MULTIZONE SOUNDFIELD REPRODUCTION USING ORTHOGONAL BASIS EXPANSION

Wenyu Jin, W. Bastiaan Kleijn

Victoria University of Wellington
New Zealand

David Virette

Huawei European Research Center
Germany

ABSTRACT

We introduce a method for 2-D spatial multizone soundfield reproduction based on describing the desired multizone soundfield as an orthogonal expansion of basis functions over the desired reproduction region. This approach finds the solution to the Helmholtz equation that is closest to the desired soundfield in a weighted least squares sense. The basis orthogonal set is formed using QR factorization with as input a suitable set of solutions of the Helmholtz equation. The coefficients of the Helmholtz solution wavefields can then be calculated, reducing the multizone sound reproduction problem to the reconstruction of a set of basis wavefields over the desired region. The method facilitates its application with a more practical loudspeaker configuration. The approach is shown effective for both accurately reproducing sound in the selected bright zone and minimizing sound leakage into the predefined quiet zone.

1. INTRODUCTION

Multizone soundfield reproduction is a technology that aims at providing an individual sound environment to each of a set of listeners using only loudspeakers. The main research goal is to find the best trade-off between reproducing the desired soundfield accurately in the pre-defined acoustic "bright" zones and minimizing the sound pressure in the specified "quiet" zones. Secondary goals include the improvement of system robustness, minimizing computational effort, and minimizing hardware requirements.

The origin of work in the area can be traced to 2002, when [1] proposed an acoustic contrast control method to maximize the ratio of the mean square sound pressure in the bright and quite zones. A personal audio system was implemented in [2], which obtained a channel separation of over 20 dB between the bright and quiet zone. Improved robustness was obtained by [3]. The focus of [1, 2, 3] is limited to the control of acoustic contrast between the different zones; the approach does not control the soundfield in the bright zone.

In [4], higher order ambisonics (HOA) was introduced to reproduce soundfield in multizones based on mode matching. In 2008, Poletti [5] proposed an approach of using standard least squares method to generate a 2-D monochromatic multizone soundfield. [6] studied the application to narrowband speech signals. Global control of the multizone soundfield was proposed in [7] using the translation of spatial harmonic coefficient between coordinate systems.

Except for [3], the fore-mentioned approaches do not separately control the leakage of energy into the quiet zone. A method for using spatial band stop filtering to suppress the inter-zone sound interference was proposed in [8]. It facilitates acoustic attenuation only outside the control region, which makes the method not useful for practical applications. In [9], the multizone sound reproduction was formulated as a constrained optimization problem to control the sound in each zone individually, while also controlling the leakage into the zones of other listeners.

Multizone audio systems often involve the inversion of a matrix representing the acoustic transfer function, e.g., [5, 7, 9]. The inversion can become ill conditioned in scenarios when the zones are not carefully positioned [10]. The ill-conditioning problem in the soundfield reproduction can be reduced in severity by introducing regularization such as Tikhonov method [10][11].

In this paper, we propose a method of describing a desired multizone soundfield to control the reproduction in each specified zone independently. Our approach consists of two distinct stages. We first match the desired multizone soundfield to an orthogonal expansion of basis functions, which are solutions of the Helmholtz equation. Thus, this first step maps the desired soundfield onto the nearest physically feasible one, which facilitates to eliminate the ill-conditioning problem that is caused by the poor positioning of the selected zones. The complicated multizone sound reproduction problem is then reduced to the reproduction of the elementary Helmholtz solution wavefields (which can be formulated as planewaves or circular waves) in the second step using an array of loudspeakers over the region and combine them to obtain the basis functions used for the feasible solution. Existing reproduction methods such as those in [12, 13, 14] can be used in the reproduction step. As we are interested in methods that facilitate practical applications, we extend the methods of [13] to include reproduction from only a segment of a circular array of loudspeakers. Our simulation results confirm the effectiveness of the resulting multizone system in terms of accurately recreating the desired soundfield in the bright zone and in terms of achieving controllable acoustic attenuation in the quiet zone.

