3D SOUNDFIELD REPRODUCTION USING HIGHER ORDER LOUDSPEAKERS

P. N. Samarasinghe† M.A. Poletti‡ S. M. A. Salehin† T. D. Abhayapala† F. M. Fazi∗

† Research School of Engineering, CECS, The Australian National University, Australia.
‡ Industrial Research Limited, PO Box 31-310, Lower Hutt, New Zealand.
∗Institute of Sound and Vibration Research, University of Southampton, United Kingdom.

ABSTRACT

Three dimensional surround sound reproduction over large areas is a prevailing challenge due to the enormous numbers of loudspeakers required. In this paper, we propose an array of higher order loudspeakers which provide a mode matching solution to the problem based on 3D wavefield translation. It is shown that for a given bandwidth, the use of $L^3$ order sources significantly brings down the minimum loudspeaker requirement by a factor of $1/(L + 1)^2$. Furthermore, the array is shown to be capable of exterior field cancellation, increasing its performance in echoing environments. Design examples are given for interior field, exterior field and interior and exterior combined field reproduction.

1. INTRODUCTION

Spatial sound reproduction systems are designed to reproduce a desired soundfield over a specified volume of space. There are two main approaches for this purpose based on distinct representations of the soundfield. First approach, the wave field synthesis method [1–3] is based on the Kirchoff-Helmholtz (K-H) integral representation of a soundfield. This method requires a continuous array of monopole and radial dipole sources on the surface of the volume of interest.

The second approach which is also used in this paper, is the higher order Ambisonics approach [4–9] based on the spherical harmonic decomposition of a soundfield.

In higher order ambisonics the minimum number of harmonics required to accurately describe a soundfield is proportional to the frequency and the radius of the concerned region. Hence, large numbers of loudspeakers are required for reproduction over greater areas. For example, at 600 Hz, 3D reproduction over a radius of 2 m requires an impractical number of at least 760 loudspeakers [10].

A second limitation of sound reproduction systems is reverberation caused by room reflections which distort the reproduced soundfield. Reverberation can only be minimized by computationally expensive pre-processing [11] or by exterior field cancellation. The latter is theoretically proven to be possible based on the K-H approach [1] but its practical application is limited due to the continuous aperture requirement.

In addressing the above mentioned issues we propose an array of 3D higher order loudspeakers where each source is capable of producing polar responses up to the $N^3$th order. We derive the reproduced soundfield using wavefield translation based on the 3D addition theorem. The reproduced field coefficients are then matched with the desired field coefficients to determine the mode amplitudes of each directional source.

1.1. Relation to prior work

A related solution was first discussed in [12] where an array of 2D general first order loudspeakers were used for height invariant interior reproduction with exterior cancellation. It extended the K-H solution for exterior field cancellation by adding an extra tangential dipole to each of the existing monopole and radial dipole sources increasing the useful bandwidth by approximately an octave. The same approach was later extended with two dimensional $N^3$ order sources in [13, 14] reducing the minimum loudspeaker requirement by a factor of $1/(2N + 1)$. Since height invariant reproduction is less applicable in practical scenarios, we consider 3D reproduction employing more realizable three dimensional higher order loudspeakers. These loudspeakers are arrays of transducers independently driven to produce variable directivity patterns up to order $N$ [15–17].

2. THEORY

2.1. Harmonic representation of a soundfield

We first describe the desired soundfield in terms of a spherical harmonic distribution. Consider an interior soundfield $\eta$, caused by one or more sources lying outside the region of concern. The wave equation can be solved in terms of spherical harmonics to represent the sound pressure at any point $x \equiv (r, \theta, \phi)$ within $\eta$ as [18]

$$P(x, k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \alpha_{nm}(k) j_n(kr) Y_{nm}(\theta, \phi)$$

(1)

where $k = 2\pi f/c$ is the wave number, $f$ is the frequency, $c$ is the speed of sound propagation, $j_n(\cdot)$ represents the spherical Bessel function of order $n$, $\alpha_{nm}(k)$ are the interior sound field coefficients, and $Y_{nm}(\theta, \phi)$ denotes the spherical harmonic of order $n$ and degree $m$. In the case of an exterior field $\zeta$, where all the sources are confined within the region of concern, the sound pressure at any point $x$ outside from the outermost source can be given by [18]

$$P(x, k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \beta_{nm}(k) h_n(kr) Y_{nm}(\theta, \phi)$$

(2)

where $h_n(\cdot)$ represents the $n^{th}$ order spherical Hankel function of the first kind and $\beta_{nm}(k)$ are the exterior sound field coefficients. Since the soundfield coefficients in (1) and (2) are both independent of spatial variations, for a given frequency $k$, the desired soundfield can be accurately reproduced if $\alpha_{nm}(k)$ and $\beta_{nm}(k)$ are successfully recreated.

