# SOUND FIELD REPRODUCTION USING MULTIPLE LINEAR ARRAYS BASED ON WAVE FIELD RECONSTRUCTION FILTERING IN HELICAL WAVE SPECTRUM DOMAIN

*Shoichi Koyama[1], Ken'ichi Furuya[2], Yusuke Hiwasaki[1], Yoichi Haneda[3], and Yôiti Suzuki[4]*

[1] NTT Media Intelligence Laboratories, NTT Corporation, Tokyo, Japan
[2] Oita University, Oita, Japan.
[3] The University of Electro-Communications, Tokyo, Japan.
[4]Tohoku University, Sendai, Japan.

**ABSTRACT**

For with-height reproduction of a sound field, an efficient technique is to record and reproduce the sound field at lower resolution at an elevation angle than that at a horizontal angle based on auditory perception. To achieve this, developing a method using multiple horizontal linear arrays of microphones and loudspeakers is necessary. We propose a sound field reproduction method for a cylindrical array configuration which is based on a wave field reconstruction (WFR) filter analytically derived in the helical wave spectrum domain. The filter is stabilized by introducing a model in which microphones are mounted on a rigid cylindrical baffle. Numerical simulation results indicated that the reproduction accuracy of the proposed method was well preserved in the near-field region along the central axis of the cylinder even when the number of elements at angular position is small.

***Index Terms*—** Sound field reproduction, wave field reconstruction filter, helical wave spectrum

## 1. INTRODUCTION

Physical reproduction of a sound field is intended to achieve more realistic audio systems. For real-time recording and reproducing systems such as telecommunication systems, it is preferable that the driving signals of loudspeakers are directly transformed from the received signals of microphones. This means that parameters used to reproduce sound fields, such as source positions, directions, and original signals, are not required. We call this type of transformation sound-pressure-to-driving-signal (SP-DS) conversion. SP-DS conversion methods require a large number of microphones and loudspeakers when a three-dimensional (3-D) sound field is to be reproduced. From a perception point of view, the sound field at a horizontal angle is more important than that at an elevation angle. For reproduction including height (with-height reproduction), therefore, it is efficient to reproduce a sound field at a lower resolution in elevation direction as opposed to higher resolution in horizontal.

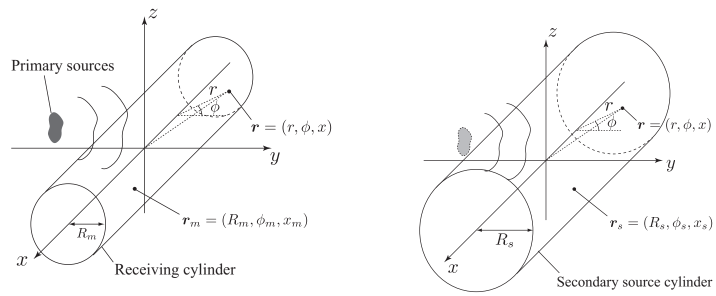
Methods based on the Kirchhoff-Helmholtz or Rayleigh integrals [1] are referred to as wave field synthesis (WFS) [2]. Based on the 3-D Rayleigh integral, WFS using a planar loudspeaker array is aimed at reproducing a with-height sound field. We have proposed an SP-DS conversion method for planar or linear microphone and loudspeaker arrays based on a *wave field reconstruction (WFR) filter* that is derived based on the Rayleigh integrals in the spatio-temporal frequency domain [3, 4]. It may seem that using WFR filter for planar arrays with a lower vertical resolution may be an immediate solution. However, a large vertical interval of the array elements leads to severe reproduction error due to spatial aliasing.

