GENERATION OF MULTIPLE SOUND ZONES BY SPATIAL FILTERING IN WAVENUMBER DOMAIN USING A LINEAR ARRAY OF LOUDSPEAKERS

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

Novel signal processing is proposed for generating acoustically bright and dark zones at arbitrary horizontal positions using a linear array of loudspeakers. Most conventional methods are based on the numerical calculation of the inverse of the spatial correlation matrix between control points and the positions of the loudspeakers. However, such methods are quite unstable because the acoustic inverse problem is very ill-conditioned. On the other hand, in the proposed method, spatial filters in the wavenumber domain are analytically derived by introducing the spectral division method and modeling sound pressures at the control line as a rectangular window corresponding to bright and dark zones. Computer simulation results show that the proposed method can generate acoustically bright and dark zones more accurately than the conventional acoustic energy difference maximization method.

***Index Terms*—** Multizone, linear array of loudspeakers, spatial Fourier transform, spectral division method, spatial filtering

1. INTRODUCTION

Generating acoustically bright and dark zones [1] using multiple loudspeakers (Fig. 1(a)) is one important and attractive acoustic communication technique. By the superposition of various bright and dark zones, multiple sound zones can also be generated (Fig. 1(b)). These techniques are useful not only for personal sound systems [2–4] but also for multilingual guide services and other virtual reality applications. This paper focuses not on beamforming methods [2, 5, 6] but on methods to generate acoustically bright and dark zones using multiple loudspeakers.

Many methods have been proposed that control both acoustic contrast or the energy between two spaces [1, 3, 4, 7–11] and reproduce multiple sound fields [12–17] using multiple loudspeakers.

Most conventional methods are based on the numerical calculation of the inverse of the spatial correlation matrix between control points and the positions of loudspeakers [1,3,4, 8–10, 12–16]. However, such methods are quite unstable because the acoustic inverse problem is very ill-conditioned [18, 19]. This problem is the same in sound field reproduction based on the least squares method [20].

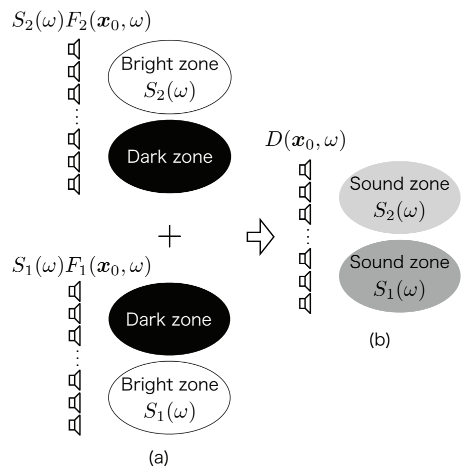


Fig. 1. Generation: (a) acoustically bright and dark zones; (b) multiple sound zones using a linear array of loudspeakers.

The acoustic energy difference maximization (EDM) [7] is based on the numerical calculation of the eigenvector of the spatial correlation matrix instead of the inverse. This method can generate acoustically bright and dark zones more accurately than the conventional acoustic contrast maximization [1]. However, it needs to calculate the reproduction results over again to select the optimal tuning factor.

To solve these problems, novel signal processing is proposed for generating acoustically bright and dark zones at arbitrary horizontal positions using a linear array of loudspeakers. The proposed method is based on analytically derived filters by introducing the spectral division method (SDM) [21], which is one sound field reproduction method for planer or linear arrays of loudspeakers. In the proposed method, an efficient spatial filter at the wavenumber domain is analytically derived by modeling sound pressures at the control line as a rectangular window corresponding to bright and dark zones.

2. GENERATION OF ACOUSTICALLY BRIGHT AND DARK ZONES BY SPATIAL FILTERING IN WAVENUMBER DOMAIN

2.1. Spectral division method in horizontal plane

The proposed method is based on the spectral division method in the horizontal plane for continuous linear distributions of secondary sources [21].

