ANGULAR SPECTRUM DECOMPOSITION-BASED 2.5D HIGHER-ORDER SPHERICAL HARMONIC SOUND FIELD SYNTHESIS WITH A LINEAR LOUDSPEAKER ARRAY

Takuma Okamoto

National Institute of Information and Communications Technology
3-5, Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-0289, Japan
okamoto@nict.go.jp

ABSTRACT

This paper derives an analytical driving function to synthesize an exterior sound field described by spherical harmonic expansion coefficients using a linear loudspeaker array. An exterior sound field is decomposed into both spherical and 3D cylindrical harmonic expansion coefficients. The spherical harmonic expansion coefficients are analytically converted into 3D cylindrical ones by plane wave decomposition. The angular spectrum coefficients at the synthesis reference line are then analytically obtained from the converted 3D cylindrical harmonic expansion coefficients and directly synthesized by the spectral division method with a linear loudspeaker array. The results of computer simulations indicate the effectiveness of the proposed analytical formulation.

Index Terms—2.5D sound field synthesis, linear loudspeaker array, plane wave decomposition, spherical harmonics, spectral division method

1. INTRODUCTION

Higher-order Ambisonics (HOA) [1–4] have been investigated for a sound field synthesis technique that can synthesize sound waves coming from all directions. In HOA, the surrounding spherical and circular loudspeaker arrays are used for synthesis. A sound field in HOA is described by spherical and circular harmonic expansion coefficients [5]. In addition, an exterior sound field with a higher-order radiation pattern, such as instrumental sound propagation, is also represented by these coefficients and synthesized by a spherical loudspeaker array mounted on a spherical rigid baffle [6].

For synthesizing sound waves from a half space, wave field synthesis (WFS) [7–9] and a spectral division method (SDM) [4,10–14] have been proposed. In WFS and SDM, the planar and linear loudspeaker arrays are introduced. A sound field in SDM is described by angular spectrum coefficients [5].

Since these methods provide analytical solutions based on the spatial Fourier transform, sound field representations must have identical formats between the recording/modeling and synthesis stages. In contrast, there is no representation consistency between these stages in pressure matching-based [15–20] and plane wave decomposition-based [21] least squares approaches. These numerical approaches, however, are quite unstable because the acoustic inverse problem is very ill-conditioned [5,13].

For these reasons, the flexibility of analytical approaches to sound field synthesis must be investigated. In addition, such systems are frequently simplified for synthesis in the horizontal plane for practical implementations. The sound sources are then arranged on a line or a circle. In actual implementations, 3D monopole sources instead of 2D line sources are usually employed as sound sources. Such approaches are called 2.5D sound field synthesis [3,8–12,14,19,22–25].

An analytical approach to 2.5D WFS with truncated linear loudspeaker arrays has been provided [22] for synthesizing interior sound fields described by spherical harmonic expansion coefficients. An analytical method of 2.5D HOA with a circular loudspeaker array for an interior sound field described by angular spectrum coefficients has also been investigated [23].

This paper focuses on the 2.5D synthesis of an exterior sound field with a higher-order radiation pattern, described by spherical harmonic expansion coefficients with a linear loudspeaker array. Some approaches have already been investigated [26–29]. This technique efficiently provides an exterior sound field in a wider area on the horizontal plane to multiple listeners and is suited to be combined with big screen video display systems [13]. The existing methods, however, introduced far field approximation and cannot effectively control the sound field near the loudspeakers.

One possible solution to this problem is extending conventional 2.5D WFS for interior field synthesis [22] into that for exterior field synthesis. In a typical 2.5D synthesis with a linear array, the synthesis reference line must be set where precise sound pressures are synthesized. In an interior field synthesis, the reference line is set on the recording/modeling center [22]. However, to synthesize an exterior field with a linear array, the reference line is not on the recording/modeling center, and the sound pressures on the reference line must be obtained from the spherical harmonic expansion coefficients. It is therefore difficult to directly apply the conventional 2.5D WFS-based method to exterior sound field synthesis.

This paper provides an alternative analytical method for the 2.5D synthesis of an exterior sound field with a linear loudspeaker array and analytically converts the spherical harmonic expansion coefficients into 3D cylindrical ones by plane wave decomposition. The angular spectrum coefficients at the synthesis reference line are then analytically obtained from the converted 3D cylindrical harmonic expansion coefficients and directly synthesized by 2.5D SDM with a linear loudspeaker array.

