AMBISONICS SYMPOSIUM 2009
June 25-27, Graz

# EVALUATION OF A CONCENTRIC RIGID AND OPEN SPHERICAL MICROPHONE ARRAY FOR SOUND REPRODUCTION

Abhaya Parthy[1], Craig Jin[2], André van Schaik[2]

[1] School of Information Technologies, The University of Sydney, Australia (aparthy@it.usyd.edu.au)
[2] School of Electrical and Information Engineering, The University of Sydney, Australia (craig@ee.usyd.edu.au)

**Abstract:** *We present an empirical evaluation of a third-order sound field recording and reproduction system which has been designed to operate over a frequency range of 900 Hz to 16 kHz. The system consists of a spherical microphone array which is a concentric rigid and open spherical microphone array (64 microphones in total) and also a spherical loudspeaker array (24 loudspeakers) in a hemi-anechoic room. The concentric rigid and open spherical microphone array was designed to improve the frequency range of the standard spherical microphone array. We present some measurements comparing performance to an ideal third-order sound field recording and reproduction system.*

Key words: microphone arrays, high-order ambisonics, sound field recording and reproduction

## 1 INTRODUCTION

A system for recording and reproducing a real sound field is useful in many applications including audio displays, virtual reality, audio-only gaming, cinema, and auditory research. The most widely used system for recording and reproducing a sound field is the first-order Ambisonics system originally developed by Gerzon [4]. The first-order Ambisonics system uses a sound field microphone to record a sound field to a first-order spherical harmonic representation. This spherical harmonic representation of the sound field is then used to drive an array of loudspeakers to recreate the original sound field, correct to a first-order approximation. The first-order Ambisonics system is in wide use and performs quite well in most applications, however, there is a significant benefit in using higher order systems as they have an increased spatial fidelity, especially at higher frequencies, and a larger listening area at the centre of the loudspeaker array [12]. Much research has been done on higher order sound field recording and reproduction systems, including work by Poletti [11], Li et al. [6], Abhayapala and Ward [12, 1], Daniel et al. [3], and Bertet et al. [2].

We have built a sound field recording and reproduction system capable of operating at up to third-order in the frequency range of 900 Hz to 16 kHz, up to second-order down to 400 Hz, and at first-order down to 60 Hz. The system uses a 24-loudspeaker spherical loudspeaker array for reproduction, and a 64-microphone concentric rigid and open spherical microphone array (CRO-SMA) for recording. The spherical loudspeaker array, shown in Figure 2, is situated inside a hemi-anechoic room, which has walls and a ceiling that are anechoic down to 100 Hz, and a floor which has thick carpet and two underlays to minimise the floor reflections. The CRO-SMA, shown in Figure 1, consists of two concentric spherical microphone arrays, one inside the other, each having 32 microphones. The inner of the two spherical microphone arrays is baffled by a rigid scatterer and the outer open array is designed to be acoustically transparent. The CRO-SMA has been previously characterised for use as a beamformer and for near-field acoustic holography, and a detailed analysis of its design and performance is presented in [8, 9, 10].

We present an empirical evaluation of the combined CRO-SMA and loudspeaker array system when it is used as a