2. PROBLEM STATEMENT

We seek to control the reproduction of the desired soundfield over each of a set of pre-defined zones individually. We consider the case where all zones lie in a 2-D plane.

As illustrated in Fig.1, the desired reproduction region $\mathbb{D}$ is the entire control zone of interest with a radius of r, which includes both the acoustic bright zone $\mathbb{D}_b$ and the quiet zone $\mathbb{D}_q$. We define the remaining area as the unattended zone. The number of employed loudspeakers is Q and the qth loudspeaker weight is denoted as $l_q(k)$, where $k = 2\pi f/c$ is the wavenumber, f is the frequency and c is the speed of sound propagation.

Our main objective is to accurately reproduce the desired soundfield inside the zone $\mathbb{D}_b$ while controlling the acoustic energy leaked into the zone $\mathbb{D}_q$. To evaluate the performance of our system we use the following two measures:

- The acoustic brightness contrast between the bright zone $\mathbb{D}_b$ and the quiet zone $\mathbb{D}_q$ to quantify sound leakage between the

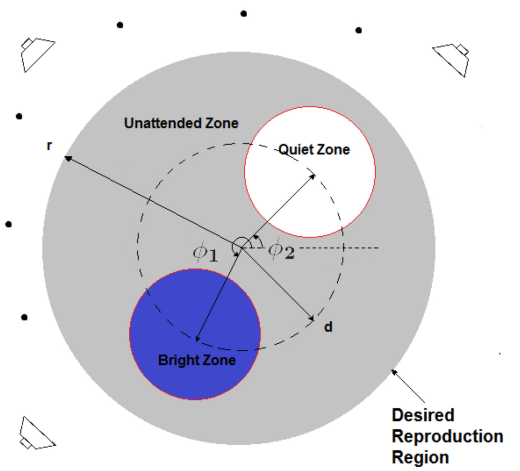


Fig. 1. The 2-D multizone soundfield reproduction in the desired reproduction region of radius r with Q loudspeakers.

two zones [2]:

$$\zeta(k) = \frac{\frac{1}{\|\mathbb{D}_b\|}\int_{\mathbb{D}_b}|S(\mathbf{x},k)|^2 d\mathbf{x}}{\frac{1}{\|\mathbb{D}_q\|}\int_{\mathbb{D}_q}|S(\mathbf{x},k)|^2 d\mathbf{x}}, \tag{1}$$

where $\mathbf{x}$ denotes an arbitrary spatial observation point. $\|\mathbb{D}_b\|$ and $\|\mathbb{D}_q\|$ mark the sizes of $\mathbb{D}_b$ and $\mathbb{D}_q$.

- The mean square error (MSE) between the desired sound $S^d(\mathbf{x},k)$ and the actually rendered sound $S^a(\mathbf{x},k)$ over the zone $\mathbb{D}_b$ to gauge the reproduction accuracy:

$$\epsilon_M(k) = \frac{\int_b |S^d(\mathbf{x},k) - S^a(\mathbf{x},k)|^2 d\mathbf{x}}{\int_b |S^d(\mathbf{x},k)|^2 d\mathbf{x}}. \tag{2}$$

3. SOUNDFIELD BASIS EXPANSION

In this section, we introduce a method to describe a desired soundfield as an orthogonal expansion of basis functions for the reproduction region. An arbitrary 2-D (height-invariant) soundfield function $S(\mathbf{x},k)$ satisfying the wave equation can be written as a superposition of an orthogonal set of solutions of the Helmholtz equation [15]. Therefore, we can write $S(\mathbf{x},k) : \mathbb{D} \times \mathbb{R} \mapsto \mathbb{C}$ as a weighted series of basis functions $\{G_n\}$

$$S(\mathbf{x},k) = \sum_n C_n G_n(\mathbf{x},k). \tag{3}$$

If the orthogonal set is *complete* in the space of feasible solutions on $\mathbb{D}$, all feasible solutions can be described.

