

3D Multizone Soundfield Reproduction in the Reverberant Room Using a Spherical Loudspeaker Array

Meng-fang Zha, Chang-chun Bao and Mao-shen Jia

Speech and Audio Signal Processing Laboratory, School of Electronic Information and Control Engineering,
Beijing University of Technology, Beijing 100124, China

E-mail: zhamengfang@emails.bjut.edu.cn, chchbao@bjut.edu.cn, jiamashen@bjut.edu.cn,

Abstract—Multizone soundfield reproduction in the reverberant room is a complex problem in acoustic signal processing. In this paper, we introduce a method of recreating a 3D multizone soundfield for reverberant environments. Based on spherical addition theorem for Bessel function, we derive the desired global soundfield by translating local desired soundfields to a single global coordinate system. We cast the global soundfield reproduction into a least squares framework and introduce the modal space approach to derive the loudspeaker weights. A practical method is used to determining the Acoustic Transfer Function (ATF) between each loudspeaker and every point of reproduction region. The simulation results indicate that the proposed method can reproduce the target 3D multizone soundfield accurately.

Index Terms—Acoustic signal processing, 3D multizone soundfield, reverberant, Acoustic Transfer Function

I. INTRODUCTION

Multizone soundfield reproduction has various applications such as entertainment systems in cars and surround sound systems in exhibition centers, where the aim is to provide listeners individual sound environment without physical isolations or use of headphones. However, in practical acoustical environments, the performance of multizone soundfield reproduction techniques is significantly degraded by the reverberation in the room.

The origin of work in this area can be traced to 2008, when Poletti proposed least squares pressure matching method to generate 2D multizone surround sound systems [1]. However, the investigation was mainly performed based on simulation results. Further investigation was made in [2], where a constrained optimization problem was formulated to control the sound in each zone individually. A method for using spatial band stop filters to suppress the unintended inter-zone sound interference and pass the desired soundfields with no distortion was introduced in [3]. The desired 2D multizone soundfield was described as an orthogonal expansion of basis functions over the reproduction region in [4].

However, the aforementioned approaches do not take into account the reverberant environments in multizone soundfield reproduction. The reverberant case is made difficult by the rapid variation of the ATF over the room [5]. A common approach for spatial sound reproduction in reverberant setting is to design some inverse filters for equalizing the ATF over

multiple points [6]. However, the resulting room correction of this technique is poor away from the design points. Alternative dereverberation techniques are the measurement of the ATF at a finite set of sample points which are later incorporated to the sound processing algorithm [7]-[8].

The techniques with regard to single zone soundfield reproduction in the reverberant room were proposed in [9]-[10]. A method of parameterizing the RTF between two arbitrary points from a predefined source and receiver regions was introduced in [9]. In [10], a practical method was provided for determining the ATF between each loudspeaker and every point in the reproduction region. This work was extended to the 2D multizone sound case in [11] using the translation of spatial harmonic coefficient between coordinate systems to achieve a global control of the multizone soundfield. In addition, a method of 2D multizone soundfield reproduction in reverberant rooms using compressed sensing technique was proposed in [12].

In this paper, we propose a method for reproducing a desired 3D multizone soundfield in reverberant environments using a single spherical loudspeaker array. First, the desired local soundfield coefficients are mapped into the equivalent global soundfield coefficients based on the spherical addition theorem of Bessel function [13]. Thus, the complicated multizone soundfield reproduction problem can be translated into a reproduction of single zone soundfield. Then, the optimal least-squares solution for the loudspeaker weights that minimize the reproduction error are derived based on the modal space approach [14]. We also describe a practical method to precisely measure the ATFs between each loudspeaker and the reproduction region. Simulation results show that the proposed approach can reproduce the target 3D spatial multizone soundfield in a reverberant room accurately.

II. MULTIZONE SOUNDFIELD'S GLOBAL CONTROL

A. Spherical Harmonic Expansion of Soundfield

As illustrated in Figure 1, the desired reproduction region Ω is the smallest sphere that encloses all spatial zones of interest with a radius of R . We first assume that there are Q non-overlapping 3D spherical zones at radius $R^{(q)}$ from the global spherical origin O . The origin and the radius of the q^{th}

zone are denoted as O_q and $R_z^{(q)}$, where O_q is located at the polar coordinate $\mathbf{R}^{(q0)} = (R^{(q0)}, \theta^{(q0)}, \phi^{(q0)})$ with respect to O . The arbitrary observation point within the q^{th} zone is denoted as $\mathbf{x}^{(q)} = (r^{(q)}, \theta^{(q)}, \phi^{(q)})$ with respect to the local origin O_q .

