which is the topic of the next section of this paper.

Extending the Low-Frequency Response of the System

First, the baseline performance is observed when a single loudspeaker is used to reproduce an octave band of noise with a centre frequency of 125 Hz. In this low-frequency band, the RMS amplitude drops almost 2 dB from the front to the rear of the targeted listening area (defined here as a square patch with each side measuring 50 centimeters). **Figure 2** shows this modulation in RMS amplitude using a surface mesh plotted over the Left-Right positioning and the Front-Rear positioning of the observation point.

Relative to the RMS amplitude of the 125-Hz octave band of noise produced in the listening area using a single loudspeaker, there is on average a 9.5 dB boost when three adjacent loudspeakers in a triangular grouping are driven with identical signals. However, as the observation point varies across the surface of the target listening area, there is a modulation of RMS amplitude of the combined loudspeaker signals due to the change in the relative time of arrival of the three loudspeaker signals as the observation point varies along the Left-Right and the Front-Rear axes. In order to graphically illustrate the nature of this modulation, the spatial deviation in amplitude was calculated by taking the difference between the RMS amplitude values of single-loudspeaker and triple-loudspeaker reproduction surfaces after each was centred by subtracting the respective spatial averages from each (which averages differed by 9.5 dB, as indicated above). **Figure 3** plots these deviation values on a surface using a different color map than that used in **Figure 2** in order to distinguish between RMS amplitude and relatively small spatial variation in RMS amplitude between two reproduction conditions. The observed amplitude deviations between reproduction conditions are of course smaller than what a human listener could detect, especially when considered in relation to the changes in RMS amplitude that occur over the listening area for a given reproduction condition (as was illustrated for the single-loudspeaker case in **Figure 2**).

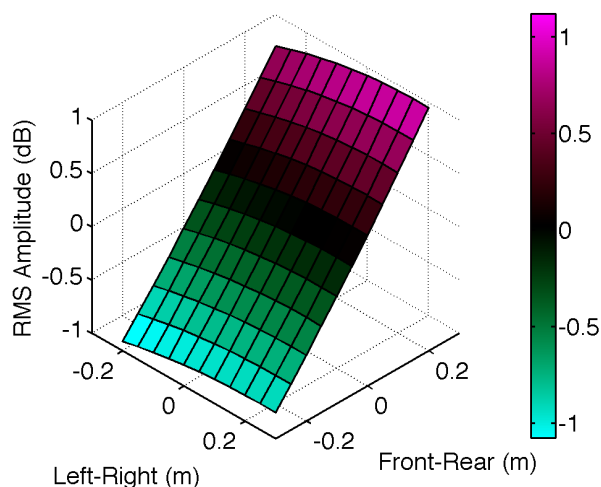


Figure 2. RMS amplitude of a single loudspeaker observed over a targeted listening area.

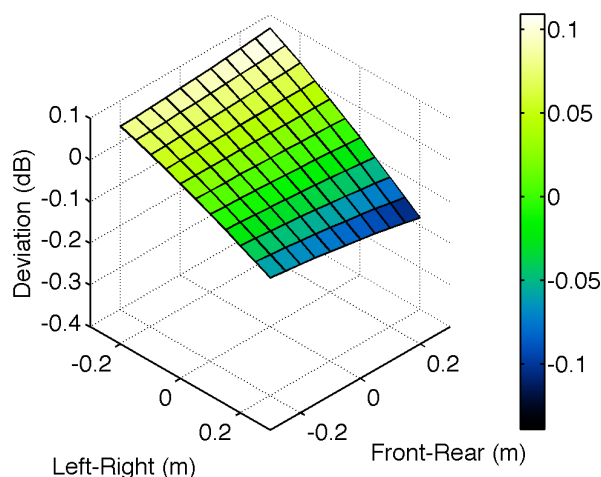


Figure 3. Deviation in the RMS amplitude reproduced by three loudspeakers in a triangular configuration, relative to that of a single loudspeaker observed over the same listening area.

Note also that the surface plotted in **Figure 3** is not symmetric on the Left-Right axis since the orientation of the listening area is aligned with one of the three loudspeakers rather than with the center of the described triangle. This allows a straightforward comparison of the triangular configuration with a hexagonal configuration that employs seven adjacent loudspeakers to achieve an even greater boost in the low-frequency performance of the system. In this case, the spatial average of RMS amplitude for the 125-Hz octave band of noise is 12 dB greater than that produced over the listening area using a single loudspeaker. As can be seen in **Figure 4**, the pattern of results in the hexagonal configuration case is quite similar to that observed in the triangular configuration case. This hexagonal case has the additional advantage that the loudspeaker in the centre of the hexagon is aligned with the average spatial direction of the whole hexagonal configuration, so that the boosted low-frequency content in a broadband signal is not likely to be perceived by the listener as shifted spatially relative to the higher-frequency components of that broadband signal. What is left out of the current analysis here is the consideration of interaural level differences (ILDs) that are introduced by the presence of a human head in the listening area. Suffice it to say that these ILDs will be affected by the deviations in RMS amplitude between different reproduction conditions (single versus triangular versus hexagonal), but it should be noted that these deviations will be quite small relative to the modulations in ILD that are associated with small rotations of a listener's head during listening.

