Audio Engineering Society

Convention Paper

Presented at the 13th Regional Convention
2007 July 19-21 Tokyo, Japan

Effects of Directivity of Microphones and Loudspeakers on Accuracy of Synthesized Wave Fronts in Sound Field Reproduction Based on Wave Field Synthesis

Toshiyuki Kimura[1], and Kazuhiko Kakehi[2]

[1] Universal Media Research Center, National Institute of Information and Communications Technology, Nukui-Kitamachi 4-2-1, Koganei, Tokyo, 184-8795, Japan
t-kimura@nict.go.jp

[2] School of Information Science and Technology, Chukyo University, Tokodachi 101, Kaizu-cho, Toyota, Aichi, 470-0393, Japan
kakehik@sist.chukyo-u.ac.jp

ABSTRACT

Sound field reproduction based on wave field synthesis is a technique that synthesizes wave fronts in a listening area using multiple microphones and loudspeakers placed at the boundary of the area. This study evaluates the effects of the directivity of microphones and loudspeakers on the accuracy of synthesized wave fronts. Three directional patterns (omnidirectional, unidirectional, and shotgun) were designed as directivity conditions of microphones and loudspeakers. The results of computer simulation show that there is almost no effect due to directivity of loudspeakers and that unidirectional or shotgun microphones can accurately reproduce wave fronts.

1. INTRODUCTION

Wave field synthesis [1-3] is a sound field reproduction technique that synthesizes wave fronts. The original sound is picked up by an array of microphones in a control area and is then reproduced in a listening area by an array of loudspeakers. The technique is based on Huygens principle. The arrays of microphones and loudspeakers are placed at the boundary of both areas. In this technique, multiple listeners can move about in a listening area or rotate their heads. This is unlike conventional sound field reproduction techniques, such as binaural [4] and transaural [5].

The theoretical background to wave field synthesis techniques was studied by Berkhout et al [3]. They assumed that the arrays of microphones and loudspeakers were placed in a line, as shown in Figure 1. On the other hand, surround systems such as 5.1ch and 7.1ch are developed for the home theater system [6] and surround systems based on wave field synthesis is proposed [7]. In cases, the arrays of microphones and loudspeakers are placed around the control area and the listening area, as in Figure 2.

The problem is that using the setup shown in Figure 2 can sometimes make the listeners feel they are in a reverberant sound field even though the original sound field is a free field—a field in which there is no reflected sound. This happens when the wave field of a free field is synthesized without using the directivity of microphones and loudspeakers so that the sound front comes from all directions in the listening area. If the setup is as in Figure 1, the sound front comes from only the front direction, so the reproduced sound field is not reverberant. Using the directivity of microphones and loudspeakers can solve the auditory problem inherent in surround systems set up as in Figure 2.

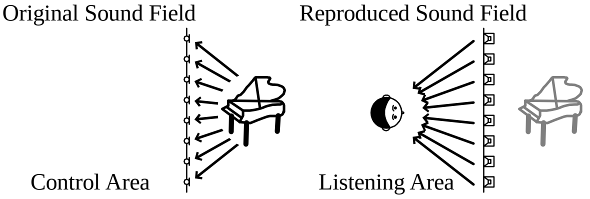

Figure 1 Setup used by Berkhout et al [3] for conventional study of original sound and reproduced sound fields

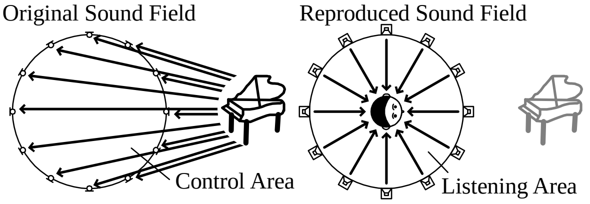

Figure 2 Original sound field and reproduced sound fields in a surround system

Although a wave field synthesis technique using directional microphones and loudspeakers was proposed by Camras [2], the effect of microphone and loudspeaker directivity was not investigated. We have investigated the effects using a computer simulation. We simulated two cases: one where the microphone and loudspeaker arrays were arranged in a circle (Section 2) and one where they were arranged in a square (Section 3).