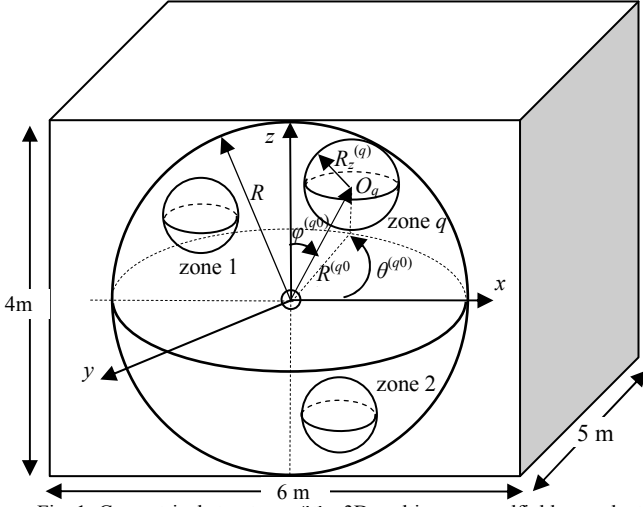


Fig. 1. Geometrical structure of the 3D multizone soundfield reproduction system in a reverberant room.

The desired soundfield of the q^{th} zone can be represented by the following spherical harmonic expansion representation [14]

$$S^{d(q)}(\mathbf{x}^{(q)}; k) = \sum_{v=0}^{V_q} \sum_{u=-v}^v \beta_{vu}^{d(q)}(k) j_v(kr^{(q)}) Y_{vu}(\theta^{(q)}, \phi^{(q)}) \quad (1)$$

where $k = 2\pi f/c$ is the wave number, f is the frequency, c is the speed of sound propagation, $j_v(kr^{(q)})$ represents the spherical Bessel function of order v , $\beta_{vu}^{d(q)}(k)$ are a set of coefficients uniquely representing the q^{th} soundfield, which are mode limited to $V_q = \lceil kR_z^{(q)} \rceil$, the superscript “d” represents the desired soundfield, and $Y_{vu}(\theta^{(q)}, \phi^{(q)})$ denotes the spherical harmonic of order v and degree u .

We attempt to find an equivalent global soundfield which consists of Q non-overlapping individual multizone soundfield. Thus, we can reproduce the global desired soundfield over the entire region instead of multiple local spatial soundfield. Let the set of coefficient $\alpha_{nm}^d(k)$ represent the desired global soundfield. The sound pressure at an arbitrary observation point $\mathbf{x} = (r, \theta, \phi)$ inside the entire region can be given with respect to the global origin O as

$$S^d(\mathbf{x}; k) = \sum_{n=0}^N \sum_{m=-n}^n \alpha_{nm}^d(k) j_n(kr) Y_{nm}(\theta, \phi) \quad (2)$$

where $\alpha_{nm}^d(k)$ are mode limited to $N = \lceil kR \rceil$.

Similarly, we write the reproduced sound pressure as

$$S^a(\mathbf{x}; k) = \sum_{n=0}^N \sum_{m=-n}^n \alpha_{nm}^a(k) j_n(kr) Y_{nm}(\theta, \phi) \quad (3)$$

where $\alpha_{nm}^a(k)$ are the coefficients of the reproduced soundfield and the superscript “a” represents the actual soundfield.

B. Translation of Spatial Harmonic Coefficients

The relationship between the desired local soundfield coefficients $\beta_{vu}^{d(q)}(k)$ and the equivalent global soundfield coefficients $\alpha_{nm}^d(k)$ can be derived based on the spherical addition theorem for Bessel function [13].

$$\beta_{vu}^{d(q)}(k) = \sum_{n=0}^N \sum_{m=-n}^n \alpha_{nm}^d(k) \hat{S}_{nv}^{mu}(\mathbf{R}^{(q0)}) \quad (4)$$

where

$$\hat{S}_{nv}^{mu}(\mathbf{R}^{(q0)}) = 4\pi i^{v-n} \sum_{l=0}^{\infty} i^l (-1)^{2m-u} j_l(kR^{(q0)}) Y_{l(u-m)}^*(\theta^{(q0)}, \phi^{(q0)}) \times \sqrt{(2n+1)(2v+1)(2l+1)/4\pi W_1 W_2}$$

with

$$W_1 = \begin{pmatrix} n & v & l \\ 0 & 0 & 0 \end{pmatrix} \text{ and } W_2 = \begin{pmatrix} n & v & l \\ m & -u & u-m \end{pmatrix}$$

represent Wigner 3-j symbols.