2. PROPOSED FORMULATION

2.1. 3D exterior sound field described by spherical harmonic expansion coefficients

In the proposed formulation, spherical coordinates relative to Cartesian coordinates for both the spatial and wavenumber k(=ω/c) domains are respectively defined in Figs. 1(a) and (b). ω = 2πf is the
The exterior expansion of a 3D sound field for regions exterior to any sound sources is given as

$$ S(r, \theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \hat{S}_n^m h_n(kr) Y_n^m(\theta, \phi), \quad (1) $$

where $\hat{S}_n^m$ and $h_n$ are the exterior sound field expansion coefficients and the $n$-th order spherical Hankel function of the first kind, and

$$ Y_n^m(\theta, \phi) = \frac{(2n+1)(n-|m|)!}{4\pi(n+|m|)!} P^m_n(\cos \theta) e^{jm\phi} $$

is the spherical harmonics, and $Y_n^m$ is the associated Legendre function [30]. The exterior field expansion coefficients are obtained from the sound pressures on a sphere with radius $r$ [5] and given as

$$ \hat{S}_n^m = \frac{1}{h_n(kr)} \int_0^{2\pi} \int_0^\pi S(r, \theta, \phi) Y_n^m(\theta, \phi)^* \sin \theta d\theta d\phi. \quad (3) $$

In data-based synthesis, $\hat{S}_n^m$ are obtained from the discretized representation of (3) where $S(r, \theta, \phi)$ are recorded by a surrounding spherical microphone array [31] (Fig. 2(a)). In contrast, $\hat{S}_n^m$ are directly given in model-based synthesis (Fig. 2(b)).

In typical spherical harmonic expansion coefficients, spherical coordinates relative to Cartesian coordinates are defined as $[x, y, z]^T = [r \sin \theta \cos \phi, r \sin \theta \sin \phi, r \cos \theta]^T$ [5]. In this case, spherical harmonic expansion rotation techniques [32] are introduced. Rotation schemes are also useful for synthesizing the arbitrary directions of the exterior sound field (Fig. 2(b)).

### 2.2. Plane wave decomposition of exterior sound field described by spherical harmonic expansion coefficients

As in the conventional interior field approach [22], a sound field is decomposed into plane waves and described by the 3D Herglotz integral [21, 30, 33]:

$$ S(r, \theta, \phi) = \frac{1}{4\pi} \int_0^{2\pi} \int_0^\pi \overline{S}(\theta_0, \phi_0) e^{ik^T \cdot \mathbf{x}} \sin \theta_0 d\theta_0 d\phi_0, \quad (4) $$

where $\overline{S}(\theta_0, \phi_0)$ is the Herglotz weight coefficient of a plane wave propagating from $[\theta_0, \phi_0]^T$, and

$$ \mathbf{r} = [k_x, k_y, k_z]^T = [k \cos \theta_0, k \sin \theta_0 \cos \phi_0, k \sin \theta_0 \sin \phi_0]^T, \quad \mathbf{x} = [x, y, z]^T = [r \cos \theta, r \sin \theta \cos \phi, r \sin \theta \sin \phi]^T. $$

A plane wave is represented by the spherical harmonic expansion [5] and given as

$$ e^{ik^T \cdot \mathbf{x}} = \sum_{j=0}^{\infty} \sum_{m=-j}^{j} j^n h_n(kr) Y_n^m(\theta, \phi) e^{jm\phi}, \quad (5) $$

where $j = \sqrt{-1}$. From (1), (4), and (5),

$$ \hat{S}_n^m = \int_0^{2\pi} \int_0^\pi \overline{S}(\theta_0, \phi_0) Y_n^m(\theta_0, \phi_0)^* \sin \theta_0 d\theta_0 d\phi_0, $$

and the inverse is obtained as

$$ \overline{S}(\theta_0, \phi_0) = \sum_{m=-\infty}^{\infty} \sum_{n=|m|}^{\infty} j^{-n} \hat{S}_n^m Y_n^m(\theta_0, \phi_0) e^{jm\phi_0}. \quad (7) $$