∗This work was supported by the Australian Research Council Discovery Grant DP 110103369.
2.2. Higher order loudspeakers

In this section, we first propose a realizable directional source design and later discuss an array of such units. An ideal 3D higher order loudspeaker placed at the origin will produce a soundfield similar to (2)

$$P(x, k) = \sum_{n=0}^{\hat{N}} \sum_{\hat{n}=0}^{\hat{n}} \hat{\beta}_{\hat{n}n}(k) h_{\hat{n}}(kr) Y_{\hat{n}n}(\theta, \phi)$$

(3)

where the loudspeaker order \( \hat{N} \) denotes the maximum directivity achievable from that particular transducer. This idealization becomes less practical when \( kr \) tends to zero due to Hankel functions. Thus, we introduce a more realizable higher order loudspeaker derived from a spherical source of radius \( r' \) vibrating with surface velocity

$$V(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{\hat{n}=0}^{\hat{n}} \chi_{\hat{n}n}(k) Y_{\hat{n}n}(\theta, \phi).$$

(4)

From (3), the radial velocity is

$$V(\theta, \phi) = -\frac{i}{pc} \sum_{n=0}^{\hat{N}} \sum_{\hat{n}=0}^{\hat{n}} \hat{\beta}_{\hat{n}n}(k) h_{\hat{n}}(kr) Y_{\hat{n}n}(\theta, \phi), \quad (r \geq r').$$

At the surface of the sphere \( r = r' \), this must equal the surface velocity (4). Hence,

$$\hat{\beta}_{\hat{n}n} = i \frac{pc}{k} \chi_{\hat{n}n}(k) h_{\hat{n}}(kr').$$

If the velocity is produced by a moving coil loudspeaker driver, then the velocity above the resonant frequency varies as \( 1/k \) and the soundfield for frequencies above the resonant frequency can be represented as

$$P(x, k) = \frac{i \frac{pc}{k}}{2} \sum_{n=0}^{\hat{N}} \sum_{\hat{n}=0}^{\hat{n}} \chi_{\hat{n}n}(k) h_{\hat{n}}(kr') Y_{\hat{n}n}(\theta, \phi), \quad (r \geq r')$$

which produces a flat response for the zeroth order mode at low frequencies. We can simplify the above response for a general \( \hat{N} \) order source as

$$P(x, k) = e^{-iwr} \sum_{n=0}^{\hat{N}} \sum_{\hat{n}=0}^{\hat{n}} w_{\hat{n}n}(k) h_{\hat{n}}(kr') Y_{\hat{n}n}(\theta, \phi), \quad (r \geq r')$$

where \( w_{\hat{n}n}(k) \) denote the modal weights. We employ an array of such higher order loudspeakers with the \( q^\alpha(q = 1 \cdots Q) \) loudspeaker located at \( x_q \equiv (r_q, \theta_q, \phi_q) \). The total array response can be expressed relative to the origin using wavefield translation based on the spherical addition theorem for Hankel functions [19, 20] as

$$P(x, k) = \left\{ \begin{array}{ll}
\sum_{n=m-n}^{\infty} \sum_{q=1}^{Q} \sum_{\hat{n}=0}^{\hat{n}} \sum_{\hat{n}=0}^{\hat{n}} \frac{w_{\hat{n}n}(q) h_{\hat{n}}(kr)}{h_{\hat{n}}(kr')} Y_{\hat{n}n}(\theta, \phi) & r \leq r_q \\
\cdot \cdot \cdot S_{\hat{n}n}^{m^*}(x_q)j_n(kr) Y_{n}(\theta, \phi) & r_q \leq r_q \\
\end{array} \right. \quad (5)

$$

where

$$S_{\hat{n}n}^{m^*}(x_q) = i^{(n-\hat{n}+4n-2m)} \sum_{l=0}^{n+\hat{n}} j_l(kr_q) Y_{l-\hat{n}}^{*}(m-n) \left( \theta_q, \phi_q \right) \
\times \sqrt{4\pi(2n+1)(2\hat{n}+1)(2l+1)} W_1 W_2,$$