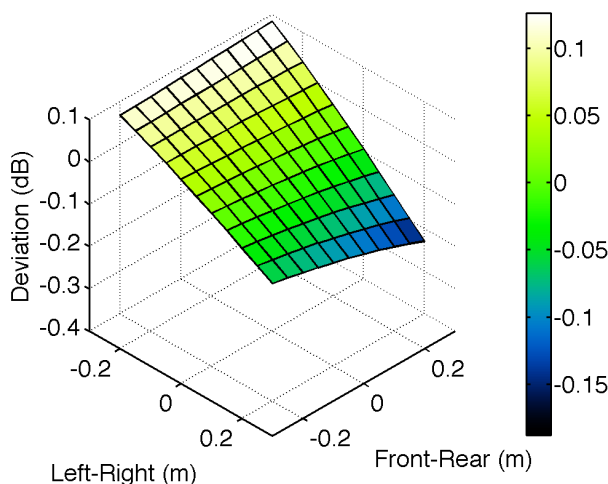


Figure 4. Deviation in the RMS amplitude reproduced by seven loudspeakers in a hexagonal configuration, relative to that of a single loudspeaker observed over the same listening area.

3. Discussion

While the triangular and hexagonal loudspeaker configurations employed here for presenting low-frequency sound are spatially extended along the vertical dimension of the hemispherical array (over a range of elevation angles), this does not present too great a problem for human localization of displayed sound sources. The acoustic cues to sound localization that are due to the transformation of incident sound by the head and pinna depend upon higher-frequency information than that in the 125-Hz octave band examined here. Indeed, such low-frequency sound tends to be heard as arriving from low elevation angles regardless of the actual angle of arrival [6]. It is particularly worth noting that the direction of arrival for reproduced low-frequency content is regarded as critically important for appreciating a natural quality for a virtual acoustic environment presented by such a loudspeaker array. While it is beyond the scope of this paper to discuss the results of listening experiments that address issues faced in multichannel loudspeaker reproduction, the reader is directed to a recent review of such studies [7]. In that review the relative value of the currently described Discrete Multichannel Simulation (DMS) is compared to two other multichannel sound-field synthesis techniques, Higher Order Ambisonics (HOA) and Wave Field Synthesis (WFS). Some advantages in terms of control of source range in simulated virtual environments have been reported [8] using a similar approach to sound-field rendering, albeit with a much smaller loudspeaker count (that study employed a 15-loudspeaker hemispherical array driven by the Pioneer Sound Field Controller (PSFC) [9]). The PSFC that was installed in the ‘Synthetic World Zone’ at the University of Aizu Multimedia Center (in the Fukushima Prefecture of Japan) presented simulated indirect sound that was based upon the spatio-temporal distribution of 480 discrete reflections within each of the 15 loudspeaker channels, sharing the spatial directions of the loudspeakers. Each of these 7200 reflections was delivered via only one of the loudspeaker channels. The simulated spatio-temporal distribution of reflections was based upon that of an existing concert hall in Japan, so that a comparison between the actual sound field and the simulated sound field was enabled. The newly developed hemispherical array of 196 independent loudspeakers that has been constructed at the University of Sydney can reproduce the spatiotemporal distribution of reflections from models of existing architectural spaces, but with reflections clustered into 196 distinct spatial angles for discretely simulated reflections, as illustrated for a small room simulation in **Figure 5**.