To define the basis functions, we use the weighted inner product

$$\langle Y_1|Y_2\rangle = \int_{\mathbb{D}} Y_1(\mathbf{x})Y_2^*(\mathbf{x})w(\mathbf{x})d\mathbf{x}, \tag{4}$$

where $Y_1 \in \mathbb{D}$ and $Y_2 \in \mathbb{D}$. Given a desired soundfield $S^d(\mathbf{x},k)$, computing the expansion coefficients $C_n = \langle S^d(\mathbf{x},k)|G_n(\mathbf{x},k)\rangle$ then gives the coefficients C_n that minimize the weighted squared error

$$\min_{C_n} \int_{\mathbb{D}} \|\sum_n C_n G_n(\mathbf{x},k) - S^d(\mathbf{x},k)\|^2 w(\mathbf{x})d\mathbf{x}. \tag{5}$$

The weighting function $w(\mathbf{x})$ specifies the relative importance of the reproduction accuracy for each point in space. In the present case, where we split $\mathbb{D}$ into discrete bright, quiet and unattended zones, it is reasonable to assign a fixed weight to each zone:

$$w(\mathbf{x}) = \begin{cases} a, & \mathbf{x} \in \text{the bright zone} \\ b, & \mathbf{x} \in \text{the quiet zone} \\ c, & \mathbf{x} \in \text{the unattended zone} \end{cases} \tag{6}$$

To construct the orthogonal set of basis functions that can be used to describe the desired multizone soundfield, it is convenient to start with a set of functions $F_n(\mathbf{x},k)$ that represent N planewaves arriving from angles $\hat{\phi}_n = (n-1)\Delta\phi,\ n = 1,\cdots,N$

$$F_n(\mathbf{x},k) = e^{ik\mathbf{x}\hat{\phi}_\mathbf{n}}, \tag{7}$$

where $\hat{\phi}_\mathbf{n}$ is the unit vector in the direction of the nth planewave. As the $F_n(\mathbf{x},k)$ are independent on $\mathbb{D}$ and as the set of physically feasible independent solutions is finite, selecting $\Delta\phi$ sufficiently small and N sufficiently large leads to a set of functions that can describe any feasible solution. We can then use a Gram-Schmidt orthogonalization to find a set of orthonormal basis functions $G_n(\mathbf{x},k)$ over $\mathbb{D}$, $n \in \{1,\cdots,N\}$. In practice this orthonormalization can be performed using a QR factorization by means of either a Householder transform or with a Gram-Schmidt process [16] on discretized functions, which results in

$$F_n(\mathbf{x},k) = \sum_j \mathbf{R}_{jn} G_j(\mathbf{x},k), \tag{8}$$

where $\mathbf{R}$ is a upper triangular matrix.

Using (8) and (3) it is now straightforward to express the feasible approximation of the desired soundfield:

$$S(\mathbf{x},k) = \sum_{n,j} C_n(\mathbf{R}^{-1})_{jn} F_j(\mathbf{x},k). \tag{9}$$

Therefore, $P_j = \sum_n C_n(\mathbf{R}^{-1})_{jn}$ specifies the coefficient for the jth ($j \in \{1,\cdots,N\}$) planewave function to construct the desired multizone soundfield.

4. THE REPRODUCTION OF THE LINEAR-COMBINED PLANEWAVES

In the first stage, we map the desired multizone soundfield onto the closest Helmholtz solution, in the least-squares sense, and express it as an expansion of planewave functions in (9). With the outcome of the first stage, which is the coefficient set P, the multizone sound reproduction problem is reduced to the reproduction of planewaves $F_n(\mathbf{x},k)$ over the region $\mathbb{D}$.

We now present a generalization of the reproduction method proposed in [13] that facilitates practical implementations. The essence of this problem is to find a set of Fourier coefficients for the aperture function, such that it can be used to approximate the desired soundfield, and such that the aperture vanishes over an angular window. The latter property facilitates a reduced number of loudspeakers in a practical configuration (e.g., a semi-circle or even quarter-circle of loudspeaker arrangement).

