Sound field synthesis based on superposition of multipoles comprising focused monopole sources

Kimitaka Tsutsumi\textsuperscript{1,2,*}, Kenta Imaizumi\textsuperscript{1,2}, Yoichi Haneda\textsuperscript{2} and Hideaki Takada\textsuperscript{1}

\textsuperscript{1}NTT Service Evolution Laboratories, NTT Corporation, 1–1, Hikarinooka, Yokosuka, 239–0847 Japan
\textsuperscript{2}The University of Electro-Communications, 1–5–1, Chofugaoka, Chofu, 162–8585 Japan

(Received 16 October 2018, Accepted for publication 29 September 2019)

Abstract: We propose a method to create a directional sound source in front of a linear loudspeaker array. The method creates clusters of focused sources to form multipoles by using a linear loudspeaker array and superposes the multipoles to synthesize a directivity pattern. We also derive an efficient multipole structure in which adjacent lower order multipoles are overlapped. The structure reduces the number of focused sources, thereby reducing the algorithmic complexity needed to create them. To further reduce complexity, we also derive a time domain implementation of the proposed method. To mitigate degradation in the reproduced directivity due to superposition of the inaccurate sound fields of focused sources, a fractional delay interpolation is applied. Computer simulation results indicate that the proposed method based on superposition of up to the third order multipoles creates a directional sound source at significantly lower complexity than a conventional method.

Keywords: Wave field synthesis, Multipole, Focused source, Directivity control, Linear loudspeaker array

PACS number: 43.60.Fg [doi:10.1250/ast.41.489]

1. INTRODUCTION

Spatial sound reproduction methods are one of the keys for providing high reality to audiences in theaters and live events since they reproduce complex sound scenes with freely movable acoustic sources. In many cases, implemented methods of this type are based on Ambisonics [1,2] or Wave Field Synthesis (WFS) [3,4]. In contrast to surround-sound reproduction methods, these methods create a sound source at an arbitrary distance as well as an arbitrary direction.

In recent live events, virtual sound sources created between an audience and loudspeakers by the focused source method [5] are being used to provide sound effects closing in upon the audience in a venue [6]. It is also known that an audience feels more realistic sound images by adding directivity patterns: the angular dependence of the strength of the radiated power from a sound source [7]. These factors demonstrate that it is more important than ever to develop a sound reproduction system that controls both distance and directivity of virtual sound images.

Researchers have been investigating methods to reproduce sound fields of directional sound sources for some time. Among them are a method that captures the sound field as a whole by using a microphone array [8], a method based on driving functions analytically derived from an angular spectrum [9], an analytical method derived from circular harmonics [10–13], and a method based on multipole superposition [14,15].

Analytical methods based on angular spectra or circular harmonics accurately reproduce sound fields even if they are created by directional sound sources [9,10]. Some of them also derived efficient time domain implementations [10]. These time domain implementations are approximately yielded by weighing and delaying the input signal and then applying filters that capture spatiotemporal characteristics of the sound field, thereby significantly reducing the algorithmic complexity. Since the filters differ between each loudspeaker, these operations need to be performed as many times as the number of loudspeakers. In addition, the filter updates for moving sources further increase computational complexity. This overhead has been introduced because the definition of the spatiotemporal filter includes the relative position of the moving source from each loudspeaker.

*e-mail: kimitaka.tsutsumi.tb@hco.ntt.co.jp
Multipole-based methods have been presented in existing researches [14,15]. [15] presented an analytical method of obtaining weighting coefficients for each multipole. Although directivity patterns were reproduced by the method, the phase of the original directional sound source was not reproduced; therefore, the sound field was degraded. Some methods can be converted into efficient time domain implementation [14]. In this method, a spatiotemporal filter is applied to reproduce the directivity pattern and the rotation or shift of a directional source is obtained by controlling the positions of focused sources in the multipoles. This spatiotemporal filter avoids the redundant filter update operations necessary to obtain the rotation or shift of a directional source, thereby further reducing algorithmic complexity.

A drawback of the multipole-based method is that the number of focused sources increases as the order of multipoles goes up [16]. This causes additional complexity even in the time domain.

To prevent boosting of the number of monopole sources in multipoles, a method with an overlapping multipole structure has been proposed [16]. Since the method is applicable only to one-dimensional cases, this paper reports how we expanded it to make it applicable to two-dimensional cases. In time domain implementation of the focused source method [5], fractional delay interpolation has been applied to accurately reproduce the sound field of a focused source [17–19]. In the work we report in this paper, we applied the interpolation, achieved by a filter known as the Thiran filter, to the proposed method to improve the accuracy of the directivity pattern.