III. SOUNDFIELD REPRODUCTION

After obtaining the desired soundfield coefficients for the global region, the multizone soundfield reproduction problem can be reduced to reproduction of the soundfield (2) over the entire regions. In this section, we introduce a modal space approach [14], which uses an efficient parametrization for ATFs to derive the loudspeaker weights.

A. Modal Space Approach

We assume all loudspeakers are placed on the sphere of radius R_l with respect to the global spherical origin O . Each loudspeaker l transmits an output signal $G_l(k)$. We define the acoustic transfer function $H_l(\mathbf{x}; k)$ as the frequency response between loudspeaker l and point $\mathbf{x}$. The pressure produced by the loudspeakers at any point $\mathbf{x}$ can be written as [13]

$$S^a(\mathbf{x}; k) = \sum_{l=1}^L G_l(k) H_l(\mathbf{x}; k) \quad (5)$$

Since the room response is equal to the sound pressure generated by the unit input signal $G_l(k)$, we can also write it as

$$H_l(r, \theta, \phi; k) = \sum_{n=0}^N \sum_{m=-n}^n \gamma_{nm}(l, k) j_n(kr) Y_{nm}(\theta, \phi) \quad (6)$$

where $\gamma_{nm}(l, k)$ are the modal coefficients of the ATF for loudspeaker l , which can be measured by microphone recordings. Next we describe a method of measuring $\gamma_{nm}(l, k)$.

The sound pressure inside Ω generated by loudspeaker l outside Ω in a reverberant enclosure as the modal expansion is

$$S(\mathbf{x}; k) = \sum_{n=0}^N \sum_{m=-n}^n \gamma_{nm}(l, k) j_n(kr) Y_{nm}(\theta, \phi) \quad (7)$$

As the sound pressure can be recorded by M microphone, the soundfield coefficients are calculated by,

$$\gamma_{nm}(l, k) = \frac{1}{j_n(kr)} \sum_{i=1}^M S(\mathbf{x}_i; k) [Y_{nm}(\theta_i, \phi_i)]^* \Delta s_i \quad (8)$$

where $\Delta s_i = 4\pi/M$.

Substituting (3) and (6) into (5), the modal coefficients of the reproduced soundfield are related to $\gamma_{nm}(l, k)$ through

$$\alpha_{nm}^a(k) = \sum_{l=1}^L G_l(k) \gamma_{nm}(l, k) \quad (9)$$

The design task of our 3D soundfield reproduction is to choose the loudspeaker weights $G_l(k)$ to minimize the normalized reproduction error η ,

$$\eta = \frac{\int_0^{2\pi} \int_0^\pi \int_0^R |S^d(\mathbf{x}, k) - S^a(\mathbf{x}, k)|^2 r^2 \sin \theta dr d\theta d\phi}{\mathcal{E}} \quad (10)$$

where the energy of the desired soundfield $\mathcal{E}$ is:

$$\mathcal{E} = \int_0^{2\pi} \int_0^\pi \int_0^R |S^d(\mathbf{x}, k)|^2 r^2 \sin \theta dr d\theta d\phi$$

An expression for the normalized error as a function of the modal coefficients can be derived

$$\eta = \frac{1}{k^3 \mathcal{E}} \sum_{n=0}^N \sum_{m=-n}^n w_n(kR) |\alpha_{nm}^a(k) - \alpha_{nm}^d(k)|^2 \quad (11)$$

$$\mathcal{E} = \frac{1}{k^3} \sum_{n=0}^N \sum_{m=-n}^n w_n(kR) |\alpha_{nm}^d(k)|^2,$$

where the modal coefficient weighting function is given by

$$w_n(kR) = \int_0^{kR} [j_n(x)]^2 x^2 dx \quad (12)$$

B. Least Squares Solution

We now derive the least squares solution for the loudspeaker weights that minimizes the error in (10). Defining the vector of loudspeaker weights $\mathbf{g} = [G_1(k), G_2(k), \dots, G_L(k)]^T$, where $[\cdot]^T$ is the matrix transpose operator, the vector of the coefficients of the reproduced soundfield $\mathbf{A}^a = [\alpha_{00}^a, \dots, \alpha_{NN}^a]^T$, and the matrix of the modal coefficients for the ATF is

$$\boldsymbol{\gamma} = \begin{bmatrix} \gamma_{00}(1, k) & \cdots & \gamma_{00}(L, k) \\ \vdots & \ddots & \vdots \\ \gamma_{NN}(1, k) & \cdots & \gamma_{NN}(L, k) \end{bmatrix} \quad (13)$$