2. COMPUTER SIMULATION—CIRCULAR AREA

2.1. Computer Simulation Conditions

The original sound field was assumed to be a free field. The original and reproduced sound fields in the computer simulation are shown in Figure 3. The shape of the control and listening areas is a circle of radius *r* meters. A sound source was placed at a position *d* meters from the center of the control area. Microphones and loudspeakers were arranged along the boundaries of the control and listening areas, respectively. The spacing of the microphones and loudspeakers was always constant and the positions of the loudspeakers relative to each other were the same as that of the microphones relative to each other. While the directivity of the microphones was toward the outside of the control area, the directivity of the loudspeakers was toward the inside of the listening area. This is the same setup used by Camras [2].

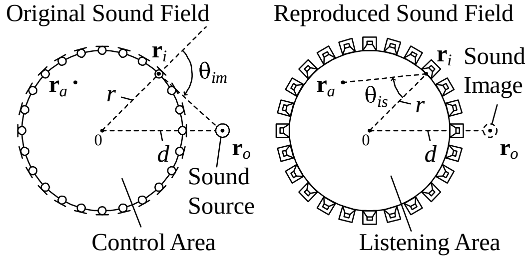

Figure 3 Original sound field and reproduced sound field using circular areas

The sound source signal was a sine-wave signal of frequency f (= $\sin 2\pi ft$). Let $\mathbf{r}_a$ be the position vector of an arbitrary point in the control area. Then, $p_o(\mathbf{r}_a, f, t)$ (sound pressure at $\mathbf{r}_a$ in original sound field) is denoted as follows,

$$p_o(\mathbf{r}_a, f, t) = \frac{1}{d_{ao}} \sin\left\{2\pi f\left(t - \frac{d_{ao}}{c}\right)\right\}, \quad (1)$$

where d_{ao} (= $\mathbf{r}_a - \mathbf{r}_o$) is the distance between the sound source and the arbitrary point, $\mathbf{r}_o$ is the position vector of the sound source, and c is the sound velocity. The $x_i(t)$ (channel signal of the ith microphone) is denoted as follows,

$$x_i(t) = \frac{D_{im}}{d_{io}} \sin\left\{2\pi f\left(t - \frac{d_{io}}{c}\right)\right\}, \quad (2)$$

where d_{io} (= $\mathbf{r}_i - \mathbf{r}_o$) is the distance between the sound source and the ith microphone, and $\mathbf{r}_i$ and D_{im} are the position vector and the directivity of the ith microphone (i = 1 ... M), respectively. Note that M is the total number of microphones. The $p(\mathbf{r}_a, f, t)$ (sound pressure of $\mathbf{r}_a$ in the reproduced sound field) is calculated as follows,

$$\begin{aligned} p(\mathbf{r}_a, f, t) &= \sum_{i=1}^{M} \frac{D_{is}}{d_{ai}} x_i\left(t - \frac{d_{ai}}{c}\right) \\ &= \sum_{i=1}^{M} \frac{D_{is} D_{im}}{d_{ai} d_{io}} \sin\left\{2\pi f\left(t - \frac{d_{ai} + d_{io}}{c}\right)\right\}, \end{aligned} \quad (3)$$

where d_{ai} (= $\mathbf{r}_a - \mathbf{r}_i$) is the distance between the ith loudspeaker and the arbitrary point, and D_{is} is the directivity of the ith loudspeaker.

Parametric conditions are shown in Table 1. The interval of microphones and loudspeakers is about 2 cm, which is less than half of the wavelength of 8000-Hz sound (=4.25 cm). Thus, the spatial sampling theorem to reproduce a wave front under 8000-Hz sound is satisfied.

Total Number (M)	630
Source frequency (f)	125, 177, 250, 354, 500, 707, 1000, 1414, 2000, 2828, 4000, 5657, 8000 Hz
Source distance (d)	3, 10, 100 m
Radius of areas (r)	2 m
Sound velocity (c)	340 m/s
Directivity (D_{im}, D_{is})	Omnidirectional, Unidirectional, Shotgun

Table 1 Parametric conditions when using circular areas

$\mathbf{r}_o$, $\mathbf{r}_a$, and $\mathbf{r}_i$ were set as follows,

$$\begin{aligned} \mathbf{r}_o &= \begin{pmatrix} d & 0 \end{pmatrix}^T \\ \mathbf{r}_a &= \begin{pmatrix} r_x & r_y \end{pmatrix}^T \\ \mathbf{r}_i &= \left(r\cos\frac{2\pi i}{M} \quad r\sin\frac{2\pi i}{M} \right)^T \end{aligned} \quad . \qquad (4)$$

Note that $r_x^2 + r_y^2 < r^2$.

