Journal of Computers Vol. 30 No. 3, 2019, pp. 213-223
doi:10.3966/199115992019063003016

A Non-central Point Sound Field Reconstruction Method for 22.2 Multichannel System

Song Wang, Cong Zhang*, Han-Chieh Chao, Heng Wang

School of Mathematics & Computer Science, Wuhan Polytechnic University, Wuhan 430048, China
{wangsongf117, co_zh2018}@163.com

Received 13 January 2019; Revised 9 March 2019; Accepted 12 March 2019

Abstract. In recent years, the success of 3D video technology makes the development of 3D audio technology more and more urgent. Most of the existing 3D audio systems aim at restoring the sound field of the central listening point, so that the sound field reconstruction at the central listening point is the best, but the sound field reconstruction at the non-central listening point is poor. Listeners may also be located at non-central listening points. In order to recover the sound field at the non-central listening point, a 3D sound field reconstruction method at the non-central listening point is proposed for the reconstruction of 22.2 multichannel system. First, a virtual sphere is constructed with the non-central point as the center and the distance between the non-central point and the virtual sound source as the radius. Second, connect the location of loudspeaker in the reconstructed system to the location of non-central point to get a segment, the intersection between the segment and the virtual sphere is the location of a virtual loudspeaker. Third, the signal of the virtual loudspeaker is solved by sound pressure and particle velocity minimization method. Then, by keeping sound pressure at the non-central point constant, the signal of real loudspeaker can be obtained from virtual loudspeaker signal. Repeating above steps, original 22 loudspeakers in 22.2 multichannel system can be replaced by 10 loudspeakers. The simulation results show that the proposed method is superior to traditional methods.

Keywords: 22.2 multichannel system, non-central point, sound field reconstruction

1 Introduction

The box office of Chinese films reached $7.9 billion in 2017, which has increased $2.7 billion than that in 2016 and led to steady growth in the global box office. In 2017, 44 3D movies were released worldwide, which has 36 more than that in 2008. With the development and promotion of 3D film and television, 3D audio technology corresponding to 3D video technology is gaining more and more attention. Currently in three-dimensional audio research area, there are several existing technologies.

In 1973, Gerzon from Mathematical Institute, University of Oxford has proposed Ambisonics technology [1-2]. By expanding the spherical harmonic function, Ambisonics can accurately reconstruct the sound pressure field near the center point [3]. When a larger central listening area is needed, the spherical harmonic function needs to be expanded to a larger order, which will lead to an increase in the number of loudspeakers. In 1988, Berkhout from Delft University of Technology has proposed Wave Field Synthesis (WFS). Wave field is represented technically as an integral form, and input of loudspeaker array is represented as explicit continuous driver function [4-7]. However, the continuous placement of loudspeaker does not conform to the reality. In practice, loudspeakers are discretely placed. To get better reconstruction effect, the gap between loudspeakers should be small, which will lead to an increase in the number of loudspeakers. The transfer function from sound source to human ear represents Head-Related Transfer Function (HRTF) [8-10]. HRTF describes the complex process of sound filtering by pinna, head, trunk and so on. Different individuals have different HRTF. The common HRTF will lead to sound image confusion between the front and the back as well as the up and the down. But

* Corresponding Author

measuring personalized HRTF requires expensive experimental equipment and a lot of time.

Vector Base Amplitude Panning (VBAP) uses three loudspeakers to synthesize a virtual sound source in three dimensions space [11-12]. Each of the three loudspeakers and the virtual sound source corresponds to a vector. Starting points of the vectors are center point, which also is the origin, and end points of vectors are three loudspeakers or a single virtual sound source. The vectors corresponding to three loudspeakers can represent the vector corresponding to a single virtual sound source linearly. VBAP mainly focuses on the restoration of sound field at the center listening point, so the sound field at the center point is well reconstructed. VBAP is popular because of its simple loudspeaker layout, cheap price, easy to use, low computational complexity, no disruption at the central point, high timbral quality and so on [13].

Ando also has proposed a method of synthesizing a virtual sound source using three loudspeakers, which is abbreviated as Ando method here [14]. Ando method is a physical property method because it maintains the sound pressure and the direction of particle velocity at the central point, which is the physical foundation of VBAP method. By Ando method, NHK 22.2 multichannel system can be simplified to 10-channel system with good listening effect, which reduces the cost of 22.2 multichannel system in the home environment. But like VBAP, the listening effect is only best at the central point.

These methods in literatures [11, 12, 14] have a common feature, that is, the positioning accuracy at the central point is the best, and the positioning accuracy at the non-central point is poor. But in practical scenario, the listener may also be located at a non-central point. At present, some techniques of non-central point sound field reconstruction already exist, such as PMSZ and PVMSZ method. PMSZ matches the sound pressure in the same region of the original sound field and the reconstructed sound field [15]. Although [15] only conducts experiments on central region sound field reconstruction, PMSZ method can also be used in three-dimensional non-central region sound field reconstruction theoretically. When loudspeakers are placed unevenly, the reconstruction effect of PMSZ method is poor. To overcome this problem, PVMSZ method has been proposed. PVMSZ method matches the particle velocity in the same region of the original sound field and the reconstructed sound field [16]. Although the experiment of central region sound field reconstruction is only carried out in reference [16], PVMSZ can also be used in three-dimensional non-central region sound field reconstruction theoretically. In the reconstruction of a sound field in a designated non-central region, PVMSZ tries to restore the original sound field inside the non-central region. However, through experiments, we find that the reconstruction error of PVMSZ is relatively large, so it is necessary to study a better method of sound field reconstruction at non-central listening points.