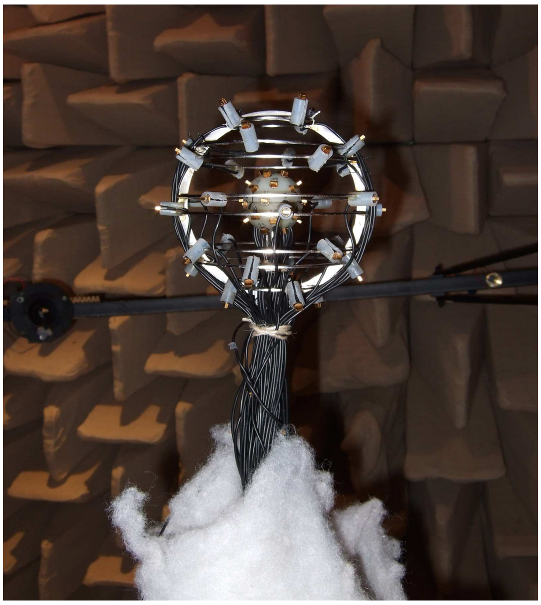


**Figure 1:** This figure shows a photo of the 64-microphone rigid and open spherical microphone array.

sound field recording and reproduction system. The performance of this system is evaluated by measuring its ability to record and reproduce a sound field generated by a single plane wave sound source. The CRO-SMA is placed at the centre of the spherical loudspeaker array and is used to measure a plane-wave sound field. The centre of the loudspeaker array is termed the "sweet-spot", because it is the position at which the recreated sound field is the most accurate. The measured plane-wave sound field is compared to ideal first, second and third-order plane-wave sound fields.

The main contribution of this paper is that we show that a dual, concentric SMA provides a solid basis for a practical and broadband three-dimensional third-order sound field recording and reproduction system, and we present empirical measurements comparing the performance of the system to an ideal third-order sound field recording and reproduction system.

## 2 METHODS

The performance of the sound field recording and reproduction system we have constructed was evaluated by measuring its ability to record and reproduce a sound field generated by a single plane wave sound source. The CRO-SMA was placed at the centre of the spherical loudspeaker array, and impulse response measurements were measured from each loudspeaker on the spherical loudspeaker array to each microphone on the CRO-SMA. An additional loudspeaker was then used to simulate a plane wave sound source and impulse response measurements were measured from it to the CRO-SMA. The recorded impulse response measurements from the loudspeaker simulating a plane-wave source were then used to calculate the speaker signals for the loudspeaker array to recreate the recorded plane-wave sound field. Using these calculated signals for the speakers in the loudspeaker array and the original impulse responses measured from each loudspeaker in the loudspeaker array to each microphone in the CRO-SMA, we simulated the recording and reproduction of the plane-wave sound field. Snapshots of images of the recorded sound field are presented in our results and compared to ideal sound fields. In this paper we use the CRO-SMA naively in that we only use the microphone signals recorded from the outer open array for frequencies below 2840 Hz and we only use the microphone signals recorded from the inner rigid array for frequencies above 2840 Hz. We have previously shown in [10] how the microphone signals recorded from both arrays can be combined to improve the overall performance of the CRO-SMA across all frequencies for near-field acoustic holography and we are now in the process of implementing the same algorithms for sound field recording and reproduction.

### 2.1. System and Equipment Setup

The CRO-SMA, shown in Figure 1, consists of a rigid inner spherical microphone array of radius 1.63 cm and an outer open spherical microphone array of radius 6.0 cm. The rigid and open spherical microphone arrays each have 32 DPA type 4060-BM omni-directional microphones mounted in an approximately equally spaced arrangement as given by [5]. When the CRO-SMA is used to beamform at third-order it has a signal-to-noise ratio (SNR) that is greater than 30.0 dB for a frequency range of 900 Hz to 16 kHz, and this SNR can be maintained by reducing the beamforming order to second-order below 900 Hz and to first-order below 400 Hz down to 60 Hz. The signals from the 64 microphones in the CRO-SMA are amplified by eight 8-channel Digidesign PRE preamplifiers, which have digitally controlled gains so that all channels are uniformly amplified. The output from the preamplifiers are fed to four 16-channel Apogee AD-16X analogue-to-digital convertors which were sampling the signals at 48 kHz with a 24-bit resolution. The output from the analogue-to-digital convertors is in ADAT format and is sent to a RME ADI-648 MADI-to-ADAT converter, which converts the eight 8-channel ADAT signals to one MADI format signal that transports all the 64-channels of audio. The 64 microphone signals, in MADI format, are recorded on a standard personal computer (PC) using a RME Hammerfall DSP MADI PCI sound card. The 64 microphone signals can then be processed on a PC to create 24 loudspeaker signals to recreate the recorded sound field. The 24 loudspeaker signals are output from the PC's RME Hammerfall DSP MADI PCI sound card in MADI format, and are converted into ADAT format using a RME ADI-648 MADI-to-ADAT converter. The ADAT format loudspeaker signals have a sample rate of 48 kHz with a 24-bit resolution, and are fed to two 16-channel Apogee DA-16X digital-to-analogue convertor. The 24 analogue signals from the digital-to-analogue converters are fed to six 4-channel Lab Gruppen C series power amplifiers which power 24 uncalibrated Tannoy V6 loudspeakers. The 24 loudspeakers are mounted approximately equally spaced around a sphere with a radius of 2.8 m, however, due to height restrictions in the hemi-anechoic room, the top and bottom loudspeakers are closer to the centre of the loudspeaker array and their signals are appropriately time delayed. There are also 4 Whise Master Entertainer 319A sub-woofers that sit on the floor around the spherical loudspeaker array which are used to reproduce the low frequencies. These sub-woofers were not ready at the time the measurements were being made and are not included in the results presented in this paper, however, they are being used now and anecdotal evidence suggests that they significantly improve the sense of presence and the spatial image.