Directional patterns of microphones and loudspeakers are shown in Figure 4.

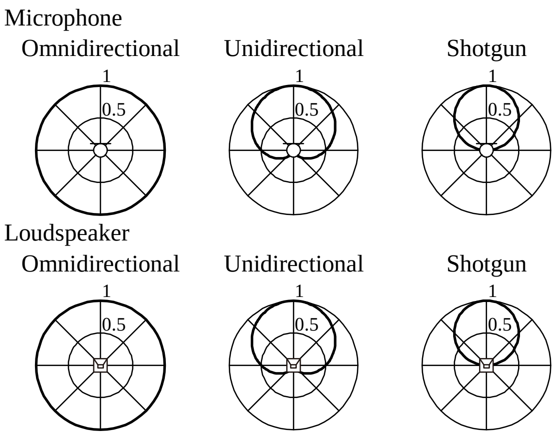

Figure 4 Directional patterns of microphones and loudspeakers

General microphones have the above directional patterns. For example, a pin microphone has an omnidirectional pattern, a dynamic microphone has a unidirectional pattern, and a shotgun microphone has a shotgun pattern. General loudspeakers have directional patterns that vary with the frequency of the output sound. The directional pattern varies in the following way: the directional pattern is omnidirectional for a low-frequency bandwidth, the directional pattern is unidirectional for a mid-frequency bandwidth, and the directional pattern is shotgun for a high-frequency bandwidth.

The D_{im} (directivity of the ith microphone) is defined as follows,

$$\begin{aligned} &\text{(Omnidirectional)} \quad D_{im} = 1 \\ &\text{(Unidirectional)} \quad D_{im} = \frac{1}{2}(1+\cos\theta_{im}) \\ &\text{(Shotgun)} \quad D_{im} = \begin{cases} \cos\theta_{im} & (|\theta_{im}| \le 90^\circ) \\ 0 & (|\theta_{im}| > 90^\circ) \end{cases} \end{aligned} \quad , \qquad (5)$$

where θ_{im} is the angle between $\mathbf{r}_{im}$ and $\mathbf{r}_{oi}$. Note that $\mathbf{r}_{im}$ (= $\mathbf{r}_i$) is the directional vector of the ith microphone and $\mathbf{r}_{oi}$ (= $\mathbf{r}_o - \mathbf{r}_i$) is the vector from the ith microphone to the sound source. The D_{is} (directivity of the ith loudspeaker) is defined as follows,

$$\begin{aligned} &\text{(Omnidirectional)} \quad D_{is} = 1 \\ &\text{(Unidirectional)} \quad D_{is} = \frac{1}{2}(1+\cos\theta_{is}) \\ &\text{(Shotgun)} \quad D_{is} = \begin{cases} \cos\theta_{is} & (|\theta_{is}| \le 90^\circ) \\ 0 & (|\theta_{is}| > 90^\circ) \end{cases} \end{aligned} \quad , \qquad (6)$$

where θ_{is} is the angle between $\mathbf{r}_{is}$ and $\mathbf{r}_{ai}$. Note that $\mathbf{r}_{is}$ (= $-\mathbf{r}_i$) is the directional vector of the ith loudspeaker and $\mathbf{r}_{ai}$ (= $\mathbf{r}_a - \mathbf{r}_i$) is the vector from the ith loudspeaker to the arbitrary point.

2.2. Results and Discussions

Results are shown in Figure 5 for simulations run with original wave fronts where f = 500 Hz. The synthesized wave fronts and the differences between the original and synthesized wave fronts at f = 500 Hz are shown in Figures 6–8, respectively. In these figures, absolute values ($|p_o|$, $|p|$, $|p - p_o|$) of pressures (p_o, p) and differences ($p - p_o$) are plotted and color bars are shown in the right of the figures. The wave fronts are accurately synthesized if the differences are white.