1.1 Goal and Structure

Aiming at above problems, this paper proposes a non-central point three-dimensional sound field reconstruction method with two times signal distributions by ten loudspeakers for 22.2 multichannel system reproduction to obtain better listening effect at some non-central point. The ten loudspeakers are also called actual loudspeakers. The proposed method constructs a virtual sphere with the non-central point as the center and selects ten points on the virtual sphere as the location of ten virtual loudspeakers. The proposed method first allocates the signal of a virtual sound source to the ten virtual loudspeakers by proposed sound pressure and particle velocity based method, and then allocates the signals from ten virtual loudspeakers to ten actual loudspeakers by keeping the sound pressure at the non-central point constant. This paper is structured as follows: Section 2 introduces the definition of two sound physical properties, Section 3 introduces the proposed method, Section 4 describes the contrastive simulation experiments between different methods, and Section 5 draws conclusions.

2 Sound Pressure and Particle Velocity of Point Sound Source

This section introduces the basic definition of sound pressure and particle velocity. A single loudspeaker can be seen as a point sound source. A single loudspeaker locates at $\vec{\xi} = (\xi_x, \xi_y, \xi_z)^T$ produces sound pressure at the receiving point $\vec{\eta} = (\eta_x, \eta_y, \eta_z)^T$ as following:

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

Where the receiving point also is listening point, $k={2\pi f}/{c}$ is wave number, f is the frequency of sound signal, c is the speed of sound signals traveling in the air, $s\left(\omega\right)$ is the Fourier transform of the loudspeaker signal, i is imaginary unit, e is natural constant, T represents transpose of a matrix, G represents the proportional coefficient between the sound pressure at unit distance from the loudspeaker and the loudspeaker signal.

A single loudspeaker locates at $\vec{\xi}=(\xi_x,\xi_y,\xi_z)^T$ produces particle velocity at the receiving point $\vec{\eta}=(\eta_x,\eta_y,\eta_z)^T$ as following [14]:

$$u(\vec{\eta},\omega)=G(1+\frac{1}{ik\left|\vec{\eta}-\vec{\xi}\right|})\frac{e^{-ik\left|\vec{\eta}-\vec{\xi}\right|}}{\left|\vec{\eta}-\vec{\xi}\right|^2}\begin{pmatrix}\eta_x-\xi_x\\ \eta_y-\xi_y\\ \eta_z-\xi_z\end{pmatrix}s(\omega)\approx G\frac{e^{-ik\left|\vec{\eta}-\vec{\xi}\right|}}{\left|\vec{\eta}-\vec{\xi}\right|^2}\begin{pmatrix}\eta_x-\xi_x\\ \eta_y-\xi_y\\ \eta_z-\xi_z\end{pmatrix}s(\omega) \tag{2}$$

When f is larger, k is larger, then ${1}/{ik\left|\vec{\eta}-\vec{\xi}\right|}\rightarrow 0$. so in equation (2), approximately equal is found.

3 Proposed Method

The original 22.2 channel system without two low frequency effect channels is shown in Fig. 1(a), the reconstructed 10 channel system is shown in Fig. 1(b). First, we will research the case that a virtual sound source (a signal loudspeaker in original system) is replaced by m loudspeakers in non-central point sound field reconstruction. As shown in Fig. 1, the central point is at $O\left(0,0,0\right)$, which is also the origin of the rectangular coordinate system; the non-central point is n, whose rectangular coordinate is $\vec{n}=\left(n_x,n_y,n_z\right)^T$, and polar coordinate is $\vec{n}=\left(\rho_x,\theta_y,\varphi_z\right)^T$; A virtual sound source and m loudspeakers all locate on sphere O, their rectangular coordinates are: $\vec{\chi}=\left(\chi_x,\chi_y,\chi_z\right)^T$, $\vec{\xi}_j=\left(\xi_{jx},\xi_{jx},\xi_{jx}\right)^T$ ($j=1,2,\cdots m$), and polar coordinates are: $\vec{\chi}=\left(\rho_0,\theta_0,\varphi_0\right)^T$, $\vec{\xi}_j=\left(\rho_j,\theta_j,\varphi_j\right)^T$, ($j=1,2,\cdots m$). These m loudspeakers are called actual loudspeakers.

