A DOWN-MIXING METHOD FOR 22.2 MULTICHANNEL SYSTEM REPRODUCTION

Song Wang, Ruimin Hu, Shihong Chen, Xiaochen Wang, Yuhong Yang, Weiping Tu

National Engineering Research Center for Multimedia Software,
School of Computer Science, Wuhan University, Wuhan, 430072, China
Research Institute of Wuhan University in Shenzhen, Shenzhen, 518000, China

ABSTRACT

This paper proposes a general multichannel system reproduction method. Firstly, relative to original multichannel system, a general global model is build up by guaranteeing sound pressure and the direction of particle velocity at the receiving point constant, and making the square error of particle velocity magnitude at the receiving point as little as possible. Then the model is equivalent to a least squares problems with non-negative constraints, it can be worked out by existing mature algorithms, and the global optimal solution of simplifying multichannel system are obtained. The proposed method can be used to simplify 22.2 multichannel system to 10.2 and 8.2 multichannel system, objective and subjective experimental results demonstrate that it performs better than traditional method.

Index Terms— Multichannel system, reproduction, global optimal solution

1. INTRODUCTION

As the development of 3D movie and 3DTV technology, 3D audio technology has becomes a research hotspot in the field of multimedia. There are following 3D audio technologies: Ambisonics [1, 2, 3, 4], Wave Field Synthesis (WFS) [5, 6, 7], Head Related Transfer Function (HRTF) [8, 9, 10], Vector Based Amplitude Panning (VBAP) [11, 12] and so on. Among them, VBAP is an important 3D audio technology.

In 1961, B. B. Bauer proposed the stereophonic law of sines which is followed by the directional perception of a virtual sound source produced by amplitude panning [13]. The stereophonic law of sines is valid if the listener's head is pointing forward. To overcome this constraint, B.Bernfeld proposed the tangent law [14] in 1973. VBAP is seen as an extension of the sine law with three dimensional loudspeaker layouts [14]. In three dimensional VBAP, a virtual sound source is synthesized by three loudspeakers using vector view [11, 12]. VBAP is of high computational efficiency and accuracy sound image reconstruction. VBAP is more flexible than Ambisonics, because loudspeakers can be freely placed; VBAP is cheaper than WFS, because it uses less loudspeakers; computational efficiency of VBAP is higher than that of HRTF 3D positioning method. Also, [15, 16] has generalized VBAP to synthesize a virtual source by four and five loudspeakers.

In 2011, Akio Ando proposed a method that provides physical underpinnings for the use of VBAP, and it can convert 22-channel signals of original 22.2 multichannel sound system without the two low-frequency effect channels into 10-, 8-channel signals [17], and perform better than conventional down-mixing method. This method localizes each channel signal of 22.2 multichannel system at the corresponding loudspeaker position as a virtual sound source. A virtual sound source is replaced by three loudspeakers in reproduced multichannel system which can form the minimum spherical triangle that includes this virtual sound source. 10 or 8 loudspeakers in reproduced multichannel system don't join in replacing a virtual sound source all in each replacement. It replaces all loudspeakers in 22.2 multichannel system by loudspeakers of reproduced multichannel system. In each replacement process, the distribution coefficients of three loudspeakers in reproduced multichannel system are calculated by guaranteeing that sound pressure and the direction of particle velocity at the receiving point produced by three loudspeakers and a virtual sound source are the same, and local solution is obtained. All loudspeakers locate on a sphere, the receiving point is at the center of this sphere. Then the final distribution coefficient of each loudspeaker in reproduced system is obtained by adding the distribution coefficient of corresponding loudspeaker in each replacement process.