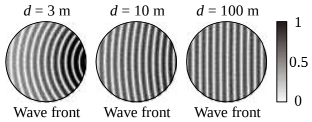

Figure 5 Wave fronts of original sound field when using circular areas (f =500 Hz)

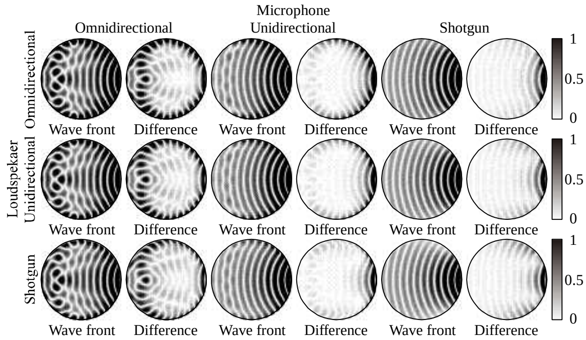

Figure 6 Synthesized wave fronts and differences when using circular areas (f = 500 Hz, d = 3 m)

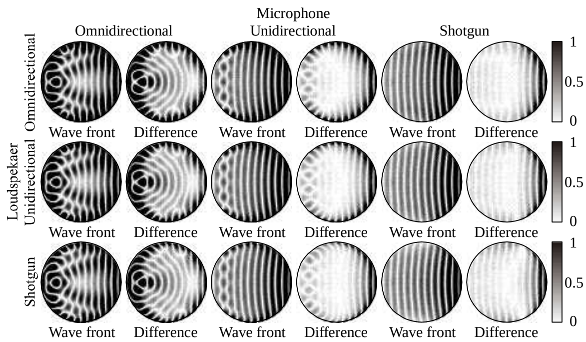

Figure 7 Synthesized wave fronts and differences when using circular areas (f = 500 Hz, d = 10 m)

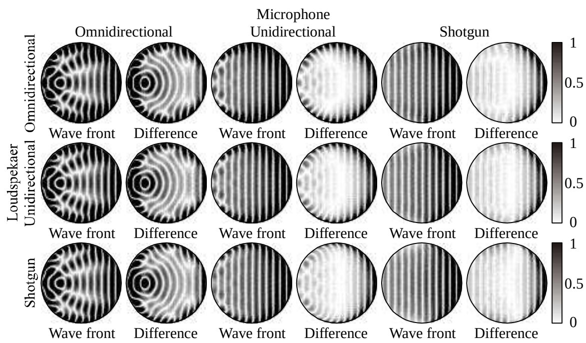

Figure 8 Synthesized wave fronts and differences when using circular areas (f = 500 Hz, d = 100 m)

When omnidirectional microphones were used, the wave fronts were not accurately reproduced because the differences were not white. This is due to the fact that all the microphones placed in the control area recorded the sound and the sound played from all directions in the listening area even though the sound was coming only from the sound source positioned at the front. When unidirectional and shotgun microphones were used, the wave fronts were accurately reproduced because the differences of the center of the circles were white. This is due to the fact that the directional microphones do not record sound coming from directions where there is no sound source.

Whatever the directivity of the loudspeakers, the results of the synthesized wave fronts were always same. This is due to the fact that the D_{is} does not vary when $\mathbf{r}_a$ is near the center of the circle according to Eq. (6). Thus, the directivity of a microphone contributes to the accuracy of synthesized wave fronts in surround systems based on wave field synthesis and the directivity of the loudspeakers does not.

In order to evaluate quantitatively the reproduced sound field, SNRs (signal-to-noise ratios) were calculated as follows,

$$\mathrm{SNR} = \frac{\sum_f \left[10\log_{10} \frac{\sum_{\mathbf{r}_a} \{p_o(\mathbf{r}_a, f, 0)\}^2}{\sum_{\mathbf{r}_a} \{p(\mathbf{r}_a, f, 0) - p_o(\mathbf{r}_a, f, 0)\}^2} \right]}{F} \quad ,(7)$$

where F(= 13) is the total number of frequencies. The range of $\mathbf{r}_a$ is limited to the inside of a circle of 1 m radius ($r_x^2 + r_y^2 < 1$). Note that $p_o(\mathbf{r}_a, f, 0)$ and $p(\mathbf{r}_a, f, 0)$ were normalized to $r_x^2 + r_y^2 < 1$ before calculating SNRs. SNR results for each directivity of microphones and loudspeakers are shown in Figure 9. Note that error bars denote the 95% confidence interval of SNRs. SNRs of unidirectional and shotgun microphones are higher than those of omnidirectional microphones under all source distance conditions. It is thus considered that the wave fronts can be accurately reproduced in surround systems using wave field synthesis if unidirectional and shotgun microphones are used.