Acoustic pressure $P(\boldsymbol{x},\omega)$ synthesized at position $\boldsymbol{x} = [x,y,0]^{\mathrm{T}}$ with $y > 0$ by secondary sources at linear array ($y = 0$) is given as

$$P(\boldsymbol{x},\omega) = \int_{-\infty}^{\infty} D(\boldsymbol{x}_0,\omega) G_{\mathrm{3D}}(\boldsymbol{x}-\boldsymbol{x}_0,\omega) dx_0, \quad (1)$$

where $\omega = 2\pi f$ denotes the radial frequency, f denotes the temporal frequency, $D(\boldsymbol{x}_0,\omega)$ denotes the secondary source (driving signal) at position $\boldsymbol{x}_0 = [x_0,0,0]^{\mathrm{T}}$, $G_{\mathrm{3D}}(\boldsymbol{x}-\boldsymbol{x}_0,\omega)$ denotes the transfer function of the secondary source placed at $\boldsymbol{x}_0$ to point $\boldsymbol{x}$, and dx_0 stands for a surface element for integration. Under the free-field assumption, $G_{\mathrm{3D}}(\boldsymbol{x}-\boldsymbol{x}_0,\omega)$ is the three-dimensional free-field Green's function [22], defined as

$$G_{\mathrm{3D}}(\boldsymbol{x}-\boldsymbol{x}_0,\omega) = \frac{\exp(-jk|\boldsymbol{x}-\boldsymbol{x}_0|)}{4\pi|\boldsymbol{x}-\boldsymbol{x}_0|}, \quad (2)$$

where $j = \sqrt{-1}$, $k = \omega/c$ denotes the wavenumber and c is the speed of sound. When applying the spatial Fourier transform to (1) with respect to the x-axis, the convolution along it is performed by the convolution theorem:

$$\tilde{P}(k_x,y,0,\omega) = \tilde{D}(k_x,\omega) \cdot \tilde{G}(k_x,y,0,\omega), \quad (3)$$

where k_x denotes the spatial frequency in the direction of x and $\tilde{G}(k_x,y,0,\omega)$ is the spatial Fourier transpose of $G_{\mathrm{3D}}(\boldsymbol{x}-\boldsymbol{x}_0,\omega)$ with respect to the x-axis calculated in [23], given as

$$\tilde{G}(k_x,y,0,\omega) = \begin{cases} -\frac{j}{4} H_0^{(2)}\left(\sqrt{k^2-k_x^2}\,y\right) \\ \qquad \text{for } 0 \leq |k_x| \leq |k| \\ \frac{1}{2\pi} K_0\left(\sqrt{k_x^2-k^2}\,y\right) \\ \qquad \text{for } 0 < |k| \leq |k_x|, \end{cases} \quad (4)$$

where $H_0^{(2)}$ denotes the zero-th order Hankel function of the second kind and K_0 is the zero-th order modified Bessel function of the second kind [22].

When the continuous receiver line is located at $y = y_{\mathrm{ref}}$, the secondary source driving function in the wavenumber domain is directly given by

$$\tilde{D}(k_x,\omega) = \frac{\tilde{P}(k_x,y_{\mathrm{ref}},0,\omega)}{\tilde{G}(k_x,y_{\mathrm{ref}},0,\omega)}, \quad (5)$$

in the spectral division method.

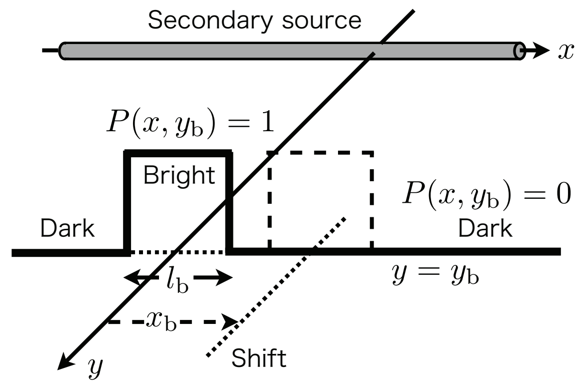


Fig. 2. $P(x,y_{\mathrm{b}})$ modeled by rectangular window for generating acoustically bright and dark zones.