There are two problem about Ando's method. First, this method uses the local optimal solution to form the global solution simply. Though it can maintain sound pressure and the direction of particle velocity at the spherical center, but it would make the sound error at non spherical center point bigger. In realistic listening, listeners keep the center of their head coinciding to the position of spherical center, their ears locates around sphere center, because the radius of human head is about 0.085 m. Second, as we know that sound can

The research was supported by National Nature Science Foundation of China (61231015, 61201169, 61201340, 61102127, U1404618), the National High Technology Research and Development Program of China (863 Program) (SS2015AA012205).

be described by sound pressure and particle velocity [18], so sound pressure and particle velocity should be well maintained in simplifying multichannel system to have a good reproduction effect. Sound pressure is scalar, particle velocity is vector which is of direction and magnitude. But Ando's method misses the magnitude of particle velocity, and we will explain it detailedly in section 5 by example.

This paper proposes a method of simplifying 22.2 multichannel system which obtains the global optimal solution. Meanwhile, it can reduce the sound error around receiving point in reproduced multichannel system.

2. BASIC DEFINITION

Our method needs to meet the following assumptions: (1) reflected sound is neglected, (2) a loudspeaker can be seen as a point source, (3) only the outgoing sound wave of the loudspeaker is considered, (4) the sound pressure at a unit distance from a loudspeaker is in proportion to the input to the loudspeaker, the proportion coefficient is recorded as G, (5) $k\sigma_{min} \gg 1$, k is the wave number, σ_{min} is the minimum distance between the virtual sound source or the loudspeakers and the receiving point. Assumption (5) is valid except for low frequency sound which does not contribute to the perception of sound location. Based on these assumptions, Fourier transform of the sound pressure produced by single loudspeaker at the receiving point is expressed as:

$$p(\vec{r},\omega) = G\frac{e^{-ik|\vec{r}-\vec{\xi}|}}{|\vec{r}-\vec{\xi}|}s(\omega) \tag{1}$$

Particle velocity produced by single loudspeaker at the receiving point is expressed as:

$$\begin{aligned} u(\vec{r},\omega) &= G(1+\tfrac{1}{ik|\vec{r}-\vec{\xi}|})\tfrac{e^{-ik|\vec{r}-\vec{\xi}|}}{|\vec{r}-\vec{\xi}|^2}\begin{pmatrix} x-\xi_x \\ y-\xi_y \\ z-\xi_z \end{pmatrix} s(\omega) \\ &\approx G\tfrac{e^{-ik|\vec{r}-\vec{\xi}|}}{|\vec{r}-\vec{\xi}|^2}\begin{pmatrix} x-\xi_x \\ y-\xi_y \\ z-\xi_z \end{pmatrix} s(\omega) \end{aligned} \tag{2}$$

where $\vec{r}=(x,y,z)^T$ is the coordinate of the receiving point, $\vec{\xi}=(\xi_x,\xi_y,\xi_z)^T$ is the coordinate of single loudspeaker, k is the wave number, $k=2\pi f/c$, f is the frequency of sound, c is the speed of sound in air, $s(\omega)$ is the Fourier transform of the input signal to the loudspeaker, T is the transposition of the matrix.

The sound pressure or particle velocity of m loudspeakers at the receiving point could be obtained by summing sound pressure or particle velocity of single loudspeakers at the receiving point.

3. GLOBAL MODEL BUILDING AND SOLVING

22.2 multichannel loudspeaker arrangement without LFE channels is shown in Figure 1(a). We first study the case that

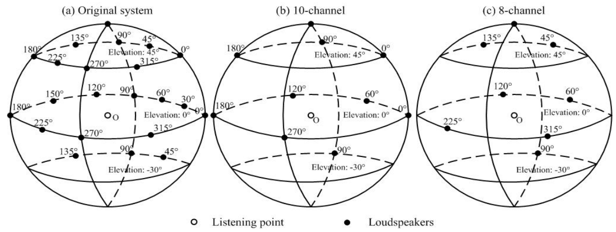


Fig. 1. (a): 22.2 multichannel loudspeaker arrangement without LFE channels in original system; (b), (c): 10-, 8-channel loudspeaker arrangements in reproduction space.