### 2.2. Sound Field Signal Processing

We assume that the CRO-SMA records a sound field in which all the sound sources are in the far-field, and we assume that the loudspeakers are ideal plane wave sources. To recreate the sound field recorded by the CRO-SMA, the loudspeakers are fed signals which are obtained by processing the signals recorded by the microphones on the CRO-SMA. This can be written as a matrix equation and is given by [12, 7]

$$\mathbf{Pa} = c\mathbf{b}, \tag{1}$$

**Figure 2:** This figure shows a photo of the spherical loudspeaker array inside the hemi-anechoic room.

where $c$ is a constant, $\mathbf{a}$ is a vector of unknown weights to be assigned to each loudspeaker, $\mathbf{P}$ is the loudspeaker decoding matrix, and for a second-order system is given by

$$\mathbf{P} = \begin{pmatrix} Y_0^0(\Omega_1) & Y_0^0(\Omega_2) & \cdots & Y_0^0(\Omega_L) \\ Y_1^{-1}(\Omega_1) & Y_1^{-1}(\Omega_2) & \cdots & Y_1^{-1}(\Omega_L) \\ Y_1^0(\Omega_1) & Y_1^0(\Omega_2) & \cdots & Y_1^0(\Omega_L) \\ Y_1^1(\Omega_1) & Y_1^1(\Omega_2) & \cdots & Y_1^1(\Omega_L) \\ \vdots & \vdots & \ddots & \vdots \\ Y_N^{-N}(\Omega_1) & Y_N^{-N}(\Omega_2) & \cdots & Y_N^{-N}(\Omega_L) \\ \vdots & \vdots & \ddots & \vdots \\ Y_N^N(\Omega_1) & Y_N^N(\Omega_2) & \cdots & Y_N^N(\Omega_L) \end{pmatrix}, \quad (2)$$

where, $\Omega_l = (\theta_l, \phi_l)$ is the spherical angular position of loudspeaker $l$ in the spherical loudspeaker array, $L$ is the number of loudspeakers in the spherical loudspeaker array, $Y_n^m(\cdot)$ are the spherical harmonic functions defined as

$$Y_n^m(\Omega) = Y_n^m(\theta, \phi) = \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) e^{im\phi}, \quad (3)$$