Notably, due to the nature of sound focusing and the finite length of secondary sources in actual implementations [5], the focused source method typically utilizes limited listening areas where the virtual source can be considered as a monopole. Therefore, we focused on creating the sound field of a directional sound source in the listening area of a focused source in 2D space. In contrast to an existing method [15], our method can reproduce both directivity patterns and phases of the target sound field to reproduce an accurate sound field.

The remainder of this paper is organized as follows. Section 2 briefly introduces the theory and implementations of WFS and the focused source method. The proposed method is described in detail in Sect. 3. Since the method reproduces the sound field based on multipoles of a cluster of focused sources, we introduce multipole sources briefly. We also explain a fractional delay filter for accurate sound field reproduction in an efficient time domain implementation in Sect. 4. In Sect. 5, the accuracies of reproduced sound fields and computational complexities are estimated and discussed. Finally, conclusions are drawn in Sect. 6.

2. WAVE FIELD SYNTHESIS

Before describing details of the proposed method, the underlying theory of wave field synthesis [4] is briefly described in this section. We also introduce a conventional method [10] to reproduce the sound field of a directional sound source. We call this conventional method “the reference method” hereafter.

2.1. The Driving Function [4]

Wave field synthesis aims at reproducing an arbitrary sound field by using secondary sources. It is implemented by using a linear distribution of secondary sources. Based on the first Rayleigh integral, the sound pressure is given by

\[ P(x, \omega) = -\int_{-\infty}^{\infty} D_{2D}(x_0, \omega) G_{2D}(x - x_0, \omega) dx_0, \]

where \( x = (x, y) \) with \( y > 0 \) and \( x_0 = (x_0, y_0) \) for the position of a secondary source. The term \( D_{2D} \) denotes a driving function of a secondary source placed at \( x_0 \). \( G_{2D} \) is the two-dimensional Green’s function, and \( \omega \) denotes angular frequency. Under the assumption that secondary sources are distributed along the \( x \)-axis, the driving function is given by

\[ D_{2D}(x_0, \omega) = 2 \frac{\partial}{\partial y} S(x, \omega) \bigg|_{x=x_0}, \]

where \( S(x, \omega) \) denotes the desired sound pressure to be reproduced. The sound field created by a secondary source is theoretically given by the two-dimensional free field Green’s function.

2.2. Focused Source Method [5]

WFS provides the ability of creating virtual point sources in between a loudspeaker array and the audience. These sources are known as focused sources due to their relation to acoustic focusing. The focused source method assumes an acoustic sink defined at the position of a virtual point source [5]. Since a sound emitted by a secondary source travels towards the virtual point source, converging and diverging parts are created in the resulting sound field. As a result, the diverging part of the reproduced sound field corresponds to the desired sound field of a virtual point source as depicted in Fig. 1. The desired sound field is given by

\[ S_{2D}(x, \omega) = \frac{j}{4} \sqrt{k} H_0^{(1)}(k|x_0 - x|), \]

where \( H_0^{(1)} \) denotes the Hankel function of the first kind of zero-th order and \( x = (x_s, y_s) \) is the position of the focused source for \( y_s > 0 \). \( j = \sqrt{-1} \) is the imaginary unit, \( k = \omega/c \) denotes the wave number, and \( c \) denotes the speed of
sound. Introducing (3) into (2) results in the following driving function:

\[ D_{2D}(x_0, x_s, \omega) = -\frac{(jk)^{3/2}}{2} \frac{y_0 - y_s}{|x_0 - x_s|} H_1^{(1)}(k|x_0 - x_s|), \]  

(4)

where \( H_1^{(1)} \) denotes the Hankel function of the first kind of the first order.

A problem with the focused source method with respect to listening area is that people in the audience who are sitting in corner seats cannot find sound images at the defined points. This is caused by the limited length of loudspeaker arrays in actual implementations [5]. A listening area where the audience perceives the virtual sound image at the defined point is depicted in Fig. 2. As can be inferred from the figure, the listening area narrows if the focused source moves away from the secondary sources.