Thus, (9) can be rewritten as $\mathbf{A}^a = \boldsymbol{\gamma} \mathbf{g}$. Additionally, define the vectors of the coefficients of the desired soundfield, $\mathbf{A}^d = [\alpha_{00}^d, \dots, \alpha_{NN}^d]^T$, and the diagonal weighting matrix

$$\mathbf{W} = \begin{bmatrix} w_0(kR) \mathbf{I}_1 & \mathbf{O}_{1(3)} & \cdots & \mathbf{O}_{1(2N+1)} \\ \mathbf{O}_{3(1)} & w_1(kR) \mathbf{I}_3 & \cdots & \mathbf{O}_{3(2N+1)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O}_{(2N+1)(1)} & \mathbf{O}_{(2N+1)(2)} & \cdots & w_N(kR) \mathbf{I}_{2N+1} \end{bmatrix}$$

where $\mathbf{O}_{nm}$ is the $n \times m$ zero matrix.

We rewrite (10) in terms of matrix as

$$\eta = \frac{(\mathbf{A}^a - \mathbf{A}^d)^H \mathbf{W} (\mathbf{A}^a - \mathbf{A}^d)}{(\mathbf{A}^d)^H \mathbf{W} \mathbf{A}^d} \quad (14)$$

where $(\cdot)^H$ is the matrix Hermitian operator.

Since $\mathbf{A}^a = \boldsymbol{\gamma} \mathbf{g}$, (14) is written as a quadratic form in the vector $\mathbf{g}$ of loudspeaker weights

$$\eta = \frac{1}{d} (\mathbf{g}^H \mathbf{B} \mathbf{g} - \mathbf{b}^H \mathbf{g} - \mathbf{g}^H \mathbf{b} + d) \quad (15)$$

where $\mathbf{B} = \boldsymbol{\gamma}^H \mathbf{W} \boldsymbol{\gamma}$, $\mathbf{b} = \boldsymbol{\gamma}^H \mathbf{W} \mathbf{A}^d$ and $d = (\mathbf{A}^d)^H \mathbf{W} \mathbf{A}^d$.

The least squares solution to (15) is $\mathbf{g} = \mathbf{B}^{-1} \mathbf{b}$.

IV. SIMULATION RESULTS

In the following example, we illustrate reproduction of a plane wave originating from direction $(\theta^{(pw)}, \phi^{(pw)})$, the desired soundfield coefficients are given by [13]

$$\beta_{vu}^{(q)}(k) = 4\pi i^v Y_{vu}(\theta^{(pw)}, \phi^{(pw)})^* \quad (16)$$

We consider a multizone reproduction in $Q = 2$ zones with $R_z^{(q)} = 0.4$ m, a zone radius assumed to be sufficient for a single listener. Zone 1 and zone 2 are both 1.3 m away from the global spherical origin O . Source 1 is a plane wave originating from $[\theta^{(pw)}, \phi^{(pw)}] = [0^\circ, 0^\circ]$ at a frequency of 500 Hz, and the target zone for this signal is zone 1. Zone 2 is set to be a silent zone. As shown in Fig. 1, the reverberant room is rectangular (size $6 \times 5 \times 4$ m) with a wall absorption coefficient of 0.51. The number of loudspeakers, L is equal to 169 and the radius of this spherical loudspeaker array is 2 m. To measure the modal coefficients for the ATF, we use $M = 64$ pressure samples on a spherical of 0.4 m.

The reverberation is generated with a 3D adaptation of the image-source method [15] including image-sources up to the fifth order, which defines the ATF in terms of

$$H(\mathbf{x}; k) = \frac{ie^{-ik\|\mathbf{x}-\mathbf{y}\|}}{k\|\mathbf{x}-\mathbf{y}\|} + \sum_{p=1}^P \varsigma_p \frac{ie^{-ik\|\mathbf{x}-\mathbf{y}_p\|}}{k\|\mathbf{x}-\mathbf{y}_p\|} \quad (17)$$

where y_p and ς_p are the position and accumulated wall reflection coefficient of the p^{th} image source.