SNRs of unidirectional and shotgun microphones were always more than 15 dB. These values are lower than the SNR value (about 30 dB) of other studies [8-9]. In our simulation, we controlled only the pressure at the boundary of the control and listening areas. In other studies, both the pressure and velocity at the boundary of these areas were controlled [8-9]. Thus, since the number of microphones and loudspeakers we used is half compared with other studies, our method of study is effective from a practical point of view.

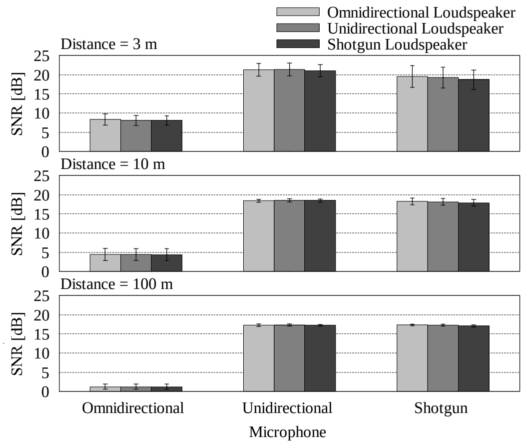

Figure 9 SNRs when using circular areas

The 95% confidence intervals of Figure 9 are the variance of SNRs in frequencies. In this simulation, the width of confidential intervals is less than 6 dB, while the average of SNRs is 15 dB when using directional microphones. Thus, the frequency has no effect on the accuracy of wave fronts.

3. COMPUTER SIMULATION—SQUARE AREA

3.1. Computer Simulation Conditions

The original sound field was assumed to be a free field. The original and reproduced sound fields simulated are shown in Figure 10. The shape of the control and listening areas is a square with sides of 2r meters in length. One sound source was placed at the position of d meters from the center of the control area and θ degrees azimuth from the center of the control area. Microphones and loudspeakers were arranged along the boundaries of control and listening areas, respectively. Note that the spacing of the microphones and loudspeakers was always constant and the positions of the loudspeakers relative to each other was the same as that of the microphones relative to each other. While the directivity of the microphones was toward the outside of the control area, the directivity of the loudspeakers was toward the inside of the listening area. This is the same setup used by Camras [2]. $p_o(\mathbf{r}_a, f, t)$ and $p(\mathbf{r}_a, f, t)$ (sound pressures in both sound fields) were calculated according to Eqs. (1) and (3).

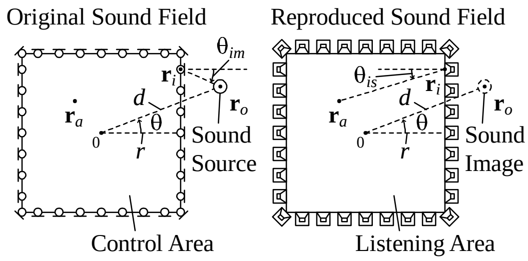

Figure 10 Original sound field and reproduced sound field using square areas

Parametric conditions are shown in Table 2. The interval of microphones and loudspeakers is 2 cm, which is less than half of the wavelength of 8000-Hz sound (= 4.25 cm). Thus, the spatial sampling theorem to reproduce the wave front under 8000-Hz sound is satisfied.

Total Number (M)	800
Source frequency (f)	125, 177, 250, 354, 500, 707, 1000, 1414, 2000, 2828, 4000, 5657, 8000 Hz
Source distance (d)	3, 10, 100 m
Source Azimuth (θ)	0, 45º
Radius of areas (r)	2 m
Sound velocity (c)	340 m/s
Directivity (D_{im}, D_{is})	Omnidirectional, Unidirectional, Shotgun

Table 2 Parametric conditions when using square areas

$\mathbf{r}_o$, $\mathbf{r}_a$, and $\mathbf{r}_i$ were set as follows,

$$\mathbf{r}_o = \left(d\cos\theta \quad d\sin\theta\right)^T$$
$$\mathbf{r}_a = \left(r_x \quad r_y\right)^T$$
$$\mathbf{r}_i = \begin{cases} \left(-r + \dfrac{ri}{100} \quad -r\right)^T & (i = 1-200) \\ \left(r \quad -r + \dfrac{r(i-200)}{100}\right)^T & (i = 201-400) \\ \left(r - \dfrac{r(i-400)}{100} \quad r\right)^T & (i = 401-600) \\ \left(-r \quad r - \dfrac{r(i-600)}{100}\right)^T & (i = 601-800) \end{cases}. \qquad (8)$$

Note that $|r_x|$, $|r_y| < r$.