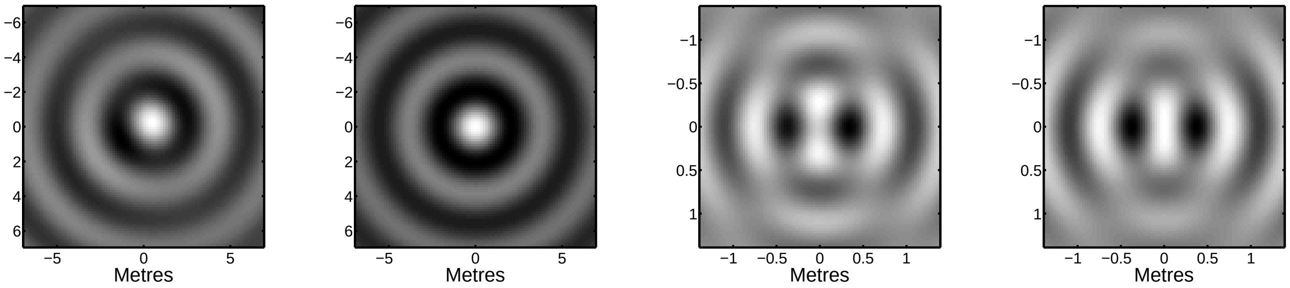
and $\mathbf{b}$ is the spherical harmonic decomposition of the sound field recorded by the CRO-SMA, and for a system with order $N$ is given by

$$\mathbf{b} = \begin{pmatrix} \sum_{j=1}^{M} \alpha_j p(k, \Omega_j) \frac{1}{b_0} Y_0^0 \\ \sum_{j=1}^{M} \alpha_j p(k, \Omega_j) \frac{1}{b_1} Y_1^{-1} \\ \sum_{j=1}^{M} \alpha_j p(k, \Omega_j) \frac{1}{b_1} Y_1^0 \\ \sum_{j=1}^{M} \alpha_j p(k, \Omega_j) \frac{1}{b_1} Y_1^1 \\ \vdots \\ \sum_{j=1}^{M} \alpha_j p(k, \Omega_j) \frac{1}{b_N} Y_N^{-N} \\ \vdots \\ \sum_{j=1}^{M} \alpha_j p(k, \Omega_j) \frac{1}{b_N} Y_N^N \end{pmatrix}, \quad (4)$$

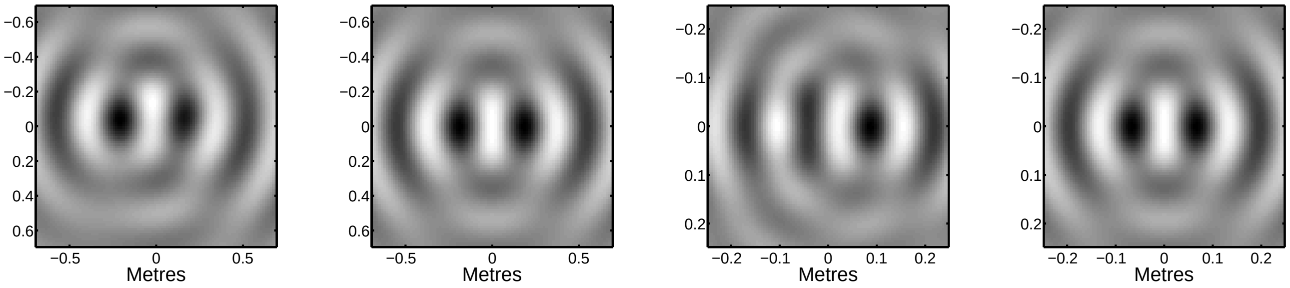
where, $k = \frac{2\pi}{\lambda}$ is the wave number, $\lambda$ is the wavelength, $\alpha_j$ is a discrete integration weighting, $\Omega_j$, is the angular position of microphone $j$ on the spherical microphone array, $M$ is the number of microphones in the spherical microphone array, $p(k, \Omega_j)$ is the pressure recorded by the microphone at position $\Omega_j$ on the spherical microphone array, and $b_n(k, r, a)$, is defined as

$$b_n(k, r, a) = 4\pi i^n \left( j_n(kr) - \frac{j_n'(ka)}{h_n'(ka)} h_n(kr) \right), \quad (5)$$