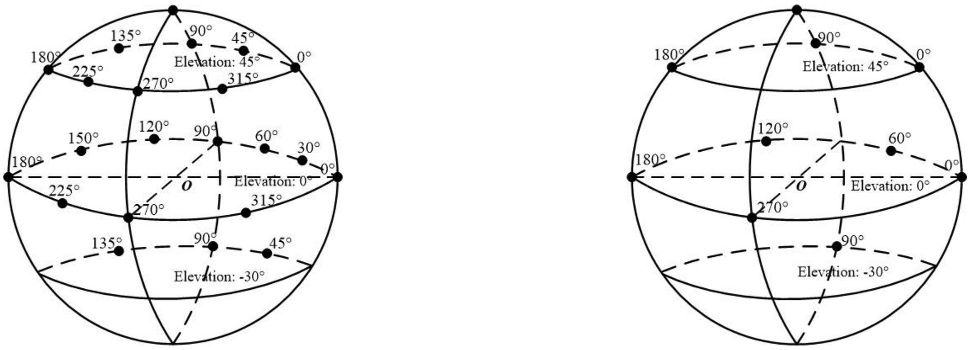
3.1 The Location of Virtual Loudspeaker

Fig. 1(c) is an example of how to find a virtual loudspeaker from an actual loudspeaker. Connect the virtual sound source located at $\vec{\chi}$ to the non-central listening point located at $\vec{n}$, the length of this line segment is denoted as:

$$R=\sqrt{\left(n_x-\chi_x\right)^2+\left(n_y-\chi_y\right)^2+\left(n_z-\chi_z\right)^2} \tag{3}$$

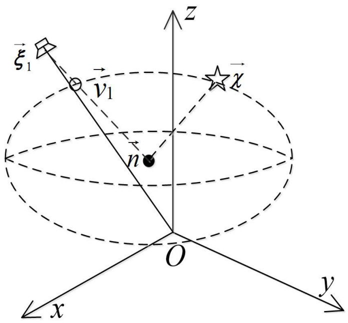
Sphere n is constructed with R as the radius and point n is the center. The sphere n is also called virtual sphere. Join point n and the ends of m vectors $\vec{\xi}_j$, $j=1,2,\cdots m$, we can get m line segments, which intersect with virtual sphere n at points $\vec{v}_j$, $j=1,2,\cdots m$ respectively. Points $\vec{v}_j$, $j=1,2,\cdots m$ are set to the location of m virtual loudspeakers.

When O is the origin of coordinated system, rectangular coordinates of $\vec{v}_j$, $j=1,2,\cdots m$ are: $\vec{v}_j\left(v_{jx},v_{jy},v_{jz}\right)$, polar coordinates of $\vec{v}_j$, $j=1,2,\cdots m$ are: $\left(\rho_j^v,\theta_j^v,\varphi_j^v\right)$. When n is the origin of coordinated system, polar coordinates of $\vec{v}_j$, $j=1,2,\cdots m$ are: $\left(\rho_j^{nv},\theta_j^{nv},\varphi_j^{nv}\right)$, the formula is as follows:



(a) Original 22.2 channel system without two low frequency effect channels

(b) 10 channel system in reconstruction system



(c) The position of virtual loudspeaker. Point O is the origin of the coordinate system, the non-central point is located at point $\vec{n}$, $\vec{n}$ is corresponds to point n, virtual sphere is ball n, the location of a virtual sound source is the point $\vec{\chi}$, the location of an actual loudspeakers is $\vec{\xi}_1$, the location of a virtual loudspeaker is $\vec{v}_1$

Fig. 1.

$$\rho_j^{nv} = R,\ \theta_j^{nv} = \arctan(\frac{\xi_{jy} - n_y}{\xi_{jx} - n_x}),\ \varphi_j^{nv} = \arcsin\left(\frac{\xi_{jz} - n_z}{R}\right),\ j = 1, 2, \cdots m \quad \textbf{(4)}$$

When O is the origin of coordinated system, rectangular coordinates of $\vec{v}_j$, $j = 1, 2, \cdots m$ can be calculated according to the following formula:

$$v_{jx} = \rho_j^{nv} \cos\theta_j^{nv} \cos\varphi_j^{nv} + n_x,\ v_{jy} = \rho_j^{nv} \sin\theta_j^{nv} \cos\varphi_j^{nv} + n_y,\ v_{jz} = \rho_j^{nv} \sin\varphi_j^{nv} + n_z,\ j = 1, 2, \cdots m \quad \textbf{(5)}$$

When n is the origin of coordinates system, the polar coordinates of $\vec{\chi}$ is marked as $\vec{\chi}'(\rho'_0, \theta'_0, \varphi'_0)$, which can be calculated by the following formula:

$$\rho'_0 = R,\ \theta'_0 = \arctan(\frac{\chi_y - n_y}{\chi_x - n_x}),\ \varphi'_0 = \arcsin\left(\frac{\chi_z - n_z}{R}\right) \quad \textbf{(6)}$$

3.2 The Solution of Virtual Loudspeaker Distribution Coefficient

On sphere n, the virtual sound source located at $\vec{\chi}$ and m virtual loudspeakers located at $\vec{v}_j$, $j = 1, 2, \cdots m$ are on a same sphere, which means $\rho_j^{nv} = \rho'_0$, $j = 1, 2, \cdots m$. The sound pressure at point $\vec{n}$ produced by the virtual sound source located at $\vec{\chi}$ is:

$$p\left(\vec{n},\omega\right)=G\frac{e^{-ik\left|\vec{n}-\vec{\chi}\right|}}{\left|\vec{n}-\vec{\chi}\right|}s\left(\omega\right)=G\frac{e^{-ik\rho'_0}}{\rho'_0}s\left(\omega\right) \quad \textbf{(7)}$$