The D_{im} and D_{is} (directivity of the ith microphone and the ith loudspeaker) were calculated according to Eqs.

(5) and (6). Note that the $\mathbf{r}_{im}$ and $\mathbf{r}_{is}$ (directional vector of the ith microphone and the ith loudspeaker) were defined as follows,

$$\mathbf{r}_{im} = \begin{cases} \left(\delta(i-200) \quad 1\right)^T & (i = 1-200) \\ \left(1 \quad \delta(i-400)\right)^T & (i = 201-400) \\ \left(-\delta(i-600) \quad 1\right)^T & (i = 401-600) \\ \left(-1 \quad -\delta(i-800)\right)^T & (i = 601-800) \end{cases}. \qquad (9)$$

$$\mathbf{r}_{is} = -\mathbf{r}_{im}$$

3.2. Results and Discussions

Results are shown in Figure 11 for simulations run where original wave fronts were f = 500 Hz. The synthesized wave fronts and the differences between the original and synthesized wave fronts at f = 500 Hz are shown in Figures 12–17, respectively. In these figures, absolute values ($|p_o|$, $|p|$, $|p - p_o|$) of pressures (p_o, p) and differences ($p - p_o$) are plotted and color bars are shown in the right of the figures. The wave fronts are accurately synthesized if the differences are white.

When unidirectional and shotgun microphones were used, wave fronts were accurately reproduced because the differences of the center of the squares were white. This is due to the fact that the directional microphone does not record sound coming from directions where there are no sound sources, as described in Section 2.

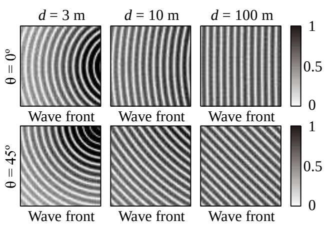

Figure 11 Wave fronts of original sound field when using square areas (f =500 Hz)

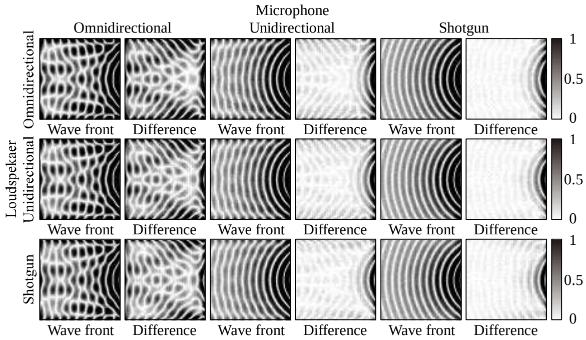

Figure 12 Synthesized wave fronts and differences when using square areas (f =500 Hz, d =3 m, θ =0º)

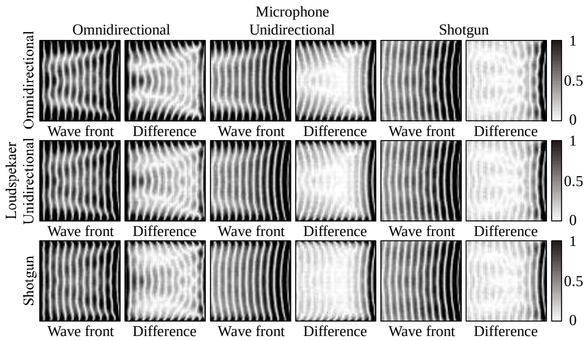

Figure 13 Synthesized wave fronts and differences when using square areas (f =500 Hz, d =10 m, θ =0º)