2.3. A Time Domain Implementation of the Focused Source Method [5]

The driving function defined in (4) can be transformed into time domain to obtain an efficient implementation following a large value approximation of the Hankel function. The driving function in time domain is given as

\[ u(x_m, x_s, t) = h(t) \ast \left\{ \frac{y_s}{r_m} \delta\left( t - \frac{r_m}{c} \right) \right\}, \]  

(5)

where \( \ast \) denotes convolution, \( t \) denotes time, \( c \) denotes the speed of sound, \( y_s \) is the distance between a focused source and a linear loudspeaker array, and \( r_m \) is the distance between the focused point and the \( m \)-th secondary source. \( h(t) \) is a WFS prefilter defined using an inverse Fourier transformation as

\[ h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\omega}{c} e^{i\omega t} d\omega \]  

(6)

The WFS prefilter is independent of the position of secondary sources. Therefore, the driving function in time domain is approximately yielded as weighting and delaying for each secondary source after applying the WFS prefilter to the input signal, thereby reducing computational complexity by avoiding inverse Fourier transform of the output signal filtered in frequency domain.

2.4. A Method to Create the Sound Field of a Directional Source [10]

A driving function to create the sound field of a directional source was derived on the basis of the Rayleigh integral and circular harmonic expansion of the sound field [10]. This method can also be applied to a focused source [13]. An arbitrary sound field can be expressed by circular harmonic expansion as follows:

\[ S(x) = \sum_{i=-\infty}^{\infty} \tilde{S}_v^{(2)}(\omega) H_v^{(2)}(kr) e^{i\alpha}, \]  

(7)

where \( \tilde{S}_v^{(2)} \) denotes harmonic expansion coefficients of order \( v \), \( H_v^{(2)} \) is the \( v \)-th order Hankel function of the second kind, and \( \alpha \) is the azimuth of an arbitrary position \( x \). The geometry system is depicted in Fig. 3. By using the Rayleigh first integral to the sound field of (7), the driving function of a secondary source at \( x \) is derived as
The time domain driving function of the reference method is defined as

\[ D(x) = jk \sum_{\nu=-\infty}^{\infty} \tilde{S}^{(2)}(\nu, \omega) \left( H^{(2)}_{\nu-1} e^{j(v-1)\omega} + H^{(2)}_{\nu+1} e^{j(v+1)\omega} \right) \]  

The time domain driving function of the reference method is given by the following equation [20,21],

\[ u_{\text{ref}}(x_0, t) = 2 \sqrt{\frac{2}{\pi r_0}} \cdot \delta \left( t - \frac{r_0}{c} \right) \otimes h'(t) \otimes \tilde{S}^{(2)}(\alpha_0, t), \]  

where \( h'(t) = F^{-1}[\sqrt{\omega/jc}] \) is a WFS prefilter, and \( \tilde{S}^{(2)}(\alpha_0, t) \) is a time domain correspondence to the plane wave decomposition of the sound field of a directional source as

\[ \tilde{S}^{(2)}(\alpha_0, t) = F^{-1} \left[ \sum_{\nu=\infty}^{\infty} \delta \cdot \tilde{S}^{(2)}(\nu, \omega) \cdot e^{j\nu \theta_0} \right]. \]  

This filter is efficiently obtained by inverse discrete Fourier transformation (DFT) of circular harmonic modes \( \tilde{S}^{(2)}(\nu, \omega) \) followed by an interpolation. This interpolation is necessary because \( \alpha_0 \) are not equidistantly located on the angle depicted in Fig. 3, while the DFT assumes equidistant locations. The block diagram of the entire process is depicted in Fig. 4. By assuming that the directivity pattern itself does not change and only shifts and rotations occur, inverse DFT can be avoided every time the sound source moves. This saves a lot of computational complexity for moving sources. Rotation of the sound source is efficiently implemented by a circular shift of circular harmonic modes \( \tilde{S}^{(2)}_\nu \) followed by an inverse Fourier transform.

### 3. PROPOSED METHOD

In this section, the proposed method is described. The method is based on multipoles created by a sum of focused sources. The concept of this method is depicted in Fig. 5. We first introduce multipoles briefly and then explain the proposed method in detail. In general, a higher order multipole has a larger number of monopole sources. Thus directivity control based on multipole superposition requires a lot of computational complexity. An efficient structure based on overlapped multipoles is derived. To further reduce the complexity, an efficient time domain implementation is also described.
in (12) are given as follows: \( N = 2, (g_s^{(0)}, g_s^{(1)}) = (-1, 1), \)
\( x_s^{(0)} = (x_c, y_c - \Delta), x_s^{(1)} = (x_c, y_c + \Delta), \) where \( x_c = (x_c, y_c) \)
the coordinate of the center of the multipole, and \( \Delta = d/2, \) a half of the distance between monopole sources.