Ambisonics [5] and higher order Ambisonics (HOA) [6–8] are methods using circular or spherical arrays and are based on circular or spherical harmonics expansion of the sound field. They can be applied as SP-DS conversion methods through encoding and decoding processes. Lower resolution at elevation angle than horizontal angle can be achieved by using lower order expansion coefficients. This type of Ambisonics is called the mixed order approach [9, 10]. Even when a small number of loudspeakers is used at an elevation angle, the reproduction accuracy at the center of the sphere is preserved at a broad frequency band based on the properties of HOA [11]. An alternative method using multiple circular loudspeaker arrays was proposed by Gupta and Abhayapala [12].

We propose an SP-DS conversion method for cylindrical arrays of microphones and loudspeakers, which is based on WFR filtering in a helical wave spectrum domain. By using a cylindrical configuration, the reproduction accuracy on the central axis of the cylinder is expected to be preserved even when a small number of linear arrays is used. The proposed WFR filter is analytically derived and is stabilized by introducing a model in which microphones are mounted on a rigid cylindrical baffle. We conducted numerical simulations to compare the proposed method with WFR filtering for planar arrays.

## 2. DERIVATION OF WFR EQUATION IN CYLINDRICAL COORDINATES

As shown in Fig. 1, the sound field captured in the source area is reproduced in the target area. It is assumed that con-




(a) Source area ($r > R_m$) (b) Target area ($r < R_s$)

**Fig. 1**. Sound field reproduction using cylindrical distributions of receivers and secondary sources

tinuous, infinitely long, cylindrical distributions of receivers and secondary sources are arranged in the source and target areas, respectively, with the central axes on the $x$-axis. The radius of the receiving and secondary source cylinders is $R_m$ and $R_s$, respectively. It is assumed that $R_m \leq R_s$ and that the sound field inside the secondary source cylinder is reproduced to correspond with that inside the receiving cylinder. The position vector is denoted as $\boldsymbol{r} = (r, \phi, x)$ in the cylindrical coordinates. When the sound pressures on the receiving cylinder and the driving signals of the secondary sources are denoted as $P_{\mathrm{rcv}}(\boldsymbol{r}_m, \omega)$ and $D(\boldsymbol{r}_s, \omega)$ in the frequency domain, respectively, it is necessary to derive an equation that relates $P_{\mathrm{rcv}}(\boldsymbol{r}_m, \omega)$ and $D(\boldsymbol{r}_s, \omega)$, which is defined as the WFR equation. Here, $\boldsymbol{r}_m = (R_m, \phi_m, x_m)$ and $\boldsymbol{r}_s = (R_s, \phi_s, x_s)$ are the position vectors on the receiving and secondary source cylinders, and $\omega$ is the frequency. The WFR equation is derived based on the synthesized and desired sound fields described in the helical wave spectrum domain [1].

### 2.1. Synthesized sound field

The synthesized sound field $P_{\mathrm{syn}}(\boldsymbol{r}, \omega)$ is described as the surface integral of $D(\boldsymbol{r}_s, \omega)$ and $G(\boldsymbol{r} - \boldsymbol{r}_s, \omega)$, the transfer function between $\boldsymbol{r}$ and $\boldsymbol{r}_s$, on the secondary source cylinder.

$$P_{\mathrm{syn}}(\boldsymbol{r}, \omega) = \int_0^{2\pi}\int_{-\infty}^{\infty} R_s D(\boldsymbol{r}_s, \omega) G(\boldsymbol{r} - \boldsymbol{r}_s, \omega) d\phi_s dx_s. \quad (1)$$

This equation can be regarded as the spatial convolution of $D(\cdot)$ and $G(\cdot)$ with respect to $\phi$ and $x$. Therefore, (1) can be transformed into the helical wave spectrum domain as [7, 13]:

$$\tilde{P}_{\mathrm{syn},n}(r, k_x, \omega) = 2\pi R_s \tilde{D}_n(k_x, \omega) \tilde{G}_n(r - R_s, k_x, \omega), \quad (2)$$

where $n$ is the order and $k_x$ is the spatial frequency. The variable in the helical wave spectrum domain is represented by a tilde. Here, the transformation into the helical wave spectrum domain is defined as [1]:

$$\tilde{S}_n(k_x) = \frac{1}{2\pi}\int_0^{2\pi} d\phi \int_{-\infty}^{\infty} S(\phi, x) e^{-jn\phi} e^{-jk_x x} dx. \quad (3)$$