a virtual sound source (single loudspeaker) is replaced by m loudspeakers. The virtual sound source and m loudspeakers are on a same sphere, the receiving point is at the center O of the sphere which also is the listening point. We suppose that the sound pressure produced by a virtual sound source at the receiving point is the same as the sound pressure produced by m loudspeakers at the receiving point, we can get:

$$G\sum_{j=1}^{m}\frac{e^{-ik\rho}}{\rho}s_j(\omega) = G\frac{e^{-ik\rho}}{\rho}s(\omega) \tag{3}$$

Together with $s_j(\omega) = w_j s(\omega)$, equation (3) could be simplified as:

$$w_1 + w_2 + \cdots + w_m = 1 \tag{4}$$

In polar coordinates, coordinate of the receiving point is $\vec{r}=(0,0,0)$, coordinate of the virtual sound source is $\vec{\xi}=(\rho,\theta,\varphi)$, where ρ is the distance to receiving point, θ is azimuth, φ is elevation, coordinates of m loudspeakers are $\vec{\xi}^{(j)}=(\rho,\theta_j,\varphi_j)$, $j = 1,2,\cdots,m$. Then Fourier transform of the particle velocity produced by the virtual sound source at the receiving point is expressed as:

$$\begin{gathered} u(\vec{r},\omega) = -\tfrac{G}{c\lambda}\tfrac{e^{-ik\rho}}{\rho}Hs(\omega), \\ H = \begin{pmatrix} \cos\theta\cos\varphi, \sin\theta\cos\varphi, \sin\varphi \end{pmatrix}^T \end{gathered} \tag{5}$$

where c is the speed of sound in air, λ is the density of air. Fourier transform of the particle velocity produced by m loudspeakers at the receiving point is expressed as:

$$\tilde{u}(\vec{r},\omega) = -\frac{G}{c\lambda}\frac{e^{-ik\rho}}{\rho}\tilde{H}Ws(\omega) \tag{6}$$

$$\tilde{H} = \begin{pmatrix} \cos\theta_1\cos\varphi_1 & \cos\theta_2\cos\varphi_2 & \cdots & \cos\theta_m\cos\varphi_m \\ \sin\theta_1\cos\varphi_1 & \sin\theta_2\cos\varphi_2 & \cdots & \sin\theta_m\cos\varphi_m \\ \sin\varphi_1 & \sin\varphi_2 & \cdots & \sin\varphi_m \end{pmatrix}$$

$$W = \begin{pmatrix} w_1, w_2, \cdots, w_m \end{pmatrix}^T$$

Let $u(\vec{r},\omega) = \tilde{u}(\vec{r},\omega)$, we can get:

$$\tilde{H}\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix} = \begin{pmatrix} \cos\theta\cos\varphi \\ \sin\theta\cos\varphi \\ \sin\varphi \end{pmatrix} \tag{7}$$

Then we divide the first and second rows by the third row in equation (7), we could get:

$$\begin{array}{l} \frac{\sum_{l=1}^{m} w_l \cos\theta_l \cos\varphi_l}{\sum_{l=1}^{m} \omega_l \sin\varphi_l} = \frac{\cos\theta\cos\varphi}{\sin\varphi} \\ \frac{\sum_{l=1}^{m} w_l \sin\theta_l \cos\varphi_l}{\sum_{l=1}^{m} \omega_l \sin\varphi_l} = \frac{\sin\theta\cos\varphi}{\sin\varphi} \end{array} \tag{8}$$

Equation (8) could guarantee that direction of particle velocity of virtual sound source and m loudspeakers at the receiving point are the same. From equation (8), we could get:

$$\begin{array}{l} \sum_{l=1}^{m} w_l(\cos\theta_l \cos\varphi_l \sin\varphi - \sin\varphi_l \cos\theta\cos\varphi) = 0 \\ \sum_{l=1}^{m} w_l(\sin\theta_l \cos\varphi_l \sin\varphi - \sin\varphi_l \sin\theta\cos\varphi) = 0 \end{array} \tag{9}$$