Focused sources can be created efficiently with low computational complexity by (5). By using this implementation, the multipoles comprising a cluster of focused sources are also efficiently implemented. From (5) and (12), the driving function for the \( m \)-th secondary source is given as
\[
\begin{align*}
  u_m(t) &= h(t) \sum_{i=0}^{N-1} \frac{g_s^{(i)}, y_s^{(i)}}{r_i^{3/2}} \delta(t - \frac{r_i}{c}), \\
&\text{(13)}
\end{align*}
\]
where \( y_s^{(i)} \) denotes the distance between the \( i \)-th focused source in a multipole and the \( x \)-axis and \( g_s^{(i)} \) denotes the sign of the \( i \)-th focused source. The number of focused sources in a multipole is \( N \) and \( r_i \) denotes the distance between the \( i \)-th focused source and the \( m \)-th secondary source.

### 3.3. An Efficient Structure of Multipoles

Multipoles can be defined on the basis of superposition of higher order partial derivatives of the Green’s function. An \( N \)-th order multipole is implemented by placing one \( (N-1) \)-th order multipole close to another with opposite phase. In this implementation, as many as \( 2^N \) monopole sources are needed to implement an \( N \)-th order multipole as depicted in Fig. 7. The number of point sources, therefore, becomes exponentially larger as the order of the multipole increases. To overcome this problem, a method to reduce the number of point sources in multipoles was proposed [22]. We briefly explain the method.

As illustrated in Fig. 8, an \( N \)-th order multipole can be implemented with \( N + 1 \) point sources by overlapping \( N - 1 \) point sources with the adjacent \( (N-1) \)-th order multipole. On top of the overlapped implementation, further reduction is possible by choosing the interval between two monopole sources nearest to the center of a multipole that has the order of odd number as twice the interval of other monopoles. This structure is shown in Fig. 9. Taking a dipole as an example, the interval between two monopoles in a dipole is chosen to be twice the interval between monopole sources in the second order multipole, by which only three point sources can reproduce a sound field obtained by superposition of a monopole, a dipole, and a quadrupole.

This method can be extended to the 2D case as shown in Fig. 10. Multipole superposition up to the fourth order can be achieved with only 21 point sources, while 129 sources are needed without the efficient method.

### 4. FRACTIONAL DELAY INTERPOLATION

Although the focused source method in time domain is efficient in terms of computational complexity, the sound field reproduced by time domain implementation is less accurate than the field reproduced by frequency domain implementation. This is because delay operation based on multiples of the sampling period is less accurate than the frequency domain implementation. For example, a system with the Nyquist frequency of 48 kHz introduces the maximum error of up to 7.1 mm (approximately equal to the sound speed of 340 m/s divided by 48,000 samples/s). To mitigate this artifact, we investigate methods of
fractional delay interpolation for accurate reproduction of the sound field.

4.1. Fractional Delay Filters

A fractional delay filter introduces a delay shorter than the sampling period. The ideal fractional delay filter is defined by the sinc function. This filter is theoretically a non-causal filter of infinite length. In practical implementations, the filter must be discontinued with a finite length and delayed so that the filter can be used as a causal filter. Although this filter exhibits an all-pass characteristic in frequency response and a linear phase characteristic, it cannot be defined with a small number of coefficients, thereby introducing a lot of computational complexity.

To achieve delay interpolation with lower complexity, alternative implementations of fractional delay filters were proposed. An example is an FIR filter based on Lagrange interpolation [23]. Since it gives filter coefficients by closed form formulas, it can be efficiently implemented for variable delay applications. Although it achieves flat group delay, Lagrange interpolation of lower order gives a low pass characteristic, thereby degrading speech quality. As the order increases, the low pass characteristic approaches a flat characteristic. However, applying a filter with higher order degrades efficiency since the speed of the characteristic evolution is slow compared to the increase in the order and its computational complexity.