We first write the loudspeaker aperture function $\rho(\phi,k)$ on a full circle as a Fourier series expansion as it is a periodic function of the angle ϕ:

$$\rho(\phi,k) = \sum_{m=-M}^{M} \beta_m(k)e^{im\phi}, \tag{10}$$

where $M = \lceil kr \rceil$ is the truncation length [12] and $H_0^{(1)}(k\|\cdot\|)$ is a zeroth-order Hankel function of the first kind [17]. $\{\beta_m(k)\}$ are the Fourier coefficients and $\beta_m^0(k) = \frac{2}{i\pi H_m^{(1)}(kR_l)}\alpha_m^{(d)}(k)$ is derived for the full circular continuous aperture function of radius R_l in [13]. $\alpha_m^{(d)}(k)$ uniquely represents the desired soundfield. To illustrate, the desired soundfield coefficients for a planewave source, $\alpha_m^{(d)}(k)$ are given by:

$$\alpha_m^{(d)}(k) = i^m e^{-im\phi_{pw}}, \tag{11}$$

where ϕ_{pw} is the polar angle of the planewave. Here in our multizone case, $\alpha_m^{(d)}(k)$ is defined as the set of linear combined coefficients for the desired multizone soundfield based on the coefficient set $\{P\}$:

$$\alpha_m^{(d)}(k) = \sum_j P_j \alpha_m^j(k), \tag{12}$$

where $\alpha_m^j(k)$ represents the desired soundfield coefficients for the jth planewave function we introduced.

In our scenario we have a constraint on the aperture function. It is then natural to find the coefficients $\{\beta_m(k)\}$ that minimize the error

$$\Gamma(\{\beta_m(k)\}) = \sum_{m=-M}^{M} |\beta_m(k) - \beta_m^0(k)|^2 \tag{13}$$

subject to an inequality constraint on the power of the aperture function $\rho(\phi, k)$ for the angular window $0 \le \phi < \phi_0$ where the aperture function should be zero ($\phi_0 = \pi$ for a semi-circle arrangement):

$$\int_0^{\phi_0} |\sum_{m=-M}^{M} (\beta_m(k)e^{im\phi})|^2 d\phi \le \eta_c. \tag{14}$$

Note that the corresponding number of required loudspeakers for a angular window of ϕ_0 is $Q_0 = \lceil \frac{\phi_0(2M+1)}{2\pi} \rceil$.

We solve the optimization problem using Lagrange multipliers. Assuming we are at the inequality bound, we minimize

$$\eta_0 = \Gamma(\{\beta_m(k)\}) + \lambda\eta_c, \tag{15}$$

where the multiplier $\lambda \le 0$. From an alternative viewpoint, it can be seen that (15) defines a weighting between the constraint and the function Γ that is determined by λ. Let $\mathbf{e_m} = [e^{-iM\phi}, \ldots, e^{iM\phi}]$. Then, the Fourier coefficients that optimize (15) are

$$\beta^{\mathbf{d}} = (\mathbf{I} - \lambda \int_0^{\phi_0} (\mathbf{e_m}^H \mathbf{e_m}) d\phi)^{-1} \beta_{\mathbf{0}}. \tag{16}$$

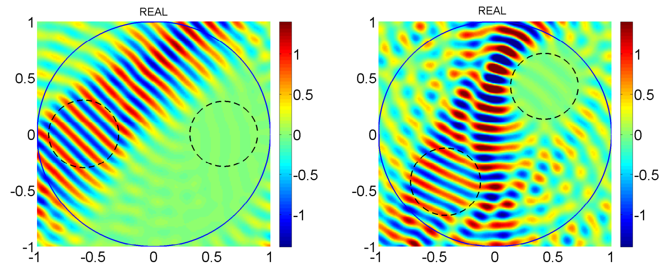
The solution of $\beta^{\mathbf{d}} = [\beta_{-M}^d(k), \ldots, \beta_M^d(k)]^T$ can then be used to calculate the qth loudspeaker weights $l_q(k)$ in the desired non-zero aperture region required for approximately reproducing the desired soundfield [13]:

$$l_q(k) = \sum_{m=-M}^{M} \beta_m^d(k) e^{im\phi_q} \Delta\phi_s. \tag{17}$$

where $\Delta\phi_s = (2\pi - \phi_0)/Q_0$ is the angular spacing of the loudspeakers and $\phi_q = \phi_0 + (q-1)\Delta\phi_s$ is the angle of the qth loudspeaker from x-axis. Let $S_{disc}^a(\mathbf{x}, k)$ be the reproduced soundfield using the part-circle method with weights provided in (17). Then

$$S_{disc}^a(\mathbf{x}, k) = \sum_q l_q(k) \frac{i}{4} H_0^{(1)}(k\|R_l\hat{\phi}_{\mathbf{q}} - \mathbf{x}\|) \tag{18}$$

where R_l is the radius of the part-circle where the loudspeakers are located.



(a) (b)

Fig. 2. Describing the desired multizone soundfields with a basis expansion. (a) and (b) are for the case when $\phi_d = 45°$ and $\phi_d = 60°$ respectively.

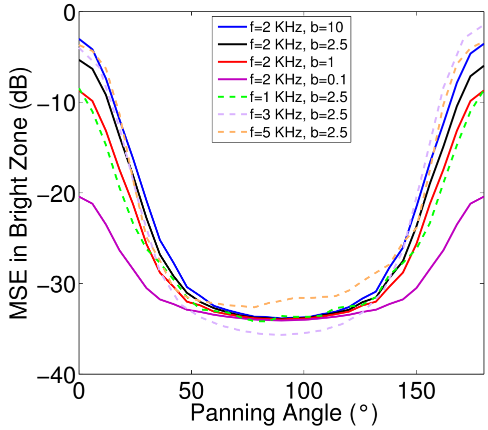
5. RESULTS

We consider a multizone reproduction with a bright zone and a quiet zone, each of radius 0.3m within a $\mathbb{D}$ of radius $r = 1m$ at a frequency of $f = 2$ KHz. The centers of $\mathbb{D}_b$ and $\mathbb{D}_q$ lie on a circle of radius $d = 0.6$m. The target bright and quiet zones are located at ϕ_1 and ϕ_2 respectively as shown in Fig. 1. We set the weighting function $w(\mathbf{x})$ as: a=1, b=2.5 and c=0.05. We attempt to recreate a planewave arriving from angle ϕ_d in $\mathbb{D}_b$, while also attenuating the sound in $\mathbb{D}_q$.

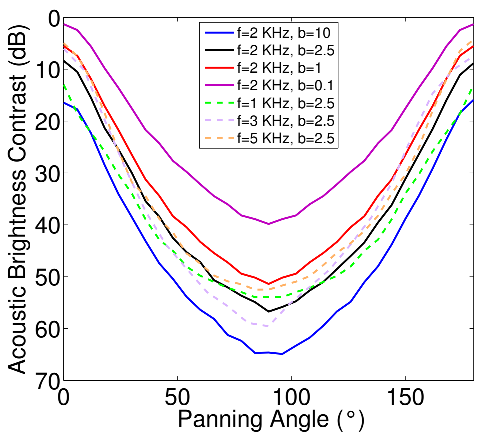
We start with simulations in Fig. 2 illustrating how the desired multizone soundfield is described by a feasible Helmholtz solution using the method introduced in Sec. 3. We set $\Delta\phi = \pi/40$ in (7) and $N = 80$. In Fig. 2(a), $\phi_1 = 180°$ and $\phi_2 = 0°$, while a planewave at $\phi_d = 45°$ is desired for $\mathbb{D}_b$. The desired quiet and unattended zones are set to zero. The synthesized multizone soundfield features an acoustic contrast of 40.5dB between $\mathbb{D}_b$ and $\mathbb{D}_q$, while the MSE of reproduction in $\mathbb{D}_b$ is -29.7 dB. In Fig. 2(b), we have $\phi_d = 60°$, $\phi_1 = 225°$ and $\phi_1 = 45°$. This more challenging multizone reproduction scenario is summarized as a "non-robustness" problem in [5] due to the occlusions of sound between $\mathbb{D}_b$ and $\mathbb{D}_q$. We achieved an acoustic contrast between the zone of $\mathbb{D}_b$ and $\mathbb{D}_q$ of 21.4 dB, while still maintaining an accurate reproduction of the desired planewave in $\mathbb{D}_b$ with a MSE of -13.8 dB.