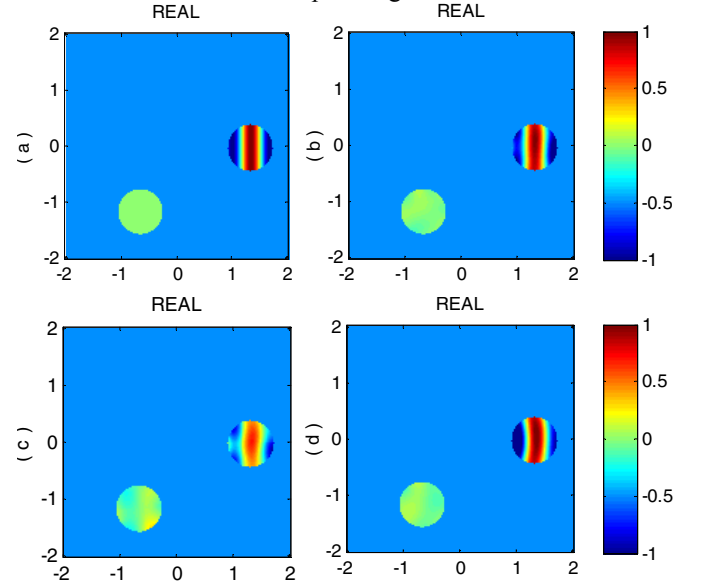


Fig. 2. Reproduction of a 3D two-zone soundfield with radius 0.4 m in a reverberant room. Two zones are located at $R^{(10)} = 1.3$ m, $\theta^{(10)} = 0^\circ$, $\phi^{(10)} = 0^\circ$ and $R^{(20)} = 1.3$ m, $\theta^{(20)} = 240^\circ$, $\phi^{(20)} = 0^\circ$, respectively. (a) Desired field, (b) reproduced field in a free field, (c) the same free field design in the reverberant room, and (d) reproduced field with the reverberation design.

The resulting reproduced field is shown in Fig. 2, displayed as slices through the sphere at height of 0 m. The top-left plot (a) shows the real part of the desired soundfield, the top-right plot (b) shows the real part of the reproduced soundfield in a free field, the bottom-left plot (c) shows the real part of the same free design in the reverberant room, and the bottom-

right plot (d) shows the real part of the reproduced field with the reverberant design. The unattended region is set to -0.5 in order to observe the performance of soundfield reproduction better.

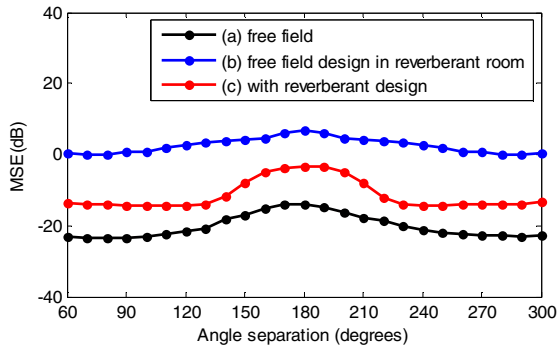


Fig. 3. Mean square error for two zones as function of different zone separation (in angles) $\theta^{(20)} - \theta^{(10)}$.

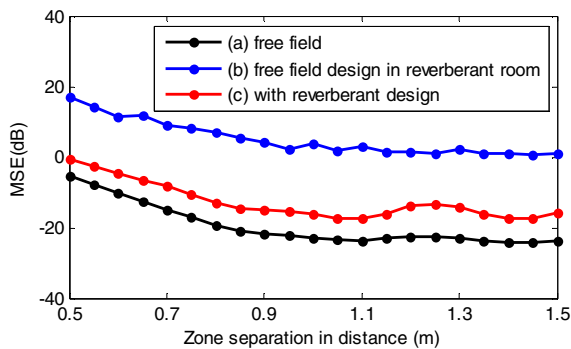


Fig. 4. Mean square error for two reproduction zones as function of different zone separation (in distance).