Another example of a fractional delay filter is an IIR filter known as the Thiran filter [24]. This IIR all-pass filter gives flat response over all frequencies as well as a lower number of operations than FIR filters. The coefficients are defined as

$$p_{Th}(t, t_{frac}) = (-1)^j \binom{M}{t} \prod_{l=0, l \neq j}^{M} \frac{t_{frac} - M - l}{t_{frac} - M + l + t}, \quad (14)$$

where $t_{frac}$ denotes the fractional delay and $M$ is the number of filter taps. Despite its flat frequency response, the group delay characteristic is not flat over all the frequency region as shown in Fig. 11. Especially at high frequency region, errors from target fractional delay are

```
N=0  N=1  N=2  N=3  N=4
```

Fig. 10 Positions and weights of monopole sources in each multipole in a 2D case. The distance between adjacent monopole sources (dotted circles) is $d$.  

```
Fig. 11 Group delay of the first order Thiran filter up to 24kHz (0.5 in normalized frequency). $N$ denotes the target group delay.
```

```
N=0  N=1  N=2  N=3  N=4
```

Fig. 10 Positions and weights of monopole sources in each multipole in a 2D case. The distance between adjacent monopole sources (dotted circles) is $d$.  

```
Fig. 11 Group delay of the first order Thiran filter up to 24kHz (0.5 in normalized frequency). $N$ denotes the target group delay.
```

Fig. 10 Positions and weights of monopole sources in each multipole in a 2D case. The distance between adjacent monopole sources (dotted circles) is $d$.  

```
Fig. 11 Group delay of the first order Thiran filter up to 24kHz (0.5 in normalized frequency). $N$ denotes the target group delay.
```

Fig. 10 Positions and weights of monopole sources in each multipole in a 2D case. The distance between adjacent monopole sources (dotted circles) is $d$.  

```
Fig. 11 Group delay of the first order Thiran filter up to 24kHz (0.5 in normalized frequency). $N$ denotes the target group delay.
```

Fig. 10 Positions and weights of monopole sources in each multipole in a 2D case. The distance between adjacent monopole sources (dotted circles) is $d$.  

```
Fig. 11 Group delay of the first order Thiran filter up to 24kHz (0.5 in normalized frequency). $N$ denotes the target group delay.
```
much larger than those by Lagrange interpolation. However, it can be considered to have a linear phase characteristic only at a frequency region lower than 2 kHz as shown in Fig. 11.

4.2. Time Domain Implementation with the Thiran Filter

Considering the fact that a sound field reproduced at frequencies higher than 2 kHz is not accurate due to spatial aliasing and speech quality degraded by the low pass characteristic of Lagrange interpolation, we applied the Thiran filter to a time domain driving function for higher frequencies higher than 2 kHz is not accurate due to spatial aliasing and speech quality degraded by the low pass filter.

Fig. 12 The block diagram of the proposed method in time domain implementation with the Thiran filter.

5. PERFORMANCE EVALUATION

We used computer simulations to analyze the performance of our method in terms of accurately reproducing sound fields and directivity patterns, as well as computational complexity. We compared the performance of our method with and without the efficient implementation of multipoles described in Sect. 3.3, the efficient time domain implementation with the Thiran filter described in Sect. 4.2, and the frequency and time domain implementations of the reference method.

5.1. Experimental Setup

In this section, we summarize the experimental setup for our experiments. We implemented simulations with MATLAB. We set the speed of sound at 343.36 m/s. We arranged a linear loudspeaker array of 65 loudspeakers along the x-axis where $-1.6 \leq x \leq 1.6$. We set the adjacent actual loudspeakers 0.05 m apart. We set the interval of monopole sources (focused sources) in multipoles at 0.01 m. The center of the multipole source formed by focused sources was $(0, 0.5)$. Positions of monopole sources in multipoles were computed using the center position of the multipole and relative coordinates depicted in Fig. 10. For example, positions of monopole sources corresponding to $\frac{2}{3\pi}$ were $(0, 0.51)$ and $(0, 0.49)$. We chose $N = 4$ as the maximum order of multipoles. We generated circular harmonics $\tilde{S}_{21}(\omega)$ as $2N + 1$ complex random numbers having amplitudes less than or equal to 1.0. We computed the original sound field using (7). We obtained the weighting coefficients for each multipole by the least square error method using the following equation [25].