### 2.2. Desired sound field obtained using receivers on open cylinder

It is necessary to describe the desired sound field $P_{\mathrm{des}}(\boldsymbol{r}, \omega)$ by using $P_{\mathrm{rcv}}(\boldsymbol{r}_m, \omega)$. It can be described as an extrapolation of the sound field in the helical wave spectrum domain as [1]:

$$\tilde{P}_{\mathrm{des},n}(r, k_x, \omega) = \tilde{P}_{\mathrm{rcv},n}(R_m, k_x, \omega) \frac{J_n(k_r r)}{J_n(k_r R_m)}, \quad (4)$$

where $k_r = \sqrt{k^2 - k_x^2}$, $k = \omega / c$ is the wave number, $c$ is the sound speed, and $J_n(\cdot)$ is the $n$-th order Bessel function.

Since the synthesized and desired sound fields are to be matched, the WFR equation can be derived by solving (2) and (4) as:

$$\tilde{D}_n(k_x, \omega) = \tilde{F}_{\mathrm{open},n}(k_x, \omega) \tilde{P}_{\mathrm{rcv},n}(R_m, k_x, \omega), \quad (5)$$

where

$$\tilde{F}_{\mathrm{open},n}(k_x, \omega) = \frac{2}{\pi j R_s H_n^{(1)}(k_r R_s) J_n(k_r R_m)}. \quad (6)$$

Here, $H_n^{(1)}(\cdot)$ is the $n$-th order Hankel function of the first kind. For simplicity, we assumed that $G(\boldsymbol{r} - \boldsymbol{r}_s, \omega)$ has monopole characteristics [14]:

$$\tilde{G}_n(r - R_s, k_x, \omega) = \frac{j}{4} H_n^{(1)}(k_r R_s) J_n(k_r r). \quad (7)$$

Equation (6) is discretized and applied as a filter for transforming the received signals into the driving signals, i.e., the WFR filter. However, the denominator of (6) includes $J_n(k_r R_m)$, which means the WFR filter diverges at the zero of $J_n(k_r R_m)$. This instability of the WFR filter derives from the assumption that the receivers are on the acoustically transparent cylinder (open cylinder).

### 2.3. Desired sound field obtained using receivers on rigid cylinder

The above instability of the WFR filter can be avoided by assuming that the receivers are on a rigid and infinitely-long cylindrical baffle. When the rigid baffle is placed in the source area, the sound field, $P_{\mathrm{rcv}}(\boldsymbol{r}, \omega)$, can be described as the summation of the incident and scattering sound fields, $P_{\mathrm{inc}}(\boldsymbol{r}, \omega)$ and $P_{\mathrm{sct}}(\boldsymbol{r}, \omega)$, as [14]:

$$P_{\mathrm{rcv}}(\boldsymbol{r}, \omega) = P_{\mathrm{inc}}(\boldsymbol{r}, \omega) + P_{\mathrm{sct}}(\boldsymbol{r}, \omega). \quad (8)$$

The sound pressure gradient on the surface becomes zeros,

$$\frac{\partial}{\partial r}\left\{P_{\mathrm{inc}}(\boldsymbol{r}, \omega) + P_{\mathrm{sct}}(\boldsymbol{r}, \omega)\right\}\Big|_{r=R_m} = 0. \quad (9)$$