The particle velocity at point $\vec{n}$ produced by the virtual sound source located at $\vec{\chi}$ is [14]:

$$u(\vec{n},\omega)=G\frac{e^{-ik\left|\vec{n}-\vec{\chi}\right|}}{\left|\vec{n}-\vec{\chi}\right|^2}\begin{pmatrix}n_x-\chi_x\\ n_y-\chi_y\\ n_z-\chi_z\end{pmatrix}s(\omega)=-\frac{G}{\tau c}\frac{e^{-ik\rho'_0}}{\rho'_0}\begin{pmatrix}\cos\theta'_0\cos\varphi'_0\\ \sin\theta'_0\cos\varphi'_0\\ \sin\varphi'_0\end{pmatrix}s(\omega) \quad \textbf{(8)}$$

Where τ is the density of air. The sound pressure at point $\vec{n}$ produced by m virtual loudspeakers located at $\vec{v}_j$, $j=1,2,\cdots m$ is:

$$p'\left(\vec{n},\omega\right)=\sum_{j=1}^{m}G\frac{e^{-ik\left|\vec{n}-\vec{v}_j\right|}}{\left|\vec{n}-\vec{v}_j\right|}s_j^v\left(\omega\right)=\sum_{j=1}^{m}G\frac{e^{-ik\rho_j^{nv}}}{\rho_j^{nv}}w_j^v s\left(\omega\right) \quad \textbf{(9)}$$

The particle velocity at point $\vec{n}$ produced by m virtual loudspeakers located at $\vec{v}_j$, $j=1,2,\cdots m$ is:

$$u'(\vec{n},\omega)=\sum_{j=1}^{m}G\frac{e^{-ik\left|\vec{n}-\vec{v}_j\right|}}{\left|\vec{n}-\vec{v}_j\right|^2}\begin{pmatrix}n_x-v_{jx}\\ n_y-v_{jy}\\ n_z-v_{jz}\end{pmatrix}s_j^v(\omega)=\sum_{j=1}^{m}-\frac{G}{\tau c}\frac{e^{-ik\rho_j^{nv}}}{\rho_j^{nv}}\begin{pmatrix}\cos\theta_j^{nv}\cos\varphi_j^{nv}\\ \sin\theta_j^{nv}\cos\varphi_j^{nv}\\ \sin\varphi_j^{nv}\end{pmatrix}w_j^v s(\omega) \quad \textbf{(10)}$$

Where $s_j^v\left(\omega\right)$ is the assigned signal of the virtual loudspeaker located at $\vec{v}_j$, $j=1,2,\cdots m$, w_j^v is the distribution coefficient and $s_j^v(\omega)=w_j^v s(\omega)$.

Let the sound pressure at point $\vec{n}$ produced by the virtual sound source located at $\vec{\chi}$ is the same as the sound pressure at point $\vec{n}$ produced by m virtual loudspeakers located at $\vec{v}_j$, $j=1,2,\cdots m$, namely $p\left(\vec{n},\omega\right)=p'\left(\vec{n},\omega\right)$, we could get:

$$G\frac{e^{-ik\rho'_0}}{\rho'_0}s\left(\omega\right)=\sum_{j=1}^{m}G\frac{e^{-ik\rho_j^{nv}}}{\rho_j^{nv}}w_j^v s\left(\omega\right) \quad \textbf{(11)}$$

Let the particle velocity at point $\vec{n}$ produced by the virtual sound source located at $\vec{\chi}$ is the same as the particle velocity at point $\vec{n}$ produced by m virtual loudspeakers located at $\vec{v}_j$, $j=1,2,\cdots m$, namely $u(\vec{n},\omega)=u'(\vec{n},\omega)$, we could get:

$$-\frac{G}{\tau c}\frac{e^{-ik\rho'_0}}{\rho'_0}\begin{pmatrix}\cos\theta'_0\cos\varphi'_0\\ \sin\theta'_0\cos\varphi'_0\\ \sin\varphi'_0\end{pmatrix}s(\omega)=\sum_{j=1}^{m}-\frac{G}{\tau c}\frac{e^{-ik\rho_j^{nv}}}{\rho_j^{nv}}\begin{pmatrix}\cos\theta_j^{nv}\cos\varphi_j^{nv}\\ \sin\theta_j^{nv}\cos\varphi_j^{nv}\\ \sin\varphi_j^{nv}\end{pmatrix}w_j^v s(\omega) \quad \textbf{(12)}$$

Combine equation (11) and (12), and simplify them, we can get:

$$AW^v=B \quad \textbf{(13)}$$

Where:

$$A=\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \cos\theta_1^{nv}\cos\varphi_1^{nv} & \cos\theta_2^{nv}\cos\varphi_2^{nv} & \cdots & \cos\theta_m^{nv}\cos\varphi_m^{nv} \\ \sin\theta_1^{nv}\cos\varphi_1^{nv} & \sin\theta_2^{nv}\cos\varphi_2^{nv} & \cdots & \sin\theta_m^{nv}\cos\varphi_m^{nv} \\ \sin\varphi_1^{nv} & \sin\varphi_2^{nv} & \cdots & \sin\varphi_m^{nv} \end{pmatrix},\ B=\begin{pmatrix} 1 \\ \cos\theta'_0\cos\varphi'_0 \\ \sin\theta'_0\cos\varphi'_0 \\ \sin\varphi'_0 \end{pmatrix},\ W^v=\begin{pmatrix} w_1^v \\ w_2^v \\ \vdots \\ w_m^v \end{pmatrix} \quad \textbf{(14)}$$