Fig. 3 illustrates the reproduction error (10) as a function of angle difference between zone 1 and zone 2. The angle difference between two reproduction zones is varied by changing their azimuth angle and making the elevation angle constant. We set $60^\circ < (\theta^{(20)} - \theta^{(10)}) < 300^\circ$ in order to reproduce two non-overlapping soundfield. The results show that the angle separation of two zones affects the system performance and the worst performance is achieved when the plane wave is in line with both zones.

Fig. 4 demonstrates the reproduction error as a function of the distance between the two zones. We can observe that the reproduction error decreases when zones are further apart. However, the error will not reduce once the zone separation beyond a certain threshold. In addition, we also provide the error with a free field design and the same free field design in the reverberant room. The results show that the reverberant performance of the free field design is poor, and the reverberant design performs well.

V. CONCLUSIONS

In this paper, we propose a new method for recreating a 3D multizone soundfield in reverberant rooms using a single spherical loudspeaker array. We present a novel approach to derive the loudspeaker weights for the desired multizone sound reproduction. We also investigate the reproduction

error in terms of the angle and distance difference between the zones. The simulation results confirm the accurate spatial multizone soundfield reproduction in a reverberant room. Reducing the number of loudspeakers required for 3D multizone soundfield reproduction shall be addressed in future research.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (Grant No. 61231015, 61201197), Scientific Research Project of Beijing Educational Committee (Grant No. KM201310005008) and the Ph.D. Programs Foundation of Ministry of Education of China (Grant No. 20121103120017).

REFERENCES

- [1] M. A. Poletti, "An investigation of 2D multizone surround sound system," in *Proc. Audio Eng. Society (AES) 125th Convention*, vol. 51, pp. 425, Oct. 2008.
- [2] T. Betlehem and P. Teal, "A constrained optimization approach for multizone surround sound," in *ICASSP*, pp. 437–440, 2011.
- [3] Y. J. Wu and T. D. Abhayapala, "Multizone 2D soundfield reproduction via spatial band stop filters," in *Workshop. Applications Signal Process. Audio Acoust. IEEE*, pp. 309–312, 2009.
- [4] W. Jin, W. B. Kleijn and D. Virette, "Multizone soundfield reproduction using orthogonal basis expansion," in *ICASSP*, pp. 311–315, May 2013.
- [5] J. Mourjopoulos, "On the variation and invertibility of room impulse response function," *J. Sound Vib.*, vol. 102, 1985.
- [6] P. A. Nelson, F. Orduna-Bustamante and H. Hamada, "Inverse filter design and equalization zones in multichannel sound reproduction," *IEEE Transactions on Speech and Audio Processing*, vol. 3, no. 3, pp. 185–192, 1995.
- [7] S. Bharitkar, "A cluster centroid method for room response equalization at multiple locations," in *Workshop. Applications Signal Process. Audio Acoust. IEEE*, pp. 55–58, 2001.
- [8] S. Tervo, J. Patynen, A. Kuusinen and T. Lokki, "Spatial decomposition method for room impulse responses," *Journal of the Audio Engineering Society*, vol. 61, pp. 17–28, 2013.
- [9] P. N. Samarasinghe, T. D. Abhayapala and M. A. Poletti, "On room impulse response between arbitrary points: An efficient parameterization," in *IEEE International Symposium on Communications, Control, and Signal Processing (ISCCSP)*, 2014.
- [10] T. Betlehem and T. D. Abhayapala, "Theory and design of sound field reproduction in reverberant rooms," *J. Acoust. Soc. Amer.*, vol. 117, pp. 2100–2111, Apr. 2005.
- [11] Y. J. Wu and T. D. Abhayapala, "Spatial multizone soundfield reproduction: Theory and design," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 19, pp. 1711–1720, Aug. 2011.
- [12] W. Jin and W. B. Kleijn, "Multizone soundfield reproduction in reverberant rooms using compressed sensing techniques," in *ICASSP*, pp. 4728–4732, 2014.
- [13] M. F. Zha, C. C. Bao and M. S. Jia, "3D multizone soundfield reproduction using spherical harmonic analysis," in *ChinaSIP*, pp. 625–629, 2015.
- [14] T. Betlehem, "Acoustic signal processing algorithms for reverberant environments," *PhD thesis, Australian National University*, pp. 122–126, 2005.
- [15] J. Allen and D. Berkley, "Image method for efficiently simulating small room acoustics," *J. Acoust. Soc. Amer.*, vol. 65, pp. 943–950, 1979.