2.2. Spatial filtering in wavenumber domain

In this subsection, spatial filters in the wavenumber domain are proposed to generate acoustically bright and dark zones.

For the reproduction of each sound source $S_i(\omega)$ at each zone using a linear array of loudspeakers, each filter $F_i(\boldsymbol{x}_0,\omega)$ for each sound source $S_i(\omega)$ at each zone is calculated. ($i = 1,2,\cdots,M$). When the number of sound sources is M, secondary source driving function $D(\boldsymbol{x}_0,\omega)$, which represents the superposition of each $S_i(\omega)F_i(\boldsymbol{x}_0,\omega)$, is given as

$$D(\boldsymbol{x}_0,\omega) = \sum_{i=1}^{M} S_i(\omega) F_i(\boldsymbol{x}_0,\omega). \quad (6)$$

The case of $M = 2$ is shown in Fig. 1.

Next the case of $M = 1$ and $S_1(\omega) = 1$ is considered. This is equivalent to generating an acoustically bright zone using a linear array of loudspeakers. In the following equations, M is omitted for simplicity. Then $D(\boldsymbol{x}_0,\omega) = F(\boldsymbol{x}_0,\omega)$, and (5) is represented as

$$\tilde{F}(k_x,\omega) = \frac{\tilde{P}(k_x,y_{\mathrm{ref}},0,\omega)}{\tilde{G}(k_x,y_{\mathrm{ref}},0,\omega)}. \quad (7)$$

From (7), filter $\tilde{F}(k_x,\omega)$ is calculated as the spatial filter in the wavenumber domain for generating an acoustically bright zone using a linear array of loudspeakers.

In sound field reproduction, $P(\boldsymbol{x},\omega)$ at position $\boldsymbol{x} = [x,y_{\mathrm{ref}},0]$ is the actual acoustic pressure received at the continuous receiver line located at $y = y_{\mathrm{ref}}$. In the proposed method, on the other hand, to generate acoustically bright and dark zones at the continuous receiver line, $P(\boldsymbol{x},\omega)$ is continuously set to 1 or 0 at all temporal frequencies. The positions at $P(\boldsymbol{x},\omega) = 1$ and 0 correspond to the bright and dark points, respectively.

For generating the bright zone of length l_{b} centered around $x = 0$ at $y = y_{\mathrm{b}}$, $P(x,y_{\mathrm{b}})$ is modeled by a rectangular window $\Pi(x/l_\mathrm{b})$ shown in Fig. 2 and given as

$$P(x,y_\mathrm{b}) = \Pi\left(\frac{x}{l_\mathrm{b}}\right) = \begin{cases} 1, & \text{for } |x| \leq l_\mathrm{b}/2 \\ 0, & \text{elsewhere.} \end{cases} \quad (8)$$

Then the Fourier transform of $P(x,y_\mathrm{b})$ with respect to x is [22]

$$\tilde{P}(k_x) = \frac{l_\mathrm{b}\sin(k_x l_\mathrm{b}/2)}{(k_x l_\mathrm{b}/2)} = l_\mathrm{b}\,\mathrm{sinc}\left(\frac{k_x l_\mathrm{b}}{2\pi}\right). \quad (9)$$

In addition, for shifting the center of the bright zone from $x = 0$ to $x = x_\mathrm{b}$, as shown in Fig. 2, the shift theorem [22] with respect to x is applied to (9) and the Fourier transform of shifted rectangular window $\tilde{P}_\mathrm{shift}(k_x)$ is represented as

$$\begin{aligned}\tilde{P}_\mathrm{shift}(k_x) &= \tilde{P}(k_x)\exp(jk_x x_b) \\ &= l_\mathrm{b}\,\mathrm{sinc}\left(\frac{k_x l_\mathrm{b}}{2\pi}\right)\exp(jk_x x_b). \end{aligned} \quad (10)$$