Fig. 3 studies system performance as the planewave angle ϕ_d over the bright zone in Fig. 2(a) is panned. We can observe that the panning angle of the desired planewave affects the system performance and that the worst performance is achieved when the planewave is in-line with both zones, while the best performance is obtained when it is perpendicular with the line drawn through the centres of the zones. We vary the weighting parameter $b \in \{10, 2.5, 1, 0.1\}$ for $\mathbb{D}_q$ (we assign a fixed value of $a = 1$ to $\mathbb{D}_b$ and $c = 0.05$ to the unattended zone) to investigate the design trade-off between the MSE of reproduction in $\mathbb{D}_b$ and the acoustic contrast. The results show that as we increase the relative reproduction importance over $\mathbb{D}_q$, it improves the acoustic contrast between the two zones at the expense of a decline in the reproduction accuracy over $\mathbb{D}_b$. In addition, we also investigate the performance of our approach at different frequencies: 1 KHz, 3 KHz and 5 KHz.

Fig. 4 demonstrates the desired multizone reproduction of the above-mentioned two scenarios using the approach of part-circle with an angular window of $\phi_0 = \pi$ at the frequency of 2000 Hz while $R = 1.5$m. We set the Lagrangian $\lambda = 10$ in (15). Overall, the number of the employed loudspeakers is 39 and only the lower part of loudspeakers are used, while a full circular array of 77

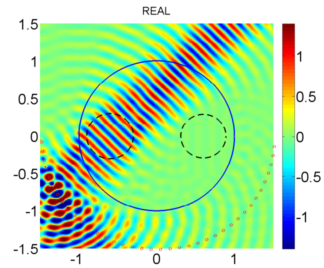


(a)

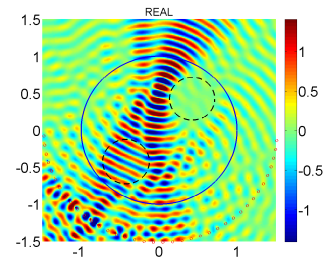


(b)

Fig. 3. Performance as the desired planewave is panned, comparing reproductions as parameters b are varied (solid curves), at different frequencies (dashed). Plotted are (a) the MSE in $\mathbb{D}_b$ and (b) acoustic contrast between the bright and quiet zone.



(a)



(b)

Fig. 4. Desired multizone reproduction using the approach of part-circle with an angular window of $\phi_0 = \pi$. 39 loudspeakers are used and the red circles demonstrate the positions of loudspeakers. (a) and (b) represent the cases with $\phi_d = 45^\circ$ and $\phi_d = 60^\circ$ respectively.

loudspeakers is required using the existing reproduction method proposed in [13]. Note that we merely adopt the orthogonal set which consists of basis planewaves arriving from 0 to π. The rationale of doing this is that physically we are not able to render sound waves travel towards the semicircle of loudspeakers and the introduction of the other half set of basis planewaves would lead to large reproduction errors overall.

We can observe that the reproduced multizone soundfields in Fig. 4 correspond well to the desired sound over $\mathbb{D}$. For the first case with a $\phi_d = 45^\circ$, the acoustic contrast between the two zones is 34.7 dB while the MSE in $\mathbb{D}_b$ is -26.8 dB. For the other case with a $\phi_d = 60^\circ$, the acoustic contrast drops to 19.5 dB while the MSE in the bright zone is now -13.7 dB. By using the semi-circle method, we achieve the same reproduction accuracy over $\mathbb{D}_b$ compared with the results using the reproduction method with a full continuous aperture function [13] in the second step. We note a reasonable decline in terms of the acoustic contrast between the two zones with the semi-circle method, which is due to the decreased number of employed loudspeakers, as well as the input array effort [3].