$$w = (G^H G + \lambda I)^{-1} G^H s,$$  \hspace{1cm} (15)

where $w$ is a vector of weighting coefficients for each multipole as $w = [w_{0,0}, \ldots, w_{0,N}]^T$. The superscript $[\cdot]^H$ is the Hermitian transpose of a matrix. $\lambda$ is a regularization parameter to prevent the inverse matrix of $G$ from becoming unstable [26]. $s$ is the vector of sound pressures observed at controlling points placed along a unit circle having a target directional sound source at its center (180 points with an interval of $2^\circ$ between adjacent controlling points). The radius was 1 m, and the transfer function matrix $G = [v_{0,0}, \ldots, v_{0,N}]$ whose columns $v_{m,n}$ were defined at the same controlling points [16] by the following equation.

$$v_{m,n}(r, k) = G_{2D}(kr)(jk)^{m+n} \cos^m \phi \sin^n \phi. \hspace{1cm} (16)$$

5.2. Objective Measures

We used the following two objective measures to evaluate the effectiveness of our method. The first evaluated the accuracy of the reproduced sound field.

$$Err_{SF}(x, y) = 10 \log 10 \left( \frac{|s_{org}(x, y) - s(x, y)|^2}{|s_{org}(x, y)|^2} \right), \hspace{1cm} (17)$$

where $s_{org}(x, y)$ and $s(x, y)$ are respectively the original and the synthesized sound pressures at position $(x, y)$. We evaluated this synthesis error $Err_{SF}(x, y)$ at the region of $-1.5 \leq x \leq 1.5$ and $1.0 \leq y \leq 3.0$ m. We set the distance between adjacent evaluation points at 0.01 m. The second objective measure evaluated the accuracy of the reproduced directivity of a target sound source. This objective measure focuses on the accuracy of radiated power for each direction rather than the phase of reproduced sound fields.

$$Err_{Dif}(x, y) = 10 \log 10 \left( \frac{|s_{org}(x, y) - s(x, y)|^2}{|s_{org}(x, y)|^2} \right), \hspace{1cm} (18)$$

where $s_{org}(x, y)$ and $s(x, y)$ are respectively the original and the synthesized sound pressures at the positions of $(x, y) = (xc + \cos \phi, yc + \sin \phi)$. $(xc, yc)$ is the position of the center of the multipole, and we chose a 1-m radius for the unit circle. We set the distance between adjacent evaluation points at $1^\circ$. 

495
5.3. Performance Dependency on the Interval of Monopole Sources in Multipoles

First, we evaluated the performance dependency on the interval between adjacent monopole sources in multipoles. We used four distances (0.01 m, 0.02 m, 0.04 m, and 0.07 m) for the evaluation. Figures 13 and 14 show the averaged synthesis error of the directivity patterns and the reproduced sound fields, respectively. Both results show that shorter intervals of adjacent monopole sources provide higher accuracies of both the reproduced sound field and the directivity patterns. Considering that the spatial Nyquist frequency is around 3.4 kHz, the interval of 0.04 m provides accuracy equal to that achieved by the intervals of 0.01 m or 0.02 m. The results confirmed that it is easier to only reproduce directivity patterns rather than reproduce sound fields by using the wide interval between monopole sources.

5.4. Accuracy of Reproduced Sound Fields and Directivity Patterns

We performed a simulation to reproduce sound fields of a target directional source by using a 1,992.2 Hz monochromatic sine wave. We also reproduced the sound fields created when the source rotated in a counterclockwise direction by \(10^\circ\). The results are shown in Fig. 15 (without rotation) and Fig. 16 (with rotation). From Fig. 15 (a) and (b) (c), in the frequency implementation, our method reproduced accurate sound fields in the listening areas of focused sources created by the loudspeaker array of finite length. Even when the efficient implementation of multipoles described in Sect. 3.3 was used, Fig. 15 (d) (e) confirmed that the accuracy was comparable with our method without efficient implementation. From Fig. 15 (f) (g), it was difficult to reproduce the sound field of the target directional source by using the time domain implementation without delay interpolation; however, Fig. 15 (h) (i) confirmed that the accuracy of the reproduced sound field can be improved by delay interpolation based on the Thiran filter. We also observed similar tendencies when the target directional source rotated by Fig. 16. Averaged synthesis errors regarding both the directivity patterns and the sound fields computed up to 3.5 kHz (around the spatial Nyquist frequency) are plotted in Fig. 17. The figure confirmed that the accuracies of both the directivity and the reproduced sound field by frequency domain implementation of our method were comparable to those by the implementation of the reference method. Regarding time domain implementation, we confirmed that the averaged synthesis error of reproduced directivity by our method was smaller than that by the reference method on average. In the frequency region lower than 600 Hz, we confirmed that the averaged synthesis error of the directional source when the maximum order of circular harmonics and multipoles were set as \(N = 3\) were plotted in Fig. 18. We confirmed that the tendency was similar to the case of the maximum order of \(N = 4\).
Fig. 15 Reproduced sound fields (Left figures) and error distributions from the sound field of the original sound source. (a): original (right figures). (b)(c): proposed method in the frequency domain. (d)(e): proposed method with the efficient implementation in the frequency domain. (f)(g): proposed method with the efficient implementation in the time domain. (h)(i): proposed method modified by the Thiran filter with the efficient implementation in the time domain.