Then $\tilde{P}_\mathrm{shift}(k_x)$ is assigned to (7), and the spatial filter in the wavenumber domain for generating bright zone $\tilde{F}(k_x,\omega)$ is analytically derived as

$$\tilde{F}(k_x,\omega) = \frac{l_\mathrm{b}\,\mathrm{sinc}\,(k_x l_\mathrm{b}/2\pi)\exp(jk_x x_\mathrm{b})}{\tilde{G}(k_x, y_\mathrm{b}, 0, \omega)}. \quad (11)$$

As a result, a bright zone of arbitrary length l_b can be generated at arbitrary horizontal position $[x_\mathrm{b}, y_\mathrm{b}]$ by the proposed spatial filter in the wavenumber domain $\tilde{F}(k_x,\omega)$ using a continuous linear secondary monopole source distribution.

2.3. Practical implementation using a linear array of loudspeakers

In a practical implementation, (11) must be discretized and truncated [22]. When the channel number of loudspeakers is N and the distance between adjacent loudspeakers is Δx, the length of the array is $L = N\Delta x$ and the spatial inverse discrete Fourier transform (IDFT) of $\tilde{F}(k_x,\omega)$ is calculated as

$$F_\mathrm{IDFT}(x,\omega) = \frac{1}{L}\sum_{m=-N/2}^{N/2-1}\tilde{F}(k_x,\omega)\exp(-2\pi jmn/N), \quad (12)$$

where $x = n\Delta x$, $k_x = 2\pi m/N\Delta x$, and $-N/2 \leq n \leq N/2-1$ [22].

In the proposed method, the number of IDFT points in (12) is set to more than $2 \times N$ to reduce the truncation error. This strategy is the same in a practical implementation of nearfield acoustical holography [22]. More than $2 \times N$ channels of $F_\mathrm{IDFT}(\boldsymbol{x},\omega)$ are calculated, and their centered N channels are finally used for filters in the temporal frequency domain.

3. COMPUTER SIMULATIONS

Computer simulations compared the proposed method with the conventional EDM [7].

3.1. Acoustic energy difference maximization

EDM is briefly introduced here. When using K channels of control points at $\boldsymbol{x}_\mathrm{co}$ and N channels of loudspeakers at $\boldsymbol{x}_\mathrm{sp}$, the spatial averaged correlation matrix between control points and the positions of loudspeakers is calculated as $\boldsymbol{R}(\omega) = 1/K\sum_{i=1}^{K}\boldsymbol{G}_i^\mathrm{H}(\omega)\boldsymbol{G}_i(\omega)$, where $\boldsymbol{G}_i(\omega) = [G_\mathrm{3D}(\boldsymbol{x}_{\mathrm{co},i} - \boldsymbol{x}_{\mathrm{sp},1},\omega)\ \cdots\ G_\mathrm{3D}(\boldsymbol{x}_{\mathrm{co},i} - \boldsymbol{x}_{\mathrm{sp},N},\omega)]$. $\boldsymbol{R}_\mathrm{b}(\omega)$ is the spatial correlation matrix between the bright points and the positions of the loudspeakers and $\boldsymbol{R}_\mathrm{d}(\omega)$ is that between the dark points and the positions of the loudspeakers. The filters of loudspeakers $\boldsymbol{F}(\omega)$ are obtained from the eigenvector of matrix $[\boldsymbol{R}_\mathrm{b}(\omega) - \alpha\boldsymbol{R}_\mathrm{d}(\omega)]$ that corresponds to the largest eigenvalue of this matrix, where α is a tuning factor [7].

3.2. Simulation condition

A sound field is assumed in a three-dimensional free field where a spherical wave of a single frequency is radiated from each loudspeaker. The speed of sound is 343.36 m/s. A linear array of loudspeakers is set along the x-axis and centered around $x = 0$. The channel number of the linear array of loudspeakers N is 64 and the distance between adjacent loudspeakers Δx is 0.05. The length of array L is 3.2 m. The position and the length of the bright zone are $[x_\mathrm{b}, y_\mathrm{b}] = [-0.8, 2]$ and $l_\mathrm{b} = L/2 = 1.6$ m, respectively.