Here, $P_{\mathrm{inc}}(\cdot)$ and $P_{\mathrm{sct}}(\cdot)$ can be represented in the helical wave spectrum domain as [1]:

$$P_{\mathrm{inc}}(\boldsymbol{r},\omega)=\frac{1}{2\pi}\sum_{n=-\infty}^{\infty}\int_{-\infty}^{\infty}\breve{P}_{\mathrm{inc},n}(k_x,\omega)J_n(k_r r)e^{jn\phi}e^{jk_x x}dk_x, \quad (10)$$

and

$$P_{\mathrm{sct}}(\boldsymbol{r},\omega)=\frac{1}{2\pi}\sum_{n=-\infty}^{\infty}\int_{-\infty}^{\infty}\breve{P}_{\mathrm{sct},n}(k_x,\omega)H_n^{(1)}(k_r r)e^{jn\phi}e^{jk_x x}dk_x. \quad (11)$$

By substituting (10) and (11) into (9),

$$\breve{P}_{\mathrm{sct},n}(k_x,\omega)=-\frac{J_n'(k_r R_m)}{H_n^{(1)\prime}(k_r R_m)}\breve{P}_{\mathrm{inc},n}(k_x,\omega). \quad (12)$$

From (8), (10), (11) and (12),

$$P_{\mathrm{rcv}}(\boldsymbol{r},\omega)=\frac{1}{2\pi}\sum_{n=-\infty}^{\infty}\int_{-\infty}^{\infty}\left(J_n(k_r r)-\frac{J_n'(k_r R_m)}{H_n^{(1)\prime}(k_r R_m)}H_n^{(1)}(k_r r)\right)\cdot \breve{P}_{\mathrm{inc},n}(k_x,\omega)e^{jn\phi}e^{jk_x x}dk_x \quad (13)$$

The desired sound field corresponds to $P_{\mathrm{inc}}(\cdot)$. Therefore, it can be described by using the received sound pressure distribution from (13) as:

$$\begin{aligned}\tilde{P}_{\mathrm{des},n}(r,k_x,\omega)&=\breve{P}_{\mathrm{inc},n}(k_x,\omega)J_n(k_r r)\\&=-\frac{\pi j k_r R_m}{2}J_n(k_r r)H_n^{(1)\prime}(k_r R_m)\\&\qquad\cdot\tilde{P}_{\mathrm{rcv},n}(R_m,k_x,\omega). \quad (14)\end{aligned}$$

The following relation is used to derive (14).

$$J_n(\psi)H_n^{(1)\prime}(\psi)-J_n'(\psi)H_n^{(1)}(\psi)=\frac{2j}{\pi\psi}. \quad (15)$$

Similar to the case with the open receiving cylinder, the WFR equation for the receivers on the rigid baffle can be derived by solving (2) and (14) as:

$$\tilde{D}_n(k_x,\omega)=\tilde{F}_{\mathrm{rigid},n}(k_x,\omega)\tilde{P}_{\mathrm{rcv},n}(R_m,k_x,\omega), \quad (16)$$

where

$$\tilde{F}_{\mathrm{rigid},n}(k_x,\omega)=-\frac{k_r R_m H_n^{(1)\prime}(k_r R_m)}{R_s H_n^{(1)}(k_r R_s)}. \quad (17)$$

The Bessel function in the denominator of (6) is eliminated in that of (17). Therefore, the WFR filter based on (17) is much more stable than that based on (6). Similar results can be obtained in the context of plane-wave decomposition of a sound field by using a spherical or circular microphone array [15–17].

## 3. WFR FILTER FOR CYLINDRICAL ARRAYS AND ITS IMPLEMENTATION

As mentioned above, the WFR filter is obtained by discretizing (6) or (17) in order to apply it as an FIR filter. The block diagram of the proposed WFR filtering system for the case of the microphone array on the rigid baffle is shown in Fig. 2. The received signals of the microphone array are transformed into the helical wave spectrum domain by using 2-D spatial DFT. When each element of the helical wave spectrum of the received signals is denoted as $\tilde{P}_{ij}(\omega)$, the driving signals of the loudspeaker array in the helical wave spectrum domain $\tilde{D}_{ij}(\omega)$ are obtained by applying the WFR filter based on (17) as:

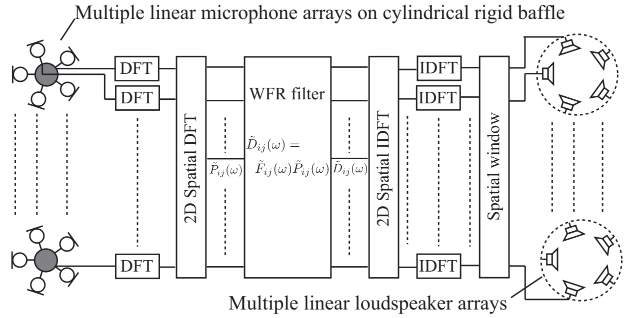



**Fig. 2**. Block diagram of proposed WFR filtering system

$$\tilde{D}_{ij}(\omega)=\tilde{F}_{ij}(\omega)\tilde{P}_{ij}(\omega), \quad (18)$$

where

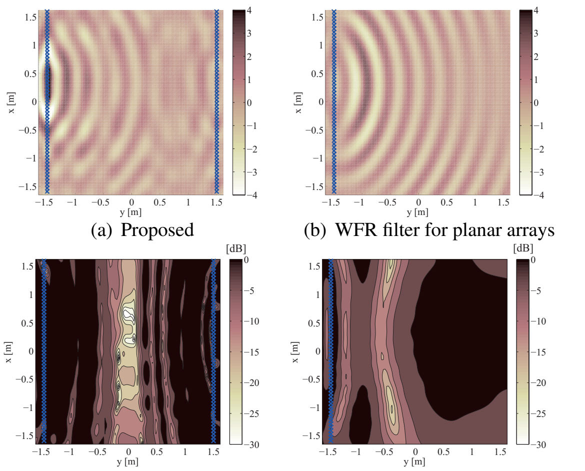
$$\tilde{F}_{ij}(\omega)=-\frac{\sqrt{k^2-k_{x,j}^2}R_m H_i^{(1)\prime}(\sqrt{k^2-k_{x,j}^2}R_m)}{R_s H_i^{(1)}(\sqrt{k^2-k_{x,j}^2}R_s)}. \quad (19)$$

If the WFR filter (19) includes evanescent wave components [1], this filter becomes very unstable, especially when $R_m \ll R_s$. Therefore, it is necessary to apply a tapering window that reduces the components of $|k_{x,j}|>|k|$ and $|i|>|kR_m|$ to (19). Additionally, in order to reduces the error derived from the finite approximation of the cylindrical arrays, the spatial tapering window is applied to the driving signals in the time domain with respect to the $x$-axis.

## 4. EXPERIMENTS

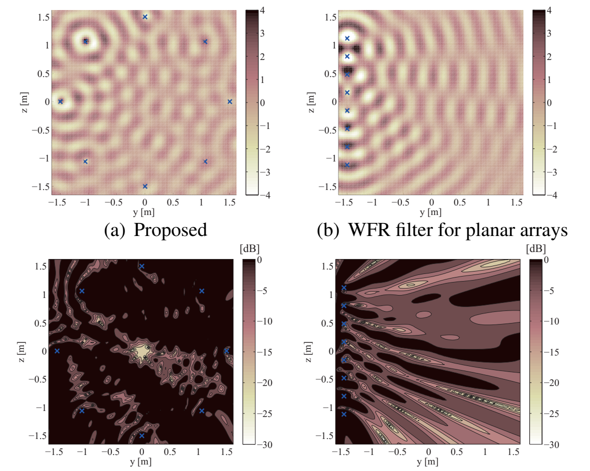
Numerical simulations were conducted under the free-field condition. We compared the proposed method with the WFR filtering for planar arrays [3]. The coordinates of the numerical simulations follow those given in Fig. 1.