4. Conclusion

In conclusion, the hemispherical loud speaker array described in this paper has enabled laboratory research on spatial sound reproduction and perception. The envisioned use in spatial hearing research, particularly in auralization and evaluation of sound-field quality, seems particularly promising. It is hoped that such high-resolution spatially quantized rendering of sound fields will allow important practical questions to be answered as well, in-

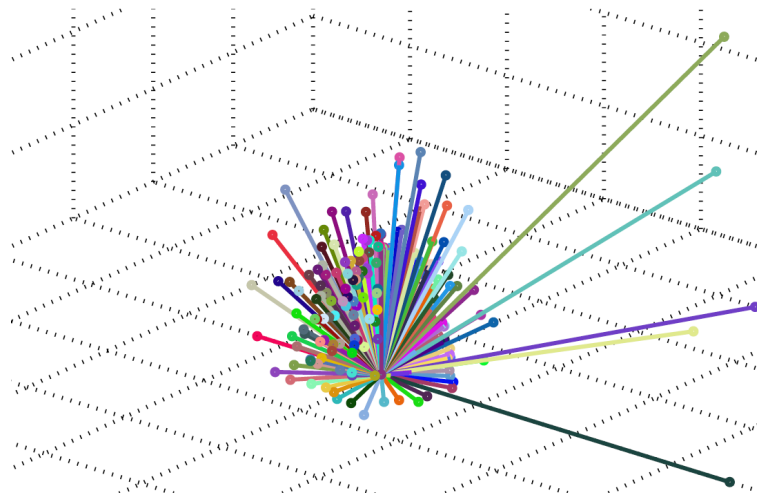


Figure 5. A hedgehog plot of RMS amplitude, spatially quantised to the 196 loudspeaker angles, based on a computer-modelled impulse response from a small room. Colours are randomly assigned (to allow the various lines to be distinguished). Each of the lines in this plot is drawn from the listening position to a point in a direction indicating the angle of incidence of the loudspeaker reproducing component indirect sound, and at a distance from the origin indicating the RMS amplitude observed for the indirect sound arriving at the listening position from that direction.

cluding the evidence-based determination of how many loudspeakers are sufficient for particular applications, since progressive increase in the number may lead to diminishing returns as the number approaches the maximum of 196 loudspeakers available.

Acknowledgements

This research was supported in part by a grant from the Australian Research Council. Additional funding was provided by The Faculty of Architecture, Design and Planning, The University of Sydney.

References

- [1] Meyer, E.W., Burgdorf, W. and Damaske, P. (1965) Eine Apparatur zur elektroakustischen Nachbildung von Schallfeldern. Subjektive Hörwirkungen beim Übergang Kohärenz-Inkohärenz. [An Apparatus for the Electroacoustic Simulation of Sound Fields. Subjective Auditory Effects at the Transition between Coherence and Incoherence]. *Acustica*, **15**, 339-344.
- [2] Kleiner, M. (1981) Speech Intelligibility in Real and Simulated Sound Fields. *Acustica*, **47**, 55-71.
- [3] Hidaka, T. (1993) Sound-Field Simulator for Room Acoustic Design and Assessment—Introduction of Wave-Theoretical Treatment to Synthesized Sound. *Journal of the Audio Engineering Society*, **41**, 914-919.
- [4] Okamoto, T., Cabrera, D., Noisternig, M., Katz, B.F.G., Iwaya, Y. and Suzuki, Y. (2010) Improving Sound Field Reproduction Based on Higher-Order Ambisonics in a Small Room with a 157-Loudspeaker Array. *Proceedings of the 2nd International Symposium on Ambisonics and Spherical Acoustics*, Paris, 6-7 May 2010.
- [5] Noisternig, M., Carpenter, T. and Warusfel, O. (2012) ESPRO 2.0—Implementation of a Surrounding 350-Loudspeaker Array for Sound Field Reproduction. *Proceedings of the 25th Audio Engineering Society UK Conference*, York, 25-27 March 2012.
- [6] Cabrera, D. and Morimoto, M. (2007) Influence of Fundamental Frequency and Source Elevation on the Vertical Localization of Complex Tones and Complex Tone Pairs. *Journal of the Acoustical Society of America*, **122**, 478-488. <http://dx.doi.org/10.1121/1.2736782>
- [7] Martens, W.L. (2014) Display of Three-Dimensional Virtual Acoustic Sound Fields for Human Listeners: Value of Discrete Multichannel Simulation. *Proceedings of the 21st International Congress on Sound and Vibration*, Beijing, 13-17 July 2014.

- [8] Honno, K., Cohen, M. and Martens, W.L. (2001) Psychophysically-Derived Control of Source Range for the Pioneer Sound Field Controller. *Proceedings of the 110th Convention of the Audio Engineering Society*, Preprint 5313, Amsterdam.
- [9] Amano, K., Matsushita, F., Yanagawa, H., Cohen, M., Herder, J., Martens, W., Koba, Y. and Tohyama, M. (1998) A Virtual Reality Sound System Using Room-Related Transfer Functions Delivered through a Multispeaker Array: The PSFC at the University of Aizu Multimedia Center. *Transactions of the Virtual Reality Society of Japan*, **3**.