Together with equation (4), we could get:

$$LW = E_1, L = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1m} \\ t_{21} & t_{22} & \cdots & t_{2m} \\ 1 & 1 & \cdots & 1 \end{pmatrix}, E_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{10}$$

where

$$\begin{array}{lcl} t_{11} & = & \cos\theta_1 \cos\varphi_1 \sin\varphi - \sin\varphi_1 \cos\theta\cos\varphi \\ t_{12} & = & \cos\theta_2 \cos\varphi_2 \sin\varphi - \sin\varphi_2 \cos\theta\cos\varphi \\ \cdots & & \\ t_{1m} & = & \cos\theta_m \cos\varphi_m \sin\varphi - \sin\varphi_m \cos\theta\cos\varphi \\ t_{21} & = & \sin\theta_1 \cos\varphi_1 \sin\varphi - \sin\varphi_1 \sin\theta\cos\varphi \\ t_{22} & = & \sin\theta_2 \cos\varphi_2 \sin\varphi - \sin\varphi_2 \sin\theta\cos\varphi \\ \cdots & & \\ t_{2m} & = & \sin\theta_m \cos\varphi_m \sin\varphi - \sin\varphi_m \sin\theta\cos\varphi \end{array} \tag{11}$$

The square error of particle velocity magnitude at the receiving point is defined as:

$$\begin{array}{l} E(w_1, w_2, w_3, \cdots, w_m) \\ = \|u(\vec{r},\omega) - \sum_{j=1}^{m} u_j(\vec{r},\omega)\|_2^2 \\ = \|G\frac{e^{-ik\rho}}{c\lambda\rho}s(\omega)\|_2^2[(\cos\theta\cos\varphi - \sum_{j=1}^{m} w_j \cos\theta_j \cos\varphi_j)^2 \\ +(\sin\theta\cos\varphi - \sum_{j=1}^{m} w_j \cdot \sin\theta_j \cos\varphi_j)^2 \\ +(\sin\varphi - \sum_{j=1}^{m} w_j \sin\varphi_j)^2] \end{array} \tag{12}$$

$u(\vec{r},\omega)$ is particle velocity of the virtual sound source at the receiving point, $u_j(\vec{r},\omega)(j{=}1,2,\cdots,m)$ is particle velocity of the j^{th} loudspeaker at the receiving point. Equation (10) could guarantee sound pressure and the direction of particle velocity at the receiving point constant, so it is necessary to make the square error of particle velocity magnitude at the receiving point as little as possible. For a given signal, virtual sound source and loudspeakers' position, $\|G\frac{e^{-ik\rho}}{c\lambda\rho}s(\omega)\|_2^2$ is a constant. So to make the error square of particle velocity magnitude E minimum, we should make the formula in bracket of equation (12) minimum. Then the global solution of replacing a virtual sound source by m loudspeakers is equivalent to solving the following question:

$$\begin{array}{l} \min\limits_{W} \frac{1}{2}\|\tilde{H}W - H\|_2^2 \\ s.t.\ \ LW = E_1 \\ \qquad w_1, w_2, \cdots, w_m \geq 0 \end{array} \tag{13}$$

Equation (13) is a least squares problems with inequality constraints and it could be worked out by existing mature algorithms such as Trust Region Algorithm [19].

Then we suppose that q-channel in original sysytem is simplified to m-channel ($q{>}m$) in reproduced system, there are q loudspeakers (recorded as $ld_1, ld_2, \cdots, ld_q$) need to be replaced by m loudspeakers (recorded as $rd_1, rd_2, \cdots, rd_m$) . In the replacement process of loudspeaker ld_s , $s{\in}\{1, 2, \cdots, q\}$, the distribution coefficients of loudspeaker $rd_1, rd_2, \cdots, rd_m$ are $w_{1s}, w_{2s}, \cdots, w_{ms}$ respectively, which could be obtained by the global model method. Because particle velocity and sound pressure are linear function of input signal, so the final distribution coefficient of loudspeaker rd_e, $e{\in}\{1, 2, \cdots, m\}$ is:

$$w_{final-e} = \sum_{g=1}^{q} w_{eg} \tag{14}$$

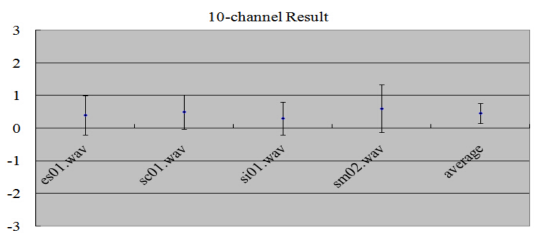
4. EXPERIMENTS

The performance of our method and Ando's method are measured by two tests. In test I, 22-channel are simplified to 10-channel, in test II, 22-channel are simplified to 8-channel (Figure 1). The radius of the spheres is 2 m. The two low-frequency effect channels are not processed. First, objective tests are done. Single frequency signals of 1000Hz is selected as the original signal for 22-channel. In test I, compared with the original 22-channel, the particle velocity error at receiving point is 6.34% by our method, the particle velocity error at receiving point is 7.74% by Ando's method. In test II, compared with the original 22-channel, the particle velocity error at receiving point is 11.81% by our method, the particle velocity error at receiving point is 13.31% by Ando's method.

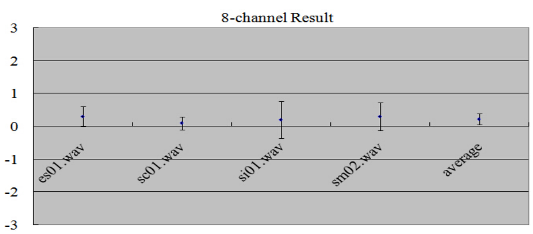
Table 1. Levels comparison standard

Comparison of the Stimuli	Score
Sound image of A is much closer to Ref than B	+3
Sound image of A is closer to Ref than B	+2
Sound image of A is slightly closer to Ref than B	+1
Sound image of A to Ref is the same as B	0
Sound image of A is slightly further to Ref than B	-1
Sound image of A is further to Ref than B	-2
Sound image of A is much further to Ref than B	-3

Then, subjective tests are done. Comparison Mean Opinion Score (CMOS) is used to test our method and Ando's method, the test material consists of Ref/A/B, in which Ref is the original sound source signal, A is signal generated by our method, and B is signal generated by Ando's method. Ref is played back by 22 loudspeakers of original 22.2 multichannel system. A and B are played back by less loudspeakers of reproduced multichannel system. We compare the sound image of A and B which is closer to Ref. The score has 7 levels, which are listed in Table 1. 10 listeners performed the listening test. All of them actively work in the audio field. The center of listener's head is at the receiving point (spherical center O in Figure 1) in testing. The test results consist



(a)



(b)

Fig. 2. CMOS score of each item for (a) 10 channel in Figure 1, test I; (b) 8 channel in Figure 1, test II.

of an average score and a 95% confidence interval. We use 4 MPEG test sequences: es01, sc01, si01, sm02 whose sample rate are 48 kHz, bit depth are 16 bits, amplitude are -40 dB, and length are 8s. The test results for each item are given in Figure 2. We can see that all these cases are statistically comparable to Ando's method in a 95% confidence interval sense. The average scores of our method are higher than Ando's method for all cases, which means that location accuracy of our method is better than Ando's method when we simplify 22-channel to 10- and 8-channel and are in accordance with objective test results. The reason is that our method can get the best solution which could maintain physical properties of sound at the receiving point as much as possible.