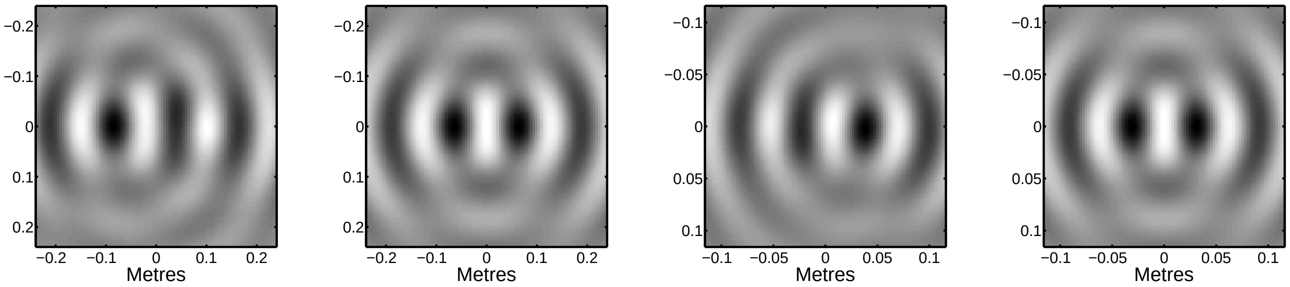
where $r$ is the radius of the spherical microphone array, $a$ is the radius of the rigid spherical baffler, $j_n(\cdot)$, $j_n'(\cdot)$, $h_n(\cdot)$, and $h_n'(\cdot)$ are the spherical Bessel function, the spherical Hankel function of the second-kind, and their derivatives, respectively.



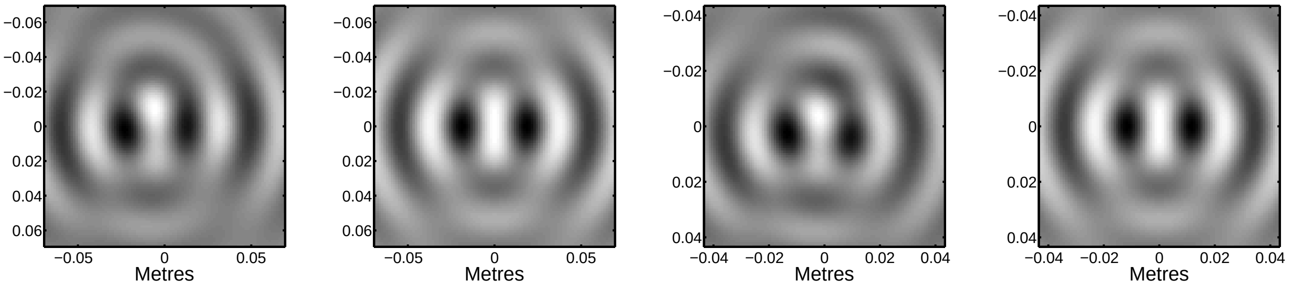
(a) Meas., 100 Hz, 1st ord. (b) Theo., 100 Hz, 1st ord. (c) Meas., 500 Hz, 2nd ord. (d) Theo., 500 Hz, 2nd ord.



(e) Meas., 1.0 kHz, 3rd ord. (f) Theo., 1.0 kHz, 3rd ord. (g) Meas., 2.8 kHz, 3rd ord. (h) Theo., 2.8 kHz, 3rd ord.



(i) Meas., 2.9 kHz, 3rd ord. (j) Theo., 2.9 kHz, 3rd ord. (k) Meas., 6.0 kHz, 3rd ord. (l) Theo., 6.0 kHz, 3rd ord.



(m) Meas., 10.0 kHz, 3rd ord. (n) Theo., 10.0 kHz, 3rd ord. (o) Meas., 16.0 kHz, 3rd ord. (p) Theo., 16.0 kHz, 3rd ord.

**Figure 3:** This figure shows, for the indicated frequencies and orders, a comparison of the measured sound field and the ideal (theoretical) sound field. The plots show the relative linear sound field pressure level across two-dimensional space. The lighter shades denote high pressure with white being the maximum pressure represented in each plot, and the darker shades denote a lower pressure with black being the minimum pressure represented in each plot. Note that the size of the area plotted is inversely proportional to the frequency, and thus becomes smaller with increasing frequency. The horizontal and vertical scales are identical and both in metres.