In the proposed method, the number of elements in the cylindrical microphone and loudspeaker arrays was 64 at the axial positions ($x$) and 8 at the angular positions ($\phi$). The microphone array was mounted on an infinitely long rigid cylindrical baffle. The radii of the arrays, $R_m$ and $R_s$, were 0.28 and 1.5 m, respectively. The interval of the array elements in the direction of $x$ was 6 cm; therefore, the length of the arrays was 3.84 m. In the WFR filtering for planar arrays, the number of elements in the planar microphone and loudspeaker arrays was 64 at the positions of $x$ and 8 at the positions of $z$. The interval of the array elements in the direction of $x$ was 6 cm, and that in the direction of $z$ was 32 cm; therefore, the arrays were $3.84 \times 2.56$ m in size. The planar arrays were set on the $x$-$z$-plane at $y=-1.5$ m. The directivity of the array elements was assumed to be omni-directional. A point source as a primary source was located at $(r,\phi,x)=$(2.5 m,



(a) Proposed (b) WFR filter for planar arrays
(c) Proposed (d) WFR filter for planar arrays

**Fig. 3**. Simulation results of (a), (b) sound pressure and (c), (d) error distributions on $x$-$y$-plane at $z = 0$



(a) Proposed (b) WFR filter for planar arrays
(c) Proposed (d) WFR filter for planar arrays

**Fig. 4**. Simulation results of (a), (b) sound pressure and (c), (d) error distributions on $y$-$z$-plane at $x = 0$

155 deg, 0.3 m) in the source area. The source signal was a 1-kHz sinusoidal wave. A Tukey window function was applied to both methods as a spatial window whose sides tapered by 10%. The sound pressure distributions were simulated in a $3.3 \times 3.3$ m$^2$ region on the $x$-$y$-plane at $z = 0$ and $y$-$z$-plane at $x = 0$ at intervals of 3.0 cm. The amplitudes were normalized at the origin.

The error distributions were calculated as:

$$\mathrm{Err}(x_k, y_l) = 10\log_{10}\frac{\sum_m |p(t_m, x_k, y_l) - p_{\mathrm{org}}(t_m, x_k, y_l)|^2}{\sum_m |p_{\mathrm{org}}(t_m, x_k, y_l)|^2} \quad (20)$$

where $p(t_m, x_k, y_l)$ and $p_{\mathrm{org}}(t_m, x_k, y_l)$ are the reproduced and original sound pressure distributions in the time domain, respectively. The total amount of time was set for 480 samples and the sampling frequency was 48 kHz.

Figures 3 and 4 are the reproduced sound pressure and error distributions on the $x$-$y$-plane at $z = 0$ and $y$-$z$-plane at $x = 0$, respectively. The blue crosses represent the positions of the loudspeakers. With the proposed method, the reproduction accuracies on the neighboring region along the central axis of the cylindrical loudspeaker array were distinctly high. In contrast, the reproduction accuracies of the WFR filtering for planar arrays were low in the entire region. This reproduction error comes from the larger intervals of the microphones and loudspeakers in the direction of $z$, i.e., spatial aliasing. Even though it is possible to reduce this artifact by bringing these intervals closer, the reproduced region becomes smaller when the number of elements in the direction of $z$ is small.

In the proposed method, the reproduction accuracies along the central axis were well preserved even when the number of elements at angular positions ($\phi$) were small. Additionally, the loudspeakers in the direction where primary sources do not exist, which were the loudspeakers at $-\pi/2 \leq \phi \leq \pi/2$ in these experimental conditions, can be thinned approximately.