5. RELATION TO PRIOR WORK

First we explain why Ando's method misses the magnitude of particle velocity. For example, suppose a loudspeaker V whose input signal is $s_0(\omega)$ in frequency domain is at $(2, 90°, 0°)$, wave number is k_0, it is replaced by loudspeakers A, B, C which are at $(2, 90°, 45°)$, $(2, 120°, 0°)$, $(2, 60°, 0°)$ by Ando's method, the distribution coefficients of loudspeakers A, B, C are 0, $\frac{1}{2}$, $\frac{1}{2}$ (the computation method is described in [17]). The pressure at the receiving point produced by loudspeakers V, A, B, C respectively are $p_V{=}G\frac{e^{-2ik_0}}{2}s_0(\omega)$, $p_A{=}G\frac{e^{-2ik_0}}{2}s_0(\omega){\cdot}0$, $p_B{=}G\frac{e^{-2ik_0}}{2}s_0(\omega){\cdot}\frac{1}{2}$, $p_C{=}G\frac{e^{-2ik_0}}{2}s_0(\omega){\cdot}\frac{1}{2}$, so $p_V{=}p_A{+}p_B{+}p_C$. The particle velocity at the receiving point produced by loudspeakers V, A, B, C respectively are:

$$\begin{aligned}
u_V &= -\tfrac{G}{c\lambda}\tfrac{e^{-2ik_0}}{2}\left(\ \cos 90°\cos 0°, \sin 90°\cos 0°, \sin 0°\ \right)^T \\
&\cdot s_0(\omega) = -\tfrac{G}{c\lambda}\tfrac{e^{-2ik_0}}{2}\left(\ 0,1,0\ \right)^T s_0(\omega), \\
u_A &= -\tfrac{G}{c\lambda}\tfrac{e^{-2ik_0}}{2}\left(\ \cos 90°\cos 45°, \sin 90°\cos 45°, \sin 45°\ \right)^T \\
&\cdot s_0(\omega)\cdot 0 = \left(\ 0,0,0\ \right)^T, \\
u_B &= -\tfrac{G}{c\lambda}\tfrac{e^{-2ik_0}}{2}\left(\ \cos 120°\cos 0°, \sin 120°\cos 0°, \sin 0°\ \right)^T \\
&\cdot s_0(\omega)\cdot\tfrac{1}{2} = -\tfrac{G}{c\lambda}\tfrac{e^{-2ik_0}}{2}\left(\ -\tfrac{1}{4}, \tfrac{\sqrt{3}}{4}, 0\ \right)^T s_0(\omega), \\
u_C &= -\tfrac{G}{c\lambda}\tfrac{e^{-2ik_0}}{2}\left(\ \cos 60°\cos 0°, \sin 60°\cos 0°, \sin 0°\ \right)^T \\
&\cdot s_0(\omega)\cdot\tfrac{1}{2} = -\tfrac{G}{c\lambda}\tfrac{e^{-2ik_0}}{2}\left(\ \tfrac{1}{4}, \tfrac{\sqrt{3}}{4}, 0\ \right)^T s_0(\omega)
\end{aligned}$$

Then the sum of particle velocity u_A, u_B, u_C is:

$$u_{sum} = u_A + u_B + u_C = -\tfrac{G}{c\lambda}\tfrac{e^{-2ik_0}}{2}\left(\ 0,1,0\ \right)^T s_0(\omega)\cdot\tfrac{\sqrt{3}}{2}$$

Though the direction of vector u_v is the same as the direction of vector u_{sum}, but $|u_v|{\neq}|u_{sum}|$. This means that Ando's method just maintain sound pressure and the direction of particle velocity at the receiving point in replacement process, it can't maintain the magnitude of particle velocity. Also in the simplifying process, local solution is obtained by three loudspeakers in reproduced multichannel system replacing a loudspeaker in 22.2 multichannel system by Ando's method. The global solution is got by simple corresponding addition of these local solution. In our method, all loudspeakers in reproduced multichannel system join in replacing a loudspeaker in 22.2 multichannel system, sound pressure and the direction of particle velocity at the receiving point is maintained, and the global solution is obtained by making the square error of particle velocity magnitude at the receiving point as little as possible. Our method is a optimization routine for searching global optimal solution.