### 2.3. Plane wave decomposition of exterior sound field described by 3D cylindrical harmonic expansion coefficients

An exterior sound field is also decomposed into 3D cylindrical harmonic expansion coefficients $\hat{S}_m(k_x)$ [5] and represented as

$$ S(R, \phi, x) = \sum_{m=-\infty}^{\infty} e^{jm\phi} \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{S}_m(k_x) H_m(k_x R) e^{ik_x x} dx, $$

where $R = \sqrt{x^2 + z^2}$, $k_x = \sqrt{k^2 - k_z^2}$, and $H_m$ is the $m$-th order Hankel function of the first kind [5]. A plane wave is also described by 3D cylindrical harmonic expansion [34] and represented as

$$ e^{ik^T \cdot \mathbf{x}} = \sum_{m=-\infty}^{\infty} e^{jm(\phi - \phi_0)} \int_{-\infty}^{\infty} H_m(k_x R) \delta(k_x - k \cos \theta_0) e^{ik_x x} dx, \quad (9) $$

where $\delta$ is a Dirac delta function [5]. From (4), (8), and (9),

$$ \hat{S}_m(k_x) = \frac{j^m}{2} \int_0^{2\pi} \int_0^\pi \overline{S}(\theta_0, \phi_0) e^{-jm\phi_0} \delta(k_x - k \cos \theta_0) \sin \theta_0 d\theta_0 d\phi_0. \quad (10) $$

### 2.4. Analytically converting sound field representation from spherical to 3D cylindrical harmonic expansion

From (7) and (10), spherical harmonic expansion coefficients $\hat{S}_n^m$ are analytically converted into 3D cylindrical ones $\hat{S}_m(k_x)$ (Fig. 2(b)) and simply represented as

$$ \hat{S}_m(k_x) = \frac{j^m \pi}{k} \sum_{n=|m|}^{\infty} j^{-n} \hat{S}_n^m Y_n^m(k_x), $$

where $\hat{S}_n^m Y_n^m(k_x)$ is the $m$-th order spherical Hankel function of the first kind, and

$$ j^{m-n} h_n(kr) Y_n^m(\theta, \phi) e^{jm\phi}. $$
where \( \sin \theta_0 = k_x / k \), the integral by substitution \( d\theta_0 \to dk_x \) and

\[
\frac{d}{dk_x} \left\{ \cos^{-1} \left( \frac{k_x}{k} \right) \right\} = \frac{-1}{\sqrt{k^2 - k_x^2}} = -\frac{1}{k_x},
\]

\[
\frac{1}{2\pi} \int_0^{2\pi} \tilde{S}(\theta_0, \phi_0) e^{-jm\phi_0} d\phi_0 = \sum_{n=|m|}^\infty j^{-n} \tilde{S}_m Y_n^m(k_x).
\]

are applied, and

\[
Y_n^m(k_x) = \sqrt{\frac{(2n+1)(n-|m|)!}{4\pi(n+|m|)!}} P_n^{|m|} \left( \frac{k_x}{k} \right).
\]

As illustrated in Fig. 2(c), the exterior sound field on a cylinder along the \( x \)-axis with radius \( R \) and angle \( \phi \) is analytically obtained by (8) where only the Fourier series expansion component is calculated and represented as

\[
\tilde{S}(R, \phi, k_x) = \sum_{m=-\infty}^{\infty} \tilde{S}_m(k_x) H_m(k_x, R) e^{jm\phi},
\]

which is obtained as the angular spectrum coefficients for \( k_x \).

### 2.5. Analytical driving function of a linear sound source

Sound pressure \( S(x) \), which is synthesized at position \( x \) by a continuous linear sound source with an infinite length along the \( x \)-axis, is given as

\[
S(x) = \int_{-\infty}^{\infty} D(x_0) G_{3D}(x, x_0) dx_0,
\]

where \( D(x_0) \) is the sound source driving function at position \( x_0 = [x_0, y_0, 0]^T \), and \( G_{3D}(x, x_0) \) is the transfer function of the sound source placed at \( x_0 \) to point \( x \). Under the free-field assumption,

\[
G_{3D}(x, x_0) = \frac{e^{j k |x - x_0|}}{4\pi |x - x_0|}
\]

is the 3D free-field Green’s function [5]. In the SDM [10], the spatial Fourier transform of (16) along the \( x \)-axis is applied and given as