## 5. CONCLUSION

An SP-DS conversion method based on cylindrical array configurations of microphones and loudspeakers was proposed in order to efficiently reproduce a with-height sound field. The proposed method is based on a WFR filter analytically derived in the helical wave spectrum domain. The WFR filter is stabilized by introducing a model in which microphones are mounted on the rigid cylindrical baffle. Numerical simulations were conducted to compare the proposed method with WFR filtering for planar arrays. The reproduction accuracies of the proposed WFR filtering for cylindrical arrays were distinctly high in the near-field region along the central axis of the cylinder. In contrast, those of WFR filtering for planar arrays were low in the entire region. Therefore, the proposed method is more efficient for reproducing with-height sound fields than WFR filtering for planar arrays.

## 6. RELATION TO PRIOR WORK

The work presented here focused on a sound field reproduction method using cylindrical arrays of microphones and loudspeakers for efficient with-height reproduction, which is based on WFR filtering in the helical wave spectrum domain. This method can be regarded as an extension of the WFR filtering for planar or linear arrays we previously proposed [3, 4]. The mixed-order Ambisonics techniques [9, 10] are for spherical arrays. The method developed by Gupta and Abhayapala [12] is for multiple circular arrays. While these previously proposed methods were derived based on spherical harmonics expansion of the sound field, the proposed method is derived in the helical wave spectrum domain; therefore, the properties of the proposed method are similar to WFS in the horizontal angle and HOA in the elevation angle.

## 7. REFERENCES


[1] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*, Academic Press, London, 1999.

[2] A. J. Berkhout, D. D. Vries, and P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Am.*, vol. 93, no. 5, pp. 2764–2778, 1993.

[3] S. Koyama, K. Furuya, Y. Hiwasaki, and Y. Haneda, "Design of transform filter for sound field reproduction using microphone array and loudspeaker array," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust. (WASPAA)*, New Paltz, Oct. 2011, pp. 5–8.

[4] S. Koyama, K. Furuya, Y. Hiwasaki, and Y. Haneda, "Analytical approach to wave field reconstruction filtering in spatio-temporal frequency domain," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 21, no. 4, pp. 685–696, 2013 (in press).

[5] M. A. Gerzon, "Periphony: With-height sound field reproduction," *J. Audio Eng. Soc.*, vol. 21, pp. 2–10, Jan. 1973.

[6] M. Poletti, "Three-dimensional surround sound systems based on spherical harmonics," *J. Audio Eng. Soc.*, vol. 56, no. 11, pp. 1004–1025, 2005.

[7] J. Ahrens and S. Spors, "An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions," *Acta Acustica United with Acustica*, vol. 94, pp. 988–999, 2008.

[8] Y. J. Wu and T. D. Abhayapala, "Theory and design of soundfield reproduction using continuous loudspeaker concept," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 17, no. 1, 2009.

[9] C. Travis, "A new mixed-order scheme for ambisonic signals," in *Proc. Ambisonics Symposium*, Graz, 2009.

[10] J. Daniel, "Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonics format," in *Proc. 23rd Conf. AES*, Copenhagen, Denmark, May 2003.

[11] S. Spors and J. Ahrens, "A comparison of wave field synthesis and higher order ambisonics with respect to physical properties and spatial sampling," in *Proc. 125th Conv. AES*, San Francisco, USA, Oct. 2008.

[12] A. Gupta and T. D. Abhayapala, "Three-dimentional sound field reproduction using mulptiple circular arrays," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 5, 2011.

[13] J. Ahrens and S. Spors, "Sound field reproduction using planar and linear arrays of loudspeakers," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 8, pp. 2038–2050, 2010.

[14] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer, 1998.

[15] B. Rafaely, "Plane-wave decomposition of the sound field on a sphere by spherical convolution," *J. Acoust. Soc. Am.*, vol. 116, no. 4, pp. 2149–2157, 2004.

[16] H. Teutsch, *Modal Array Signal Processing: Principles and Applications of Acoustic Wavefield Decomposition*, Springer, 2007.

[17] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthogonal decomposition of the soundfield," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process. (ICASSP)*, 2002, pp. 1781–1784.