with

$$W_1 = \begin{pmatrix} \hat{n} & n & l \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad W_2 = \begin{pmatrix} \hat{n} & n & l \\ m & m & (m-\hat{n}) \end{pmatrix}$$

representing Wigner 3 – j symbols, and

$$\hat{S}_{\hat{n}n}^{m^*}(x_q) = i^{(n-\hat{n}+4n-2m)} \sum_{l=0}^{n+\hat{n}} j_l(kr_q) Y_{l-\hat{n}}^{*}(m-n) \left( \theta_q, \phi_q \right) \
\times \sqrt{4\pi(2n+1)(2\hat{n}+1)(2l+1)} W_1 W_2.$$

In practice, the infinite summations in (5) can be replaced with finite sums, \( n \in [-N, N] \), due to the high pass nature of Bessel functions [5]. For computational simplicity we assume a spherical array of radius \( R_q \) with \( r_q = R_0 \) for all \( q \). Since \( j_n(kr) \) is negligible for \( n \geq kr \), for interior reproduction \( (r \leq R_0) \), \( N_l = [kR_0] \), where \([\cdot]\) denotes rounding up to the closest upper integer. Similarly, for exterior reproduction \( (r \geq R_0) \), \( r_n + \lambda(kr_q) \) is negligible for \( n \geq (\hat{n} + kR_0) \) yielding \( N_E = \hat{N} + [kR_0] \).

3. SOUNDFIELD REPRODUCTION

3.1. Interior field reproduction

For interior field reproduction with no exterior control, we require the higher order array response to recreate the desired interior field coefficients \( a_{nm}(k) \). Comparison of (1) and (5) for \( r \leq r_q \) yields

$$\sum_{n=0}^{Q} \sum_{q=1}^{Q} \sum_{\hat{n}=0}^{\hat{n}} \sum_{\hat{n}=0}^{\hat{n}} w_{\hat{n}n}(q) h_{\hat{n}}(kr) S_{\hat{n}n}^{m^*}(x_q) = a_{nm}(k), n \in [-N_l, N_l].$$

(6)

For all \( n \), this forms a a set of linear equations

$$Hw = \alpha$$

(7)

where \( H \) is an \( (N_l + 1)^2 \) by \( Q(N_l + 1)^2 \) translation matrix, \( w \) is a \( Q(N_l + 1)^2 \) by one vector of loudspeaker weights and \( \alpha \) is a \( (N_l + 1)^2 \) by one vector of desired field coefficients. Since \( H \) and \( \alpha \) are known, the required modal weights of higher order sources can be solved using

$$w = H^\dagger \alpha$$

(8)

where \( H^\dagger \) denotes the pseudoinverse. To avoid spatial aliasing \( Q(N_l + 1)^2 \geq (N_l + 1)^2 \) has to be satisfied with (7) then yielding the minimum energy weight solution. This condition also implies that for a given frequency \( k \), the minimum requirement of \( Q \) will decrease with increasing \( \hat{N} \) by a factor \( 1/(\hat{N} + 1)^2 \). The maximum soundfield order that can be controlled by an array of \( Q \) number of \( \hat{N} \)th order loudspeakers is \( N_l = \sqrt{Q(\hat{N} + 1)} - 1 \). Since \( N_l = [kR_0] \), the array’s maximum achievable frequency is limited to

$$f_{\text{max}l} = \left( \frac{\sqrt{Q(\hat{N} + 1)} - 1}{2\pi R_0} \right)$$

(9)

where \( \gamma \) is an empirical oversampling factor enabling accurate reproduction in the vicinity of \( f_{\text{max}l} \).

3.2. Exterior field reproduction

For exterior field reproduction with no interior control, the desired field coefficients \( \beta_{nm}(k) \) in (2) needs to be matched with the source array response in (5) for \( r \geq r_q \). Hence,

$$\sum_{n=0}^{Q} \sum_{q=1}^{Q} \sum_{\hat{n}=0}^{\hat{n}} \sum_{\hat{n}=0}^{\hat{n}} w_{\hat{n}n}(q) h_{\hat{n}}(kr) S_{\hat{n}n}^{m^*}(x_q) = \beta_{nm}(k), n \in [-N_E, N_E].$$
For all \( n \), this forms a set of linear equations

\[
Jw = \beta
\]