## 3 RESULTS

The sound field recording and reproduction system has been tested at a number of frequencies of interest, and the recorded sound field is compared to an ideal sound field recording and reproduction system which has no error. The incoming wave direction for the tests is at an azimuth and elevation of 0° which is on the horizontal plane directly in front. Figure 3 shows the ideal and recorded sound fields for the sound field recording and reproduction system at the frequencies and orders indicated. The plots show horizontal cross sections of the sound field at an elevation of 0°. It should be noted that the term ideal is used to refer to an ideal or theoretical decoding of a plane-wave sound field at the indicated order and not an ideal plane-wave sound field. The open spherical microphone array is used for recording and reproducing frequencies in the range 60 Hz to 2840 Hz and the rigid spherical microphone array is used for recording and reproducing frequencies in the range 2840 Hz to 16.0 kHz.

## 4 CONCLUSION

We have presented an initial evaluation of a third-order sound field recording and reproduction system, with a frequency range of 60 Hz to 16 kHz, which uses as 64-microphone CRO-SMA for recording and 24-loudspeaker spherical loudspeaker array for reproduction. The initial results look promising and we are in the process of performing further perceptual and analytical tests on the system to more fully characterise its performance.

## REFERENCES

[1] T.D. Abhayapala and D.B. Ward. Theory and design of high order sound field microphones using spherical microphone array. In *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Proc.*, volume 2, pages 1949–1952, 2002.

[2] S. Bertet, J. Daniel, E. Parizet, L. Gros, and O. Warusfel. Investigation of the perceived spatial resolution of higher order ambisonic sound fields: A subjective evaluation involving virtual and real 3d microphones. In *Proc. 30th AES Int. Conf.*, Saariselka, Finland, 2007.

[3] J. Daniel, J. Rault, and J. Polack. Ambisonics encoding of other audio formats for multiple listening conditions. In *105th AES Convention*, San Francisco, CA, USA, 1998.

[4] M.A. Gerzon. Periphony: With-height sound reproduction. *J. Audio Eng. Soc.*, 21(1):2–10, 1973.

[5] R.H. Hardin and N.J.A. Sloane. Mclaren's improved snub cube and other new spherical designs in three dimensions. *Discrete and Computational Geometry*, 15(4):429–441, 1996.

[6] Z. Li, R. Duraiswami, and L.S. Davis. Recording and reproducing high order surround auditory scenes for mixed and augmented reality. In *Proc. Third IEEE and ACM Int. Symposium on Mixed and Augmented Reality*, pages 240–249, 2004.

[7] Z. Li, R. Duraiswami, and N.A. Gumerov. Capture and recreation of higher order 3d sound fields via reciprocity. In *10th Meeting of the International Conference on Auditory Display*, Sydney, Australia, 2004.

[8] A. Parthy, C. Jin, and A. van Schaik. Optimisation of co-centred rigid and open spherical microphone arrays. In *120th AES Conv.*, Paris, France, 2006.

[9] A. Parthy, C. Jin, and A. van Schaik. Measured and theoretical performance comparison of a co-centred rigid and open spherical microphone array. In *Int. Conf. Audio, Language and Image Proc.*, pages 1289–1294, July 2008.

[10] A. Parthy, C. Jin, and A. van Schaik. Acoustic holography with a concentric rigid and open spherical microphone array. In *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Proc.*, pages 2173–2176, 2009.

[11] M.A. Poletti. A unified theory of horizontal holographic sound systems. *J. Audio Eng. Soc.*, 48(12):1155–1182, Dec. 2000.

[12] D.B. Ward and T.D. Abhayapala. Reproduction of a plane-wave sound field using an array of loudspeakers. *IEEE Trans. on Speech and Audio Proc.*, 9(6):697–707, 2001.