Fig. 16 Reproduced sound fields (Left figures) and error distributions from the sound field of the original sound source rotated toward counter-clockwise direction by \( \phi = 10^\circ \). Methods to plot these figures correspond to those used in Fig. 15.
In calculating the computational complexity, we implemented both our method and the reference method in C language and then incorporated the floating-point complexity counter provided in ITU-T G.192 [27].

In this simulation, we assumed that only shift and rotation of the target directional source occurred every 20 ms. For both the reference method and the proposed method, we chose to use 65 loudspeakers and chose respectively 513 and 129 as the length of the WFS prefilter and the spatiotemporal filter. Prior to the experiment, the spatiotemporal filters for each method were calculated and stored in the memory.

Table 1 shows the complexity of each method. First, we estimated the complexities to create monopole sources in the frequency domain and in the time domain. We confirmed that the complexity of a monopole source in the

5.5. Computational Complexity

In calculating the computational complexity, we implemented both our method and the reference method in C language and then incorporated the floating-point complexity counter provided in ITU-T G.192 [27].

In this simulation, we assumed that only shift and rotation of the target directional source occurred every 20 ms. For both the reference method and the proposed method, we chose to use 65 loudspeakers and chose respectively 513 and 129 as the length of the WFS prefilter and the spatiotemporal filter. Prior to the experiment, the spatiotemporal filters for each method were calculated and stored in the memory.

Table 1 shows the complexity of each method. First, we estimated the complexities to create monopole sources in the frequency domain and in the time domain. We confirmed that the complexity of a monopole source in the
time domain was only 3.43% compared to that of the frequency domain. Since the time domain implementation is much more efficient than the frequency domain implementation, we computed only the complexities of time domain implementations. The results showed us that complexities of the reference and our method up to the forth order were lower than what was needed to create a monopole source in the frequency domain. The complexity of our method increased as the maximum order of multipoles went up, while that of the reference method did not change as the maximum order of circular harmonics increased. This is because the length of the spatiotemporal filter for the reference method does not change if the order of circular harmonics changes.

The proposed method had complexity lower than that of the reference method up to the third order. With the third order multipole the proposed method achieved complexity 29.5% less than that of the reference method.

### 6. CONCLUSION

In this paper, we proposed a method of creating a sound field of a directional sound source in front of a linear loudspeaker array. The method was based on multipoles formed by multiple focused sources created by a linear loudspeaker array. We also applied an efficient implementation of multipoles by overlapping the adjacent multipoles. To further reduce complexity, we derived a time domain implementation. Since the implementation created multiple focused sources close to each other, errors of delays shorter than the sampling period degraded the accuracy of the reproduced sound field. To mitigate the accuracy degradation, we applied a fractional delay filter known as the Thiran filter and confirmed that we could improve the accuracy of the reproduced sound field. We also confirmed that complexities of our method were lower than those of the conventional method up to the third order multipole superposition in time domain implementation.