\[
\tilde{S}(y_{syn}, 0, k_x) = \hat{D}(k_x) \tilde{G}_{3D}(k_x, y, y_0)
\]

\[
= \hat{D}(k_x) \frac{1}{4} H_0(k_x (y_{syn} - y_0)),
\]

where the synthesis reference line is set to \( y = y_{syn} > 0 \), \( z = 0 \) and \( y_0 < y_{syn} \) (Fig. 2(d)). From (15) and (18), the driving function for synthesizing exterior sound field \( \tilde{S}(y_{syn}, 0, k_x) \) in the angular spectrum domain is then analytically derived as

\[
\hat{D}(k_x) = \frac{-4j}{k H_0(k_x (y_{syn} - y_0))} \sum_{m=-\infty}^{\infty} j^m H_m(k_x y_{syn}) \sum_{n=|m|}^\infty j^{-n} \tilde{S}_m Y_n^m(k_x),
\]

where \( y_{syn} = y_0 \). The driving function of the proposed method in the temporal frequency domain is finally derived by the inverse spatial Fourier transform:

\[
D(x_0) = \int_{-\infty}^{\infty} \hat{D}(k_x) e^{j k_x x_0} dk_x,
\]

where only the propagation wave components are considered and evanescent components \( |k_x| > |k| \) are discarded to calculate stable driving functions [13, 14].

For actual implementations, a linear loudspeaker array is used instead of a continuous linear sound source, and (20) must be discretized and truncated.

### 2.6. Comparison with conventional virtual point source synthesis by SDM

The driving function for virtual point source synthesis, which is located at \( x_s = [x_s, y_s, 0]^T \) by the SDM [11, 25], is given as

\[
\hat{D}(k_x)_{ps} = \hat{P}_{ps} e^{-jk_x x_s} \frac{H_0(k_x (y_{syn} - y_s))}{H_0(k_x y_{lodm})},
\]

where \( \hat{P}_{ps} \) is the temporal spectrum of the virtual point source. When \( \hat{P}_{ps} = -j \sqrt{\pi} \tilde{S}_0^2 / k \) and \( x_s = y_0 = 0 \), \( D(k_x)_{ps} \) is equivalent to proposed driving function \( \hat{D}(k_x) \) in (19) for the 0-th order.
The results of computer simulations showed the effectiveness of the proposed analytical driving function. Averaged synthesis error [dB]

\[
\begin{array}{cccccc}
\text{Temporal frequency [kHz]} & 0 & 0.5 & 1 & 1.5 & 2 & 2.5 & 3 & 3.5 \\
\text{Averaged synthesis error [dB]} & -30 & -25 & -20 & -15 & -10 & -5 & 0 & 5 \\
\end{array}
\]

Computer simulations evaluated the proposed analytical approach and compared it with the conventional numerical SDM [12]. In all the simulations, a three-dimensional free field was assumed. The speed of sound \( c \) was 343.36 m/s.

The original exterior sound field was modeled by \( \hat{S}_{n}^{m} \) up to \( n = 8 \)-th order with 81 random numbers and calculated by (1). A linear array of loudspeakers was set on \( y_0 = 0.5 \) or \(-0.5 \) m and centered around \( x = 0 \). The number of loudspeakers was \( L = 64 \), and the distance between adjacent loudspeakers was \( \Delta x = 0.05 \) m. Its spatial Nyquist frequency was then about 3.4 kHz. The synthesis reference line was set to \( y_{0y} = 2.0 \) m. \( dk_x \) in (20) was discretized into \( \Delta k_x = 2\pi/4L\Delta x = 0.4909 \).

In the numerical SDM, the sound pressures on the synthesis reference line were obtained by (1), and \( \hat{S}(y_{0y}, 0, k_x) \) were numerically calculated by the discrete spatial Fourier transform where \( \Delta x = 0.05 \) m and \(-6.4 \leq x \leq 6.35 \) m.

To estimate the synthesized sound field, the synthesis error at position \( x \) was defined as

\[
E(x) = 10 \log_{10} \left| \frac{S_{\text{org}}(x) - S_{\text{syn}}(x)}{|S_{\text{org}}(x)|^2} \right|^2,
\]

where \( S_{\text{org}}(x) \) and \( S_{\text{syn}}(x) \) are respectively the original and synthesized sound pressures at position \( x \).