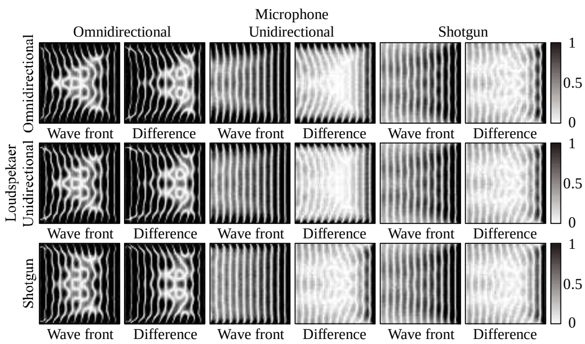

Figure 14 Synthesized wave fronts and differences when using square areas (f =500 Hz, d=100 m, θ=0º)

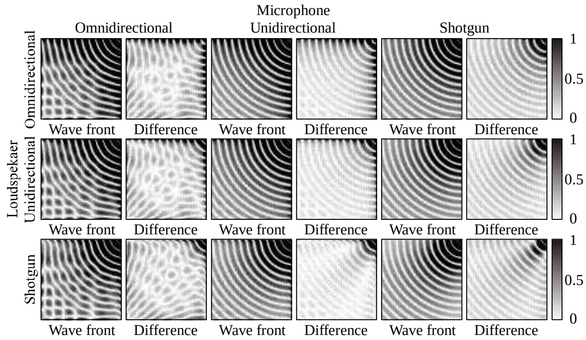

Figure 15 Synthesized wave fronts and differences when using square areas (f =500 Hz, d =3 m, θ =45º)

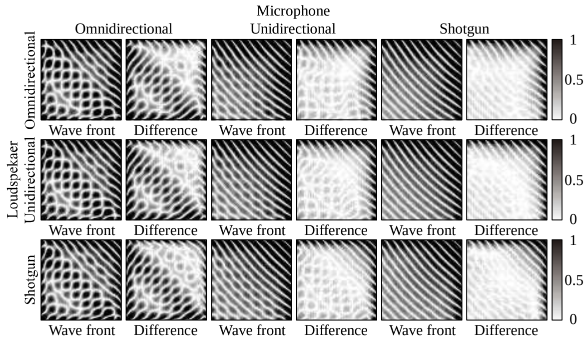

Figure 16 Synthesized wave fronts and differences when using square areas (f =500 Hz, d =10 m, θ =45º)

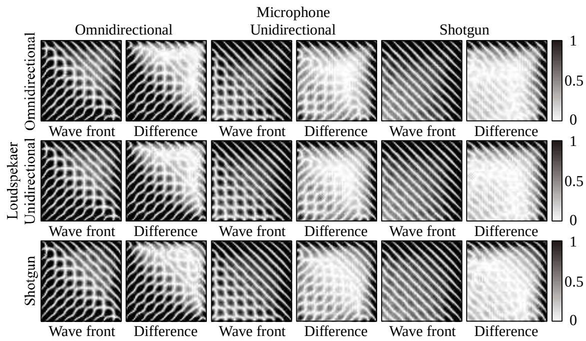

Figure 17 Synthesized wave fronts and differences when using square areas (f =500 Hz, d=100 m, θ=45º)

In order to evaluate quantitatively the reproduced sound field, SNRs were calculated according to Eq. (7). The range of $\mathbf{r}_a$ is limited to the inside of the 2-m square with ($|r_x|$, $|r_y|$ < 1). Note that $p_o(\mathbf{r}_a, f, 0)$ and $p(\mathbf{r}_a, f, 0)$ were normalized in $|r_x|$, $|r_y|$ < 1 before calculating SNRs. SNR results for each directivity of microphones and loudspeakers are shown in Figures 18 and 19. Note that error bars denote a 95% confidence interval for SNRs. SNRs of unidirectional and shotgun microphones are higher than those of omnidirectional microphones except in the case of d = 3 m, θ = 45 degrees. It is thus considered that wave fronts can be accurately reproduced in surround systems based on wave field synthesis if unidirectional and shotgun microphones are used.