Making the sound pressure and particle velocity at point $\vec{n}$ produced by the virtual sound source located at $\vec{\chi}$ as close as possible to the sound pressure and particle velocity at point $\vec{n}$ produced by m virtual loudspeakers located at $\vec{v}_j$, $j=1,2,\cdots m$ is equivalent to the following question:

$$\begin{aligned} &\min_{W^v}\frac{1}{2}\left\|AW^v-B\right\|_2^2 \\ &s.t.\ \ w_1^v,w_2^v,\cdots w_m^v\geq 0 \end{aligned} \quad \textbf{(15)}$$

Equation (15) is a least squares problems with inequality constraints and could be worked out by many existing algorithms.

3.3 The Solution of Actual Loudspeaker Distribution Coefficient

When O is the origin of coordinated system, the sound pressure at point $\vec{n}$ produced by actual loudspeakers located at $\vec{\xi}_j=\left(\xi_{jx},\xi_{jx},\xi_{jx}\right)^T$, $j=1,2,\cdots m$ is:

$$p_j^t\left(\vec{n},\omega\right)=G\frac{e^{-ik\left|\vec{n}-\vec{\xi}_j\right|}}{\left|\vec{n}-\vec{\xi}_j\right|}w_j^t s\left(\omega\right) \quad \textbf{(16)}$$

where w_j^t, $j=1,2,\cdots m$ represent the signal distribution coefficients of m actual loudspeakers.

When O is the origin of coordinated system, the sound pressure at point $\vec{n}$ produced by virtual loudspeakers located at $\vec{v}_j\left(v_{jx},v_{jy},v_{jz}\right)$, $j=1,2,\cdots m$ is:

$$p_j^v\left(\vec{n},\omega\right)=G\frac{e^{-ik\left|\vec{n}-\vec{v}_j\right|}}{\left|\vec{n}-\vec{v}_j\right|}w_j^v s\left(\omega\right) \quad \textbf{(17)}$$

where w_j^v, $j=1,2,\cdots m$ represent the signal distribution coefficients of m virtual loudspeakers, which can be calculated from equation (15).

By ensuring that the sound pressure at point $\vec{n}$ produced by the j^{th} virtual loudspeaker located at $\vec{v}_j$ is equal to the sound pressure at point $\vec{n}$ produced by the j^{th} actual loudspeaker located at $\vec{\xi}_j$, we can obtain the signal distribution coefficient of the j^{th} actual loudspeaker. The formula is as following:

$$\begin{pmatrix} p_1^v\left(\vec{n},\omega\right) \\ p_2^v\left(\vec{n},\omega\right) \\ \vdots \\ p_m^v\left(\vec{n},\omega\right) \end{pmatrix}=\begin{pmatrix} p_1^t\left(\vec{n},\omega\right) \\ p_2^t\left(\vec{n},\omega\right) \\ \vdots \\ p_m^t\left(\vec{n},\omega\right) \end{pmatrix} \quad \textbf{(18)}$$

Then we can get: $$CW^t=D \quad \textbf{(19)}$$

Where:

$$C=\begin{pmatrix} \frac{e^{-ik\left|\vec{n}-\vec{\xi}_1\right|}}{\left|\vec{n}-\vec{\xi}_1\right|} & 0 & \cdots & 0 \\ 0 & \frac{e^{-ik\left|\vec{n}-\vec{\xi}_2\right|}}{\left|\vec{n}-\vec{\xi}_2\right|} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \frac{e^{-ik\left|\vec{n}-\vec{\xi}_m\right|}}{\left|\vec{n}-\vec{\xi}_m\right|} \end{pmatrix},\ W^t=\begin{pmatrix} w_1^t \\ w_2^t \\ \vdots \\ w_m^t \end{pmatrix},\ D=\begin{pmatrix} \frac{e^{-ik\left|\vec{n}-\vec{v}_1\right|}}{\left|\vec{n}-\vec{v}_1\right|} w_1^v \\ \frac{e^{-ik\left|\vec{n}-\vec{v}_2\right|}}{\left|\vec{n}-\vec{v}_2\right|} w_2^v \\ \vdots \\ \frac{e^{-ik\left|\vec{n}-\vec{v}_m\right|}}{\left|\vec{n}-\vec{v}_m\right|} w_m^v \end{pmatrix} \tag{20}$$

Then we can get:

$$W^t=\left(C^H C+\alpha U\right)^{-1} C^H D \tag{21}$$

Where α is the regularization factor, -1 represents the inverse matrix, U is the identity matrix, H is Hermitian transpose of a matrix.