Figure 3 shows the results of the original exterior sound field, the synthesized sound field, and the synthesis error in the horizontal plane for the proposed method at a temporal frequency of \( f = 2 \) kHz with \( y_0 = 0.5 \) and \(-0.5 \) m, respectively. The results of the averaged synthesis error calculated on a plane with \( z = 0 \) m, \(-1.5 \leq x \leq 1.5 \) m, and \( 1.0 \leq y \leq 3.0 \) m for both the numerical and proposed methods up to the spatial Nyquist frequency were plotted in Fig. 4. Those of the virtual point source synthesis discussed in Section 2.6 were also plotted as references. These results indicated that the proposed method achieved 2.5D exterior sound field synthesis where the original higher-order sound source is located both in front of and behind the linear loudspeaker array with identical accuracy as the numerical approach.

Consequently, based on angular spectrum decomposition, the proposed analytical driving function can effectively synthesize an exterior sound field described by the spherical harmonic expansion coefficients using a linear loudspeaker array.

3. COMPUTER SIMULATIONS

Computer simulations evaluated the proposed analytical approach and compared it with the conventional numerical SDM [12]. In all the simulations, a three-dimensional free field was assumed. The speed of sound \( c \) was 343.36 m/s.

The original exterior sound field was modeled by \( \hat{S}_{n}^{m} \) up to the \( n = 8 \)-th order with \( (n+1)^2 = 81 \) random numbers and calculated by (1). A linear array of loudspeakers was set on \( y_0 = 0.5 \) or \(-0.5 \) m and centered around \( x = 0 \). The number of loudspeakers was \( L = 64 \), and the distance between adjacent loudspeakers was \( \Delta x = 0.05 \) m. Its spatial Nyquist frequency was then about 3.4 kHz. The synthesis reference line was set to \( y_{0y} = 2.0 \) m. \( dk_x \) in (20) was discretized into \( \Delta k_x = 2\pi/4L\Delta x = 0.4909 \).

In the numerical SDM, the sound pressures on the synthesis reference line were obtained by (1), and \( \hat{S}(y_{0y}, 0, k_x) \) were numerically calculated by the discrete spatial Fourier transform where \( \Delta x = 0.05 \) m and \(-6.4 \leq x \leq 6.35 \) m.

To estimate the synthesized sound field, the synthesis error at position \( x \) was defined as

\[
E(x) = 10 \log_{10} \left| \frac{S_{\text{org}}(x) - S_{\text{syn}}(x)}{|S_{\text{org}}(x)|^2} \right|^2,
\]

where \( S_{\text{org}}(x) \) and \( S_{\text{syn}}(x) \) are respectively the original and synthesized sound pressures at position \( x \).

Figure 4 shows the results of the original exterior sound field, the synthesized sound field, and the synthesis error in the horizontal plane for the proposed method at a temporal frequency of \( f = 2 \) kHz with \( y_0 = 0.5 \) and \(-0.5 \) m, respectively. The results of the averaged synthesis error calculated on a plane with \( z = 0 \) m, \(-1.5 \leq x \leq 1.5 \) m, and \( 1.0 \leq y \leq 3.0 \) m for both the numerical and proposed methods up to the spatial Nyquist frequency were plotted in Fig. 4. Those of the virtual point source synthesis discussed in Section 2.6 were also plotted as references. These results indicated that the proposed method achieved 2.5D exterior sound field synthesis where the original higher-order sound source is located both in front of and behind the linear loudspeaker array with identical accuracy as the numerical approach.

Consequently, based on angular spectrum decomposition, the proposed analytical driving function can effectively synthesize an exterior sound field described by the spherical harmonic expansion coefficients using a linear loudspeaker array.

4. CONCLUSIONS

This paper provided an analytical approach to the 2.5D synthesis of an exterior sound field described by spherical harmonic expansion coefficients using a linear loudspeaker array. The spherical harmonic expansion coefficients were analytically converted into 3D cylindrical ones by plane wave decomposition. The angular spectrum coefficients at the synthesis reference line were then analytically obtained from the converted 3D cylindrical harmonic expansion coefficients and directly synthesized by 2.5D SDM with a linear loudspeaker array. The results of computer simulations showed the effectiveness of the proposed analytical driving function.
5. REFERENCES