### REFERENCES

[1] M. A. Gerzon, “Periphony: With-height sound reproduction,” J. Audio Eng. Soc., 21, 2–10 (1973).
[2] M. A. Poletti, “Three-dimensional surround sound systems based on spherical harmonics,” J. Audio Eng. Soc., 53, 1004–1025 (2005).
[3] A. J. Berkhout, D. de Vries and P. Vogel, “Acoustic control by wave field synthesis,” J. Acoust. Soc. Am., 93, 2764–2778 (1993).
[4] J. Ahrens, R. Rabenstein and S. Spors, “The theory of wave field synthesis revisited,” Audio Eng. Soc. Conv. 124 (2008).
[5] S. Spors, H. Wierstorf, M. Geier and J. Ahrens, “Physical and perceptual properties of focused virtual sources in wave field synthesis,” Audio Eng. Soc. Conv. 127 (2009).
[6] K. Tsutsumi and H. Takada, “Powerful sound effects at audience seats by wave field synthesis,” NTT Tech. Rev., 15(12) (2017).
[7] K. Maki, T. Kimura and M. Katsumoto, “Reproduction of real musical instrument sounds by a spherical loudspeaker with 26 transducers: Acoustic comparison on reproducibility of sound radiation patterns of musical instruments,” J. Acoust. Soc. Jpn., 67, 66–69 (2007) (in Japanese).
[8] E. Corteel, “Synthesis of directional sources using wave field synthesis, possibilities, and limitation,” EURASIP J. Adv. Signal Process., Article No. 090509 (2007).
[9] T. Okamoto, “Angular spectrum decomposition-based 2.5D higher-order spherical harmonic sound field synthesis with a linear loudspeaker array,” Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, pp. 180–184 (2017).
[10] J. Ahrens and S. Spors, “Implementation of directional sources in wave field synthesis,” Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, pp. 447–458 (2008).
[11] J. Ahrens and S. Spors, “Rendering of virtual sound sources with arbitrary directivity in higher order ambisonics,” Audio Eng. Soc. Conv. 123. (2007).
[12] J. Ahrens and S. Spors, “Wave field synthesis of moving virtual sound sources with complex radiation properties,” J. Acoust. Soc. Am., 130, 2807–2816 (2011).
[13] J. Ahrens and S. Spors, “Notes on rendering of focused directional sound sources in wave field synthesis,” Proc. 34rd German Annu. Conf. Acoustics (DAGA), pp. 229–230 (2011).
[14] K. Tsutsumi, H. Takada and Y. Haneda, “Directivity control with multi-poles comprising plural focused sources in wave field synthesis,” Proc. Spring Meet. Acoust. Soc. Jpn., pp. 469–470 (2018) (in Japanese).
[15] K. Tsutsumi, H. Takada, K. Noguchi and Y. Haneda, “Directivity synthesis with multipoles comprising a cluster of focused sources using a linear loudspeaker array,” Proc. ICASSP, pp. 496–500 (2018).
[16] Y. Haneda, K. Furuya and S. Shimachi, “Directivity synthesis using multipole sources based on spherical harmonic expansion,” J. Acoust. Soc. Jpn., 69, 577–588 (2013) (in Japanese).
[17] A. Franck, K. Brandenburg and U. Richter, “Efficient delay interpolation for wave field synthesis,” Audio Eng. Soc. Conv. 125 (2008).
[18] F. Winter and S. Spors, “On fractional delay interpolation for local wave field synthesis,” Proc. 24th Eur. Signal Processing Conf., pp. 2415–2419 (2016).
[19] K. Tsutsumi, H. Takada and Y. Haneda, “Fractional delay filter for time domain implementation of directivity synthesis using multipoles comprising a cluster of focused sources,” Proc. Autumn Meet. Acoust. Soc. Jpn., pp. 489–490 (2018) (in Japanese).
Japanese).

[20] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography* (Elsevier Science, San Diego, 1999).

[21] P. A. Nelson and S. J. Elliott, *Active Control of Sound* (Academic Press, London, 1992).

[22] Y. Haneda, K. Furuya and H. Itou, “Design of multipole loudspeaker array based on spherical harmonic expansion,” *Proc. ICASSP*, pp. 141–144 (2011).

[23] T. I. Laakso, V. Valimaki, M. Karjalainen and U. K. Laine, “Splitting the unit delay [FIR/all pass filters design],” *IEEE Signal Process. Mag.*, 13, 30–60 (1996).

[24] J. P. Thiran, “Recursive digital filters with maximally flat group delay,” *IEEE Trans. Circuit Theory*, 18, 659–664 (1971).

[25] F. M. Fazi, M. Shin, F. Olivieri, S. Fontana and Y. Lang, “Comparison of pressure-matching and mode-matching beamforming for methods for circular loudspeaker arrays,” *Audio Eng. Soc. Conv.* 137 (2014).

[26] M. Shin, F. M. Fazi, P. A. Nelson and F. C. Hirono, “Controlled sound field with a dual layer loudspeaker array,” *J. Sound Vib.*, 333, 3794–3817 (2014).

[27] International Telecommunication Union (ITU), “ITU-T Recommendation G.192: A common digital parallel interface for speech standardization activities,” Geneva, (1996).