3.4 The Final Signal of m Actual Loudspeakers When Multiple Virtual Sound Sources Are Replaced by m Loudspeakers

From Section 3.3, we can know that the actual loudspeakers signal can be got by $s_j(\omega)=w_j^t s(\omega)$, $j=1,2,\cdots m$ when a virtual sound source (a signal loudspeaker in original system) is replaced by m loudspeakers in non-central point sound field reconstruction. When multiple virtual sound sources are replaced by m loudspeakers, each time we consider that a single virtual sound source to be replaced by m loudspeakers and calculate the signals of m actual loudspeakers by the method in Section 3.1 - Section 3.3 until all virtual sound sources are replaced by m actual loudspeakers, then the signal of the corresponding actual loudspeaker obtained in each replacement can be added to obtain the final distribution signal of the corresponding actual loudspeaker.

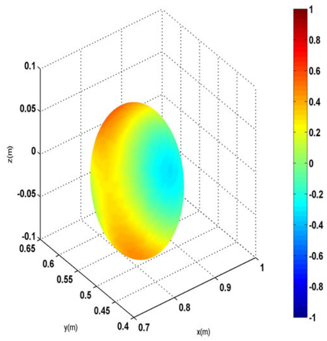
4 Simulation Experiments

Ando [14], PMSZ [15], PVMSZ [16] and proposed method will be compared in this section. 22 loudspeakers in original system is shown in Fig. 1(a), 10 loudspeakers in reconstructed system is shown is Fig. 1(b), these loudspeaker layouts refer to loudspeaker placement in [14], which are provided by researcher in NHK Science and Technology Research Laboratories to test the reconstructed effect of 22 channel system. The frequency of virtual sound source is 1000Hz. Sound speed in the air is 340 m/s. Radius of human head is about 0.085m. The rectangular coordinate of the non-central point is $\vec{n}=(0.8,0.5,0)^T$.

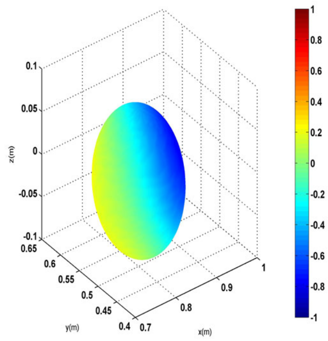
The sound field picture reflects the recovery of sound pressure. The comparison of sound field picture on a ball with radius 0.085m is shown in Fig. 2. We can see that the sound field of Ando, PMSZ and PVMSZ method is obviously different from the original sound field, and the sound field produced by proposed method is closer to the original sound field.

The relative mean square error (RMSE) in the desired reconstruction region is used as the measure of reconstruction error, its formula is:

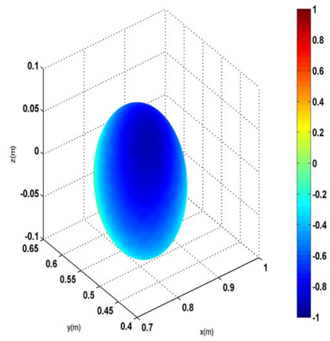
$$\kappa=\frac{\iiint_V \left|p_o-p_r\right|^2 dv}{\iiint_V \left|p_o\right|^2 dv} \tag{22}$$



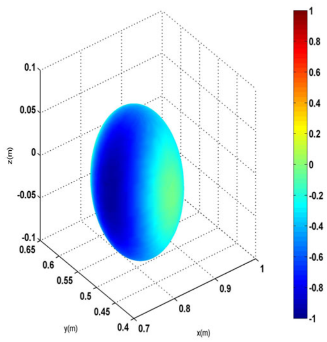
(a) Original field



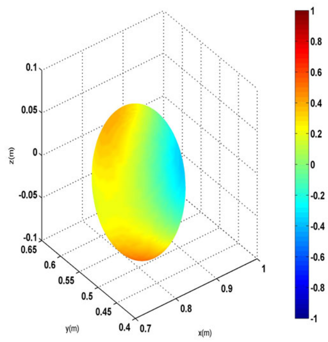
(b) Ando method



(c) PMSZ method



(d) PVMSZ method



(e) Proposed method

Fig. 2. Sound field pictures of different methods

Where the integration region is V, which is a 3D sphere with n as its center and r as its radius, p_o represents the sound pressure within the integral region V produced by original sound source, p_r represents the sound pressure within the integral region V produced by sound source in reconstruction system.

The relative mean square error comparison of different methods is shown in Fig. 3. From it we can see, the radius of the 3D sphere V varies from 0.085m to 1.02 meter, because the center of 3D sphere V is at $\vec{n} = (0.8, 0.5, 0)^T$. The sampling interval of radius r is 0.085m. The relative mean square error of Ando method is lower than PMSZ and PVMSZ method for all values of r; the relative mean square error of PVMSZ method is lower than that of PMSZ method when the radius of 3D sphere V is 0.085m, but higher than that of PMSZ method when the radius of 3D sphere V is larger than 0.085m; the relative mean square error of proposed method is the lowest for all values of r. When r is 0.085m, it means the 3D sphere V can contain just one head, the relative mean square error of different methods is shown in Table 1. From Table 1, we can see that the relative mean square error of proposed method is obviously lower than that of other methods, which means in a human head region, the proposed method performances better than traditional methods.