In the EDM, a linear array of 64 control points is set parallel to an array of loudspeakers apart from y_b along the y-axis. 32 control points at $x < 0$ are bright points $\boldsymbol{x}_\mathrm{b}$, and the other half are dark points $\boldsymbol{x}_\mathrm{d}$. Turning factor α is set to 0.9999 for maximizing the acoustic energy difference between two zones from the pre-simulations.

In the proposed method, the number of IDFT points in Eq. (12) is $N \times 2 = 128$, and the 64 centered channels of $F_\mathrm{IDFT}(\boldsymbol{x},\omega)$ are used as filters of the 64 loudspeakers.

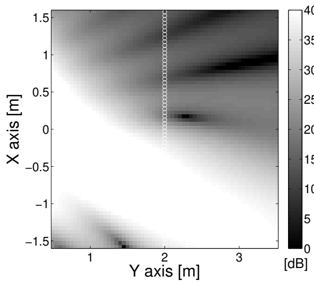
The reproduced sound pressure level at position $\boldsymbol{x}$ is defined as

$$P_\mathrm{SPL}(\boldsymbol{x},\omega) = 10\log_{10}\left|\hat{P}(\boldsymbol{x},\omega)\right|^2, \quad (13)$$

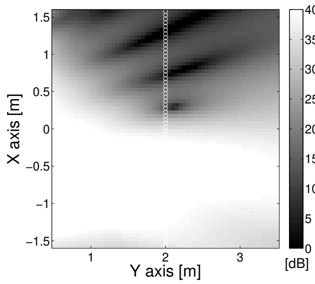
where $\hat{P}(\boldsymbol{x},\omega)$ is the reproduced sound pressure at position $\boldsymbol{x}$.

Furthermore, for evaluating the sound pressure level between bright zone $\boldsymbol{x}_\mathrm{b}$ and dark zone $\boldsymbol{x}_\mathrm{d}$, bright to dark ratio $BDR(\omega)$ is also defined as

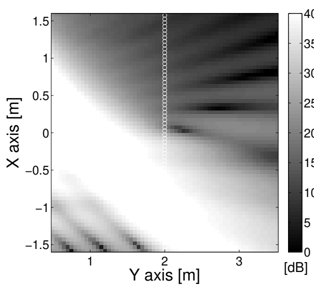
$$BDR(\omega) = 10\log_{10}\frac{\sum_{\boldsymbol{x}_\mathrm{b}}\left|\hat{P}(\boldsymbol{x}_\mathrm{b},\omega)\right|^2}{\sum_{\boldsymbol{x}_\mathrm{d}}\left|\hat{P}(\boldsymbol{x}_\mathrm{d},\omega)\right|^2}. \quad (14)$$



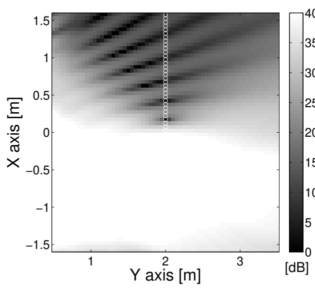
(a) EDM ($f = 500$ Hz)



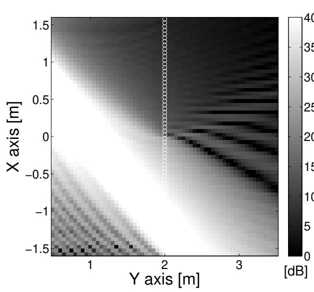
(b) Proposed ($f = 500$ Hz)



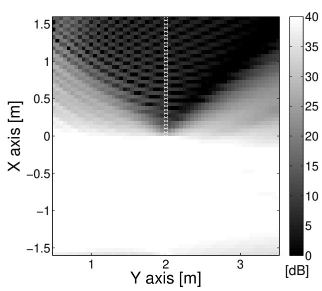
(c) EDM ($f = 1$ kHz)



(d) Proposed ($f = 1$ kHz)



(e) EDM ($f = 3$ kHz)



(f) Proposed ($f = 3$ kHz)

Fig. 3. Experimental results of reproduced sound pressure level $P_{\mathrm{SPL}}(\boldsymbol{x}, \omega)$ defined in (13). White circles parallel with $y = 2$ are control points.