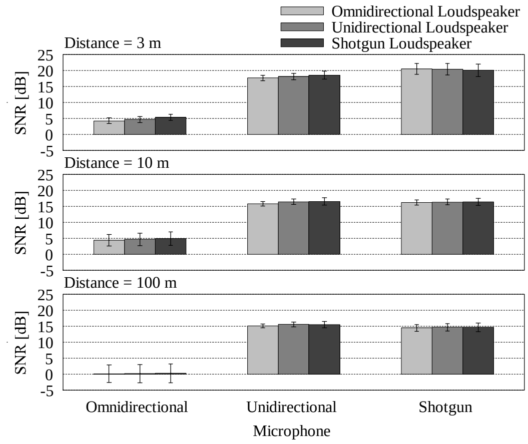

Figure 18 SNRs in the case of square areas (θ = 0º)

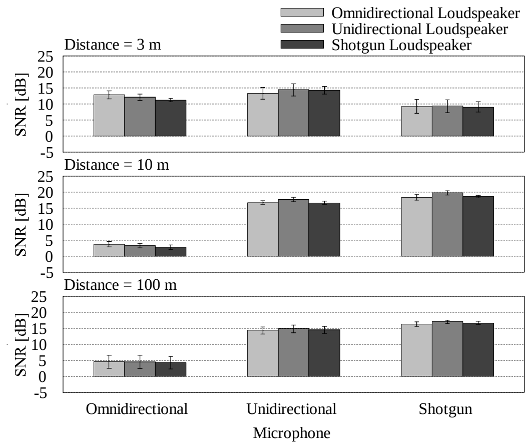

Figure 19 SNRs in the case of square areas (θ = 45º)

SNRs of unidirectional and shotgun microphones are always more than 15 dB. Although these values are lower than the SNR value (about 30 dB) of other studies [8-9], our method of study is effective from a practical point of view, as described in Section 2.

The 95% confidence intervals of Figures 18 and 19 are the variance of SNRs in frequencies. In this simulation, the width of confidential intervals is less than 6 dB, while the average of SNRs is 15 dB when using directional microphones. Thus, frequencies have no effect on the accuracy of wave fronts.

4. CONCLUSIONS

We performed a computer simulation in order to evaluate the effect of the directivity of microphones and loudspeakers on the accuracy of synthesized wave fronts produced by surround systems that use wave field synthesis. The results of our simulations of two different cases showed that the directivity of loudspeakers has almost no effect on the accuracy of the wave fronts. The results also showed that accurate wave fronts can be reproduced when either a unidirectional or a shotgun microphone are used.

In the future, the control and listening areas should be evaluated in three-dimensional space. To do this, it will be necessary to evaluate the effect of the directivity of microphones and loudspeakers on the accuracy of synthesized wave fronts in three-dimensional space.

5. ACKNOWLEDGEMENTS

The authors would like to thank Prof. K. Takeda, Prof. F. Itakura and Prof. M. Nakagawa for their continued support and encouragement during our research.

6. REFERENCES

[1] H. Fletcher, "Symposium on wire transmission of symphonic music and its reproduction on auditory perspective: Basic requirement," Bell System Technical Journal, vol. 13, no. 2, pp. 239-244 (1934)

[2] M. Camras, "Approach to recreating a sound field," Journal of the Acoustical Society of America, vol. 43, no. 6, pp. 1425-1431 (1968)

[3] A. J. Berkhout, D. de Vries, P. Vogel, "Acoustic control by wave field synthesis," Journal of the Acoustical Society of America, vol. 93, no. 5, pp. 2764-2778 (1993)

[4] J. Blauert, "Spatial Hearing," p. 78, revised edition, MIT Press, Cambridge, Mass. (1997)

[5] M. R. Schroeder, D. Gottlob, K. F. Siebrasse, "Comparative study of european concert halls: Correlation of subjective preference with geometric and acoustic parameters," Journal of Acoustical Society of America, vol. 56, no. 4, pp. 1195-1201 (1974)

[6] ITU-R Recommendation BS.775-1, "Multichannel Stereophoneic Sound System with and without Accompanying Picture" (1992-1994).

[7] http://www.iosono-sound.com/

[8] S. Ise, "A principle of sound field control based on the Kirchhoff-Helmholtz integral equation and the theory of inverse systems," ACUSTICA-Acta Acustica, vol. 85, no.1, pp. 78-87 (1999)

[9] S. Takane, Y. Suzuki, T. Miyajima, T. Sone, "A new theory for high definition virtual acoustic display named ADVISE," Acoustical Science & Technology, vol. 24, no. 5, pp. 276-283 (2003)