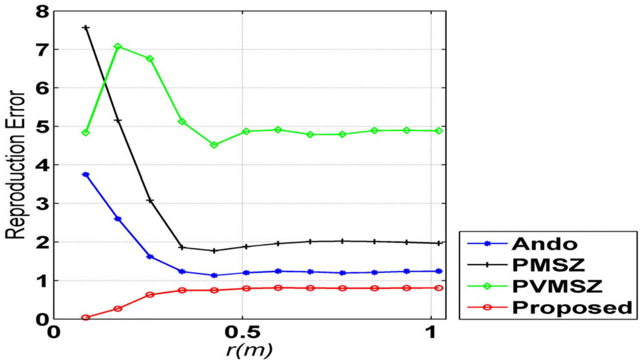


Fig. 3. The relative mean square error comparison

Table 1. The relative mean square error of different methods when r is 0.085m

Method	RMSE	Difference of RMSE with proposed method
Ando	375.00%	370.73%
PMSZ	756.62%	752.34%
PVMSZ	483.42%	479.14%
Proposed	4.28%	0.00%

Besides relative mean square error, the direction of intensity flow is another important measure of acoustic image localization. The time average sound intensity represents the average instantaneous intensity over a period of time, and is used to analyze the sound intensity flow. The definition of time averaged acoustic intensity is:

$$\vec{I}\left(\vec{n},\omega\right)=\frac{1}{2}\mathrm{Re}\left(p\left(\vec{n},\omega\right)u\left(\vec{n},\omega\right)^{*}\right) \tag{23}$$

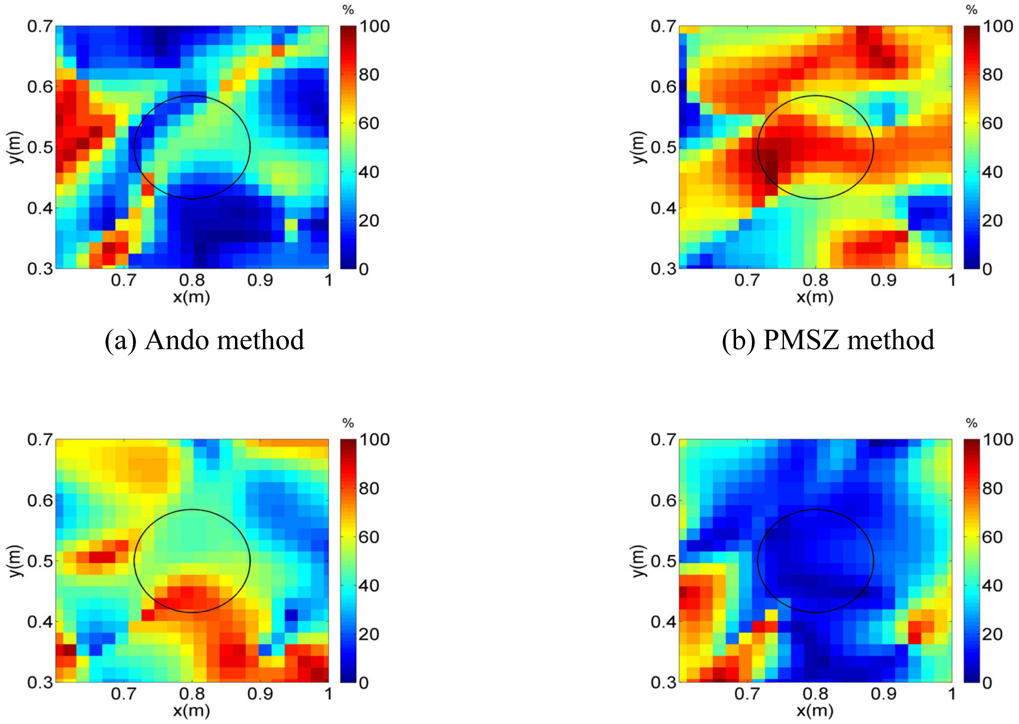
Where "*" denotes complex conjugate, 'Re' means the real part of the complex number. Intensity flow is the direction of the time average sound intensity, and its formula is:

$$\vec{D}_{I}\left(\vec{n},\omega\right)=\frac{\vec{I}\left(\vec{n},\omega\right)}{\left|\vec{I}\left(\vec{n},\omega\right)\right|} \tag{24}$$

Then the definition of intensity flow error is:

$$e_{IF}\left(\vec{n},\omega\right)=\frac{\cos^{-1}\left(\vec{D}_{Id}\left(\vec{n},\omega\right)\bullet\vec{D}_{Ir}\left(\vec{n},\omega\right)\right)}{\pi}\times 100\big/100 \tag{25}$$

Where $\vec{D}_{Id}\left(\vec{n},\omega\right)$ and $\vec{D}_{Ir}\left(\vec{n},\omega\right)$ are intensity flow of desired sound field and reconstructed sound field respectively. The intensity flow error comparison of different methods is shown in Fig. 4. The radius of black circle is human head radius, about 0.085m. From Fig. 4, we can see that in a human head area, the intensity flow error of PMSZ method is the highest, the intensity flow error of Ando method is lower than that of PVMSZ method. The intensity flow error of proposed method is the lowest of all.