6. CONCLUSIONS

To promote the application of 3D multichannel system in family, this paper proposes a 3D multichannel system reproduction method. Relative to original multichannel system, this method can reduce channel number, at the same time, it can guarantee sound pressure and the direction of particle velocity at the receiving point constant, and make the square error of particle velocity magnitude at the receiving point as little as possible. The method build a general global model which is equivalent to a least squares problems with non-negative constraints, and it can be worked out by existing mature algorithms. Compared with Ando's method, the proposed method could get the global optimal solution of simplifying multichannel system, which can maintain sound pressure and particle velocity around receiving point better. Subjective experimental results show that the proposed method have better localization effect than traditional methods in simplifying 22.2 multichannel system to 10- and 8-channel system. The future work is to promote localization effect of our method futher in simplifying 22.2 multichannel system.

7. REFERENCES

[1] M. A. Gerson, "Periphony: With-height sound reproduction," *Journal of the Audio Engineering Society*, vol. 21, no. 1, pp. 2–10, 1973.

[2] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 1, pp. 135–143, Jan. 2005.

[3] 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, no. 6, pp. 988–999, 2008.

[4] W. Zhang and T. D. Abhayapala, "Three dimensional sound field reproduction using multiple circular loudspeaker arrays: Functional analysis guided approach," *IEEE/ACM Transactions on Audio, Speech,and Language Processing*, vol. 22, no. 7, pp. 1184–1194, 2014.

[5] A. J. Berkhout, "A holographic approach to acoustic control," *Journal of the Audio Engineering Society*, vol. 36, no. 12, pp. 977–995, 1988.

[6] A. J. Berkhout et al., "Acoustic control by wave field synthesis," *Journal of the Acoustic Society of America*, vol. 93, no. 5, pp. 2764–2778, 1993.

[7] J.-M. Lee et al., "Wave field synthesis of a virtual source located in proximity to a loudspeaker array," *Journal of the Acoustic Society of America*, vol. 134, no. 3, pp. 2106–2117, 2013.

[8] J. Blauert, "Spatial hearing," Cambridge: MIT Press, 1983.

[9] Fuliang Yin et al., "Review on 3d audio technology," *Journal on Communications*, vol. 32, no. 2, pp. 130–138, 2011.

[10] E. Rasumow et al., "Smoothing individual head-related transfer functions in the frequency and spatial domains," *Journal of the Acoustic Society of America*, vol. 135, no. 4, pp. 2012–2025, 2014.

[11] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *Journal of Audio Engineering Society*, vol. 45, no. 6, pp. 456–466, 1997.

[12] V. Pulkki, "Localization of amplitude-panned virtual sources ii: Two-and three-dimensional panning," *Journal of Audio Engineering Society*, vol. 49, no. 9, pp. 753–767, 2001.

[13] B. B. Bauerl, "Phasor analysis of some stereophonic phenomena," *The Journal of the Acoustical Society of America*, vol. 33, no. 11, pp. 1536–1539, 1973.

[14] B. Bernfeld, "Attempts for better understanding of the directional stereophonic listening mechanism," *Proceedings of the 44th Audio Engineering Society Convention*, Feb. 1973.

[15] Song Wang et al., "Sound intensity and particle velocity based three-dimensional panning methods by five loudspeakers," in *2013 IEEE International Conference on Multimedia and Expo, ICME 2013*, San Jose, CA, United States, July. 2013.

[16] Song Wang et al., "Three-dimensional panning by four loudspeakers and its solution," in *2014 IEEE International Conference on Multimedia and Expo Workshops, ICMEW 2014*, Chengdu, China, July. 2014.

[17] A. Ando, "Conversion of multichannel sound signal maintaining physical properties of sound in reproduced sound field," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 6, pp. 1467–1475, Aug. 2011.

[18] A. D. Pierce, "Acoustics, an introduction to its physical principles and applications," *Acoustical Society of America*, 1989.

[19] Moré J.J. and D.C. Sorensen, "Computing a trust region step," *SIAM Journal on Scientific and Statistical Computing*, vol. 3, pp. 553–572, 1983.