(9)

where \( J \) is an \((N_E + 1)^2\) by \( Q(N + 1)^2\) translation matrix and \( \beta \) is an \((N_E + 1)^2\) by one vector of desired field coefficients. An exact solution to the system exists similar to (7) where \( Q(N + 1)^2 \geq (N_E + 1)^2 \) still holds to avoid spatial aliasing. The maximum soundfield order that can be controlled by an array of \( Q \) number of \( N_m^\text{th} \) order loudspeakers is \( N_E = \sqrt{Q(N + 1)} - 1 \). Since \( N_E = N + [kR_0] \), the array’s maximum achievable frequency is limited to

\[
f_{\text{max(IE)}} = (N + 1)(\sqrt{Q} - 1) - 1 \frac{c}{2\pi R_0}.
\]

(10)

### 3.3. Interior/Exterior field reproduction

For interior/exterior combined reproduction, we require the higher order source array response to equal both \( \alpha_{nm}(k) \) and \( \beta_{nm}(k) \) reproducing two separate desired fields inside and outside the spherical source array respectively. Combining (6) and (9) we derive

\[
\begin{bmatrix}
H
\end{bmatrix}w = \Psi w = \begin{bmatrix}
\alpha \\
\beta
\end{bmatrix}
\]

(11)

where \( \Psi \) is an \((N_I^2 + 1)^2\) by \((N + 1)^2\) translation matrix. Exterior cancellation (i.e. \( \beta = [0 \cdots 0]^T \)) is a special case of interior/exterior reproduction providing an effective solution in reverberant environments. When \( Q(N + 1)^2 \geq (N_I^2 + 1)^2 + (N_E + 1)^2 \) is satisfied to prevent spatial aliasing the system can be solved using (7) yielding a minimum energy weight solution. For an array of \( Q \) number of \( N_m^\text{th} \) order loudspeakers, the maximum number of soundfield modes are controlled when \((N_E + 1)^2 + (N_I + 1)^2 = Q(N + 1)^2\). Hence, the array’s maximum achievable frequency is limited to

\[
f_{\text{max(IE)}} = (N + 1)(\sqrt{2Q} - 1) - 1 \frac{c}{4\pi R_0}.
\]

(12)

### 3.4. Reproduction error

The reproduction accuracy of the proposed design may be quantified in terms of angle-averaged relative error [10] as

\[
\epsilon = \frac{\int_0^{2\pi} \int_0^\pi \left( P(x, k) - \tilde{P}(x, k) \right)^2 \sin \theta \, d\theta \, d\phi}{\int_0^{2\pi} \int_0^\pi |P(x, k)|^2 \sin \theta \, d\theta \, d\phi}.
\]

(13)

For interior field reproduction, substituting from (1) and (5) in \( P \) and \( \tilde{P} \), and using the orthonormal property of spherical harmonics, the above expression can be simplified to

\[
\epsilon_I = \frac{N_I}{\int_0^{2\pi} \int_0^\pi |P(x, k)|^2 \sin \theta \, d\theta \, d\phi} \left( \sum_{n=0}^{N_I} \sum_{m=-n}^{n} |\alpha_{nm}(k) - \alpha'_{nm}(k)|^2 j_n(kr)^2 \right)
\]

(14)

where

\[
\alpha'_{nm}(k) = \sum_{q=1}^Q \sum_{n=0}^N \sum_{m=-n}^{n} \frac{w_{nm}(q)}{H_n(kr)} S_{nm}^q(x_q)
\]

(15)

are the soundfield coefficients reproduced after solving for \( w \) in (6). Similarly for exterior fields, substituting for (13) from (2) and (5) yields a reproduction error of

\[
\epsilon_E = \frac{\sum_{n=0}^{N_E} \sum_{m=-n}^{n} |\beta_{nm}(k) - \beta'_{nm}(k)|^2 j_n(kr)^2}{\int_0^{2\pi} \int_0^\pi |P(x, k)|^2 \sin \theta \, d\theta \, d\phi}
\]

(16)

\[
\beta'_{nm}(k) = \sum_{q=1}^Q \sum_{n=0}^N \sum_{m=-n}^{n} \frac{w_{nm}(q)}{H_n(kr)} S_{nm}^q(x_q).
\]

(17)