(a) Ando method

(b) PMSZ method

(c) PVMSZ method

(d) Proposed method

Fig. 4. Sound intensity flow error of different methods

From above comparison of sound field picture, relative mean square error and sound intensity flow error, we can see that the test results of the proposed method are superior to traditional methods. These traditional methods mainly focus on the sound field recovery at the central point and neglect the sound field recovery at non-central point. Therefore, the reconstructed sound field at the central listening point is the best, while the reconstructed sound field at non-central listening point is poor. In non-central point sound field reconstruction, the proposed method references the advantages of traditional methods to maintain the best sound field effect at the central listening point, transforms the non-central point into the central point by constructing virtual sphere with the non-central point as the center. On the virtual sphere, the signal of original virtual sound sources are allocated to virtual loudspeakers one by one by keeping the sound pressure and particle velocity at the center of virtual sphere constant. By keeping the sound pressure at the non-central point constant, the signals of these virtual loudspeakers are converted into the signals of actual loudspeakers. Through the above operations, the physical property of sound at the non-central point is restored and maintained optimally, so the proposed method performs better than traditional methods.

5 Conclusions

Traditional methods can only reconstruct the sound field perfectly at the central point, but poorly at non-central point. Aiming at this problem, this paper proposed a non-central point three-dimensional sound field reconstruction method by multiple loudspeakers with two times signal distributions. In simulation experiments, 10 loudspeakers are used to reconstruct the original 22 channel system. The test results show that the proposed method is superior to traditional methods in comparisons of sound field picture, relative mean square error and sound intensity flow error. The proposed method is beneficial to the application of 22.2 multichannel system in home environment with less channel. When a listening point is given, the proposed method can intelligently calculate the distribution signals of loudspeakers and keep the optimal sound field reconstruction effect at the given listening point, which can improve the user's listening experience.

Acknowledgements

The authors would like to thank Natural Science Foundation of Hubei Province (No. 2018CFB408), Hubei Provincial Education Department Science and Technology Research Project (No. Q20181807), Hubei Provincial Major Science and Technology Special Projects (No. 2018ABA099).

References

[1] D.H. Cooper, T. Shiga, Discrete-matrix multichannel stereo, Journal of the Audio Engineering Society 20(5)(1972) 346-360.

[2] M.A. Gerzon, Periphony: with-height sound reproduction, Journal of the Audio Engineering Society 21(1973) 2-10.

[3] M. Frank, F. Zotter, Exploring the perceptual sweet area in ambisonics, in: Proc. the 142nd Audio Engineering Society Convention, 2017.

[4] A.J. Berkhout, A holographic approach to acoustic control, Journal of the Audio Engineering Society 36(12)(1988) 977-995.

[5] J.-M. Lee, J.-W. Choi, Y.-H. Kim, Integral formula for reproducing higher order sound field with polar phase variation, in: Proc. the 22nd International Congress on Sound and Vibration, 2015.

[6] S. Koyama, K. Furuya, Y. Hiwasaki, Y. Haneda, Y. Suzuki, Wave field reconstruction filtering in cylindrical harmonic domain for with-height recording and reproduction, IEEE/ACM Transactions on Audio, Speech, and Language Processing 22(10)(2014) 1546-1557.

[7] H. Wierstorf, C. Hold, A. Raake, Listener preference for wave field synthesis, stereophony, and different mixes in popular music, Journal of the Audio Engineering Society 66(5)(2018) 385-396.

[8] J. Blauert, Spatial Hearing, MIT Press, Cambridge, MA, 1983.

[9] E. Rasumow, M. Blau, M. Hansen, S. Par, S. Doclo, V. Mellert, D. Püschel, Smoothing individual head-related transfer functions in the frequency and spatial domains, Journal of the Acoustic Society of America 135(4)(2014) 2012-2025.

[10] H.M. Fayek, L. Maaten, G.D. Romigh, R. Mehra, On data-driven approaches to head-related transfer function personalization, in: Proc. the 143rd Audio Engineering Society Convention, 2017.

[11] V. Pulkki, Virtual sound source positioning using vector base amplitude panning, Journal of Audio Engineering Society 45(6)(1997) 456-466.

[12] V. Pulkki, Localization of amplitude-panned virtual sources II: two-and three-dimensional panning, Journal of Audio Engineering Society 49(9)(2001) 753-767.

[13] A. Franck, W. Wang, F.M. Fazi, Sparse l1-optimal multiloudspeaker panning and its relation to vector base amplitude panning, IEEE/ACM Transactions on Audio, Speech, and Language Processing 25(5)(2017) 996-1010.

[14] A. Ando, Conversion of multichannel sound signal maintaining physical properties of sound in reproduced sound field, IEEE Transactions on Audio, Speech, and Language Processing 19(6)(2011) 1467-1475.

[15] J. Seo, J. Yoo, K. Kang, F.M. Fazi, 21-Channel surround system based on physical reconstruction of a three dimensional target sound field, in: Proc. 128th International Convention of the Audio Engineering Society, 2010.

[16] M. Shin, P.A. Nelson, F.M. Fazi, J. Seo, Velocity controlled sound field reproduction by non-uniformly spaced loudspeakers, Journal of Sound and Vibration 370(2016) 444-464.