6. CONCLUSION

We proposed a new approach for describing a desired 2-D multizone soundfield using an orthogonal expansion based on a weighted inner product over the desired reproduction region. The method enables us to tailor the overall system performance by adjusting the parameters in the weighting function. The approach was shown to be effective both in controlling the rendering accuracy in the bright zone and the leakage into the quiet zone, particularly when occlusions occur between the two zones. The approach was formulated to eliminate the ill-conditioning problems due to the positioning of the selected zones. We used a generalization of the reproduction method of [13]. This approach allows us to diminish the number of employed loudspeakers and obtain a more practical configuration.

7. REFERENCES

[1] J. W. Choi and Y. H. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources," *J. Acoustics Society America*, vol. 111, pp. 1695–1700, 2002.

[2] 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. Audio Engineering Society,*, vol. 125, pp. 2091–2097, Apr. 2009.

[3] S. J. Elliott, J. Cheer, J-W. Choi, and Y. Kim, "Robustness and regularization of personal audio systems," *IEEE Trans. Speech, Audio and Language Process.*, vol. 20, pp. 2123–2133, Sept. 2012.

[4] J. Daniel, R. Nicol, and S. Moreau, "Further investigations of high order ambisonics and wavefield synthesis for holophonic sound imaging," in *Proc. Audio Eng. Society 114th Convention*, 2003, vol. 51, p. 425.

[5] M. Poletti, "An investigation of 2D multizone surround sound system," in *Proc. Audio Eng. Society 125th Convention*, 2008, vol. 51, p. 425.

[6] N. Radmanesh and I. S. Burnett, "Reproduction of independent narrowband soundfields in a multizone surround system and its extension to speech signal sources," in *Proc. IEEE Int Conf. Acoust. Speech Signal Process.*, 2011, vol. 11, pp. 598–610.

[7] Y. J. Wu and T. D. Abhayapala, "Spatial multizone soundfield reproduction: Theory and design," *IEEE Trans. Acoust.,Speech, Signal Processing*, vol. 19, no. 6, pp. 1711–1720, August 2011.

[8] Y. J. Wu and T. Abhayapala, "Multizone 2D soundfield reproduction via spatial band stop filters," in *Workshop. Applications Signal Process. Audio Acoust.* IEEE, 2009, pp. 309–312.

[9] T. Betlehem and P. Teal, "A constrained optimization approach for multizone surround sound," in *Proc. IEEE Int Conf. Acoust. Speech Signal Process.*, 2011, pp. 437–440.

[10] F. M. Fazi and P. A. Nelson, "The ill-conditioning problem in sound field reconstruction," in *Proc. Audio Eng. Society 123rd Convention*, 2007.

[11] P. D. Teal, T. Betlehem, and M. Poletti, "An algorithm for power constrained holographic reproduction of sound," in *Proc. IEEE Int Conf. Acoust. Speech Signal Process.*, March 2010, pp. 101–104.

[12] D. B. Ward and T. D. Abhayapala, "Reproduction of the plane wave sound field using an array of loudspeakers," *IEEE Trans. Acoust.,Speech, Signal Processing*, vol. 9, no. 6, pp. 697–707, Sep. 2001.

[13] Y. J. Wu and T. D. Abhayapala, "Theory and design of soundfield reproduction using continuous loudspeaker concept," *IEEE Trans. Acoust.,Speech, Signal Processing*, vol. 17, no. 1, pp. 107–116, January 2009.

[14] T. Betlehem and C. Withers, "Sound field reproduction with energy constraint on loudspeaker weights," *IEEE Trans. Acoust.,Speech, Signal Processing*, vol. 20, no. 8, pp. 2388–2392, Oct. 2012.

[15] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*, Academic, New York, 1999.

[16] G. H. Golub and C. Van Loan, *Matrix Computation*, Johns Hopkins Univ. Press, 3rd edition, Oct. 1996.

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