For calculating (14), both zones are centered around control line $y = y_{\mathrm{b}}$. The length of each zone is $L/2 = 1.6$ m, and the width of each zone is 0.4 m.

3.3. Simulation results

The results of the reproduced sound pressure level $P_{\mathrm{SPL}}(\boldsymbol{x}, \omega)$ and the bright to dark ratio $BDR(\omega)$ are shown in Figs. 3 and 4, respectively. Both results show that the proposed method generated bright zones more accurately than the conventional EDM, especially at more than 1 kHz. The results of Fig. 3 also show that the length of the bright zone generated by the EDM at each frequency is different. In contrast, the lengths in the proposed method are constant at all frequencies. Consequently, the proposed spatial filtering can efficiently control sound pressures at the target line.

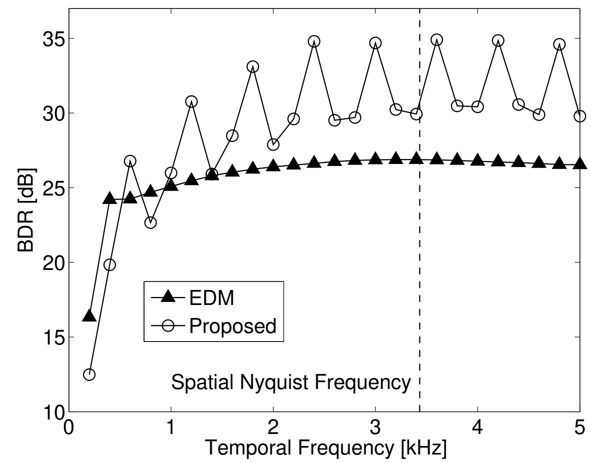


Fig. 4. Experimental results of bright to dark ratio $BDR(\omega)$ defined in (14).

4. CONCLUSION

A method was proposed to generate acoustically bright and dark zones at arbitrary horizontal positions using a linear array of loudspeakers. In the proposed method, the spatial filter in the wavenumber domain was analytically derived by introducing the spectral division method in the horizontal plane for continuous linear distributions of secondary sources and modeling sound pressures at the control line as a rectangular window corresponding to the bright and dark zones. The computer simulation results show that the proposed method generated bright and dark zones more accurately than conventional acoustic energy difference maximization.

5. RELATION TO PRIOR WORK

The work presented in this paper focused on a method for generating acoustically bright and dark zones using many loudspeakers. Most conventional methods are based on the numerical calculation of the inverse of the spatial correlation matrix between control points and the positions of loudspeakers [1,3,4,8–10,12–16]. However, they are quite unstable because the acoustic inverse problem is very ill-conditioned [18, 19]. On the other hand, the proposed method is based on analytically derived filters and introduces the spectral division method [21]. An efficient spatial filter at the wavenumber domain is analytically derived by modeling sound pressures at the control line as a rectangular window for generating acoustically bright and dark zones at arbitrary horizontal positions using a linear array of loudspeakers.

6. ACKNOWLEDGEMENT

This study was partly supported by Grant-in-Aid for Young Scientists B (No. 25871208) from the JSPS.

7. REFERENCES

[1] J.-W. Choi and Y.-H. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources," *J. Acoust. Soc. Am.,* vol. 111, no. 4, pp. 1695–1700, Apr. 2002.