### 4. SIMULATIONS

To demonstrate the applicability of the theory developed, we performed Matlab simulations for interior, exterior and interior/exterior reproduction using a fixed \( Q \). The expected performance increase with increasing source order was analyzed by varying \( N \) from 0 to 3. The sources were assumed to be equally distributed on a spherical array of radius \( R_0 = 2 \text{ m} \). Since a mathematical explanation for

![Fig. 1. Interior field reproduction: Desired soundfield at \( f = 500 \text{ Hz} \) and the reproduced soundfield by an array of 49 third order loudspeakers at elevations (a) \( z = 0 \text{ m} \) and (b) \( z = 1 \text{ m} \).](image)

![Fig. 2. Interior field: Array reproduction error for source orders \( N = 0, 1, 2 \) and 3.](image)
an equal division of a spherical surface does not exist we positioned the loudspeakers at locations determined by [21] which confines $Q$ to $g^3$ where $g$ is an integer and in our examples we fix $Q$ at 49. An oversampling factor of $\gamma = 0.7$ was used to ensure increased performance above $f_{\text{max}}$. For simplicity, reproduction plots were confined to horizontal planes at different elevations of the 3D soundfield.

For interior only reproduction we assumed a desired field due to a point source at $x_s = (3, \pi/3, \pi/4)$ and expected recreation over the entire spherical region within the source array. Figure 1 shows the desired and reproduced soundfields at $f = 500$ Hz for an array of 49 third order loudspeakers with $f_{\text{max}(t)} = 610$ Hz. Successful reproduction is observed within the source array beyond which the field is not controlled. To reach the same frequency, we require at least 760 zeroth order loudspeakers indicating the significant reduction of $Q$ with increasing $N$. A more general display of the array performance is given in figure 2 where the reproduction error (14) was plotted against frequency for all $N$. According to (8), their respective $f_{\text{max}(t)}$ were 132, 291, 451 and 610 Hz demonstrating a noticeable increase in the array’s useful bandwidth with increasing source order. The figure shows minimal error up to $f_{\text{max}(t)}$ beyond which spatial aliasing error dominates the array performance. When solving (6) the desired soundfield order $N_I$ was varied with $f$ which may have caused the step increase in error beyond $f_{\text{max}(t)}$.

![Fig. 3. Exterior field reproduction: Desired soundfield at $f = 500$ Hz and the reproduced soundfield for an array of 49 third order loudspeakers at elevations (a) $z = -1.5$ m and (b) $z = 1.5$ m.](image)

![Fig. 4. Exterior field: Array reproduction error for source orders $N = 0, 1, 2$ and 3.](image)

![Fig. 5. Interior field reproduction with exterior cancellation: Desired soundfield at $f = 200$ Hz and the reproduced soundfield by an array of 49 third order loudspeakers at elevations (a) $z = 0$ m and (b) $z = 1$ m.](image)

For exterior only reproduction, we assumed a desired field due to 16 point sources equally distributed on a spherical array of radius 1 m. Reproduction was expected over the entire region outside the higher order loudspeaker array. Figure 3 shows an example of desired and successfully reproduced soundfields at $f = 500$ Hz for an array of 49 third order loudspeakers with $f_{\text{max}(t)} = 528$ Hz. A more general representation of the array’s performance is given in figure 4 where reproduction error (16) is plotted against frequency for all $N$. According to (12) their respective $f_{\text{max}(t)}$ are 132, 264, 396 and 528 Hz, up to which error is observed to be minimal.

For interior/exterior reproduction we used the same desired field described in interior reproduction combined with exterior field cancellation to avoid reverberation. Figure 5 shows a reproduction example at $f = 200$ Hz for an array of 49 third order loudspeakers with $f_{\text{max}(IE)} = 280$ Hz where the desired interior field is successfully recreated within the source array while canceling the exterior field. The additional requirement of controlling $(N_E + 1)^2$ exterior modes is followed by a decrease in $f_{\text{max}(IE)}$ from $f_{\text{max}(t)}$.

5. CONCLUSION

We have theoretically formulated a 3D surround sound system capable of interior field, exterior field and interior/exterior combined field reproduction. From simulations we have showed that the use of loudspeakers with up to $L^6$ order directivity allows reproduction over approximately $L$ times the bandwidth while producing a significantly attenuated exterior soundfield. It also suggests that the use of higher order sources results in a significant reduction in the minimum loudspeaker requirement at the expense of more sophisticated source units. Therefore, implementation of such arrays in commercial surround systems is more feasible compared to large monopole loudspeaker arrays.
6. REFERENCES