[2] W. F. Druyvesteyn and J. Garas, "Personal sound," *J. Audio Eng. Soc.*, vol. 45, no. 9, pp. 685–701, Sept. 1997.

[3] J.-H. Chang, C.-H. Lee, J.-Y. Park, and Y.-H. Kim, "A realization of sound focused personal audio system using acoustic contrast control," *J. Acoust. Soc. Am.,* vol. 125, no. 4, pp. 2091–2097, Apr. 2009.

[4] S. J. Elliott, J. Cheer, J.-W. Choi, and Y. Kim, "Robustness and regularization of personal audio systems," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 7, pp. 2123–2133, Sept. 2012.

[5] J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing*, Berlin: Springer-Verlag, 2008.

[6] M. M. Boone, W.-H. Cho, and J.-G. Ih, "Design of a highly directional endfire loudspeaker array," *IEEE J. Audio Eng. Soc.*, vol. 57, no. 5, pp. 309–325, May 2009.

[7] M. Shin, S. Q. Lee, F. M. Fazi, P. A. Nelson, D. Kim, S. Wang, K. H. Park, and J. Seo, "Maximization of acoustic energy difference between two spaces," *J. Acoust. Soc. Am.,* vol. 128, no. 1, pp. 121–131, July 2010.

[8] J.-H. Chang and F. Jacobsen, "Sound field control with a circular double-layer array of loudspeakers," *J. Acoust. Soc. Am.,* vol. 131, no. 6, pp. 4518–4525, June 2012.

[9] Y. Cai, M. Ming, and J. Yang, "Design of a time-domain acoustic contrast control for broadband input signals in personal audio systems," *Proc. ICASSP 2013*, pp. 341–345, May 2013.

[10] P. Coleman, P. Jackson, M. Olik, and J. A. Pedersen, "Optimizing the planarity of sound zones," *Proc. 52nd Int. Conf. AES*, Sept. 2013.

[11] Y. Nakayama, M. Adachi, K. Ishimoto, and Y. Tatekura, "Individual sound image generation for multiple users based on loudspeaker array with NBSFC," *Proc. DSP 2013*, July 2013.

[12] M. Poletti, "An investigation of 2D multizone surround sound systems," *Proc. 125th Conv. AES*, 7551, Oct. 2008.

[13] T. D. Abhayapala and Y. J. Wu, "Spatial soundfield reproduction with zones of quiet," *Proc. 127th Conv. AES*, 7887, Oct. 2009.

[14] Y. J. Wu and T. D. Abhayapala, "Spatial multizone soundfield reproduction: theory and design," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 6, pp. 1711–1720, Aug. 2011.

[15] M. B. Moller, M. Olsen, and F. Jacobsen, "A hybrid method combining synthesis of a sound field and control of acoustic contrast," *Proc. 132nd Conv. AES*, Apr. 2012.

[16] N. Radmanesh and I. S. Burnett, "Generation of isolated wideband sound fields using a combined two-stage lasso-ls algorithm," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 21, no. 2, pp. 378–387, Feb. 2013.

[17] W. Jin, W. B. Kleijn, and D. Virette, "Multizone soundfield reproduction using orthogonal basis expansion," *Proc. ICASSP 2013*, pp. 311–315, May 2013.

[18] W. A. Veronesi, and J. D. Maynard, "Digital holographic reconstruction of sources with arbitrarily shaped surfaces," *J. Acoust. Soc. Am.,* vol. 85, no. 2, pp. 588–598, Feb. 1989.

[19] F. M. Fazi, "Sound field reproduction," Ph.D. dissertation, Univ. of Southampton, Feb. 2010.

[20] O. Kirkeby and P. Nelson, "Reproduction of plane wave sound fields," *J. Acoust. Soc. Am.,* vol. 94, no. 5, pp. 2992–3000, Nov. 1993.

[21] 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, Nov. 2010.

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

[23] I. S. Gradshteyn, and I. M. Ryzhik, *Table of Integrals, Series, and Products*, San Diego: Academic Press, 2000.