Subjective Listening Experiments on a Front and Rear Array-Based WFS System

Jae-hyoun Yoo, Jeongil Seo, Hwan Shim, Hyunjoo Chung, Koeng-Mo Sung, and Kyeongok Kang

Wave field synthesis (WFS) has been gathering more and more attention recently due to its ability to perfectly reproduce an original sound field. However, to realize theoretically perfect WFS, a four-sided loudspeaker array that encloses the listener is required. However, it is difficult to build such a system except in large listening spaces, such as a theater or concert hall. In other words, if the listening space is a home, installing a side loudspeaker array is impractical. If the two side walls located to the left and right of the listener can be omitted, a setup using only front and rear loudspeaker arrays may be a solution. In this letter, we present a subjective listening experiment of sound localization/distance based on a WFS using a front and rear loudspeaker array system which is conducted on two listening points and shows average localization errors of 6.1° and 9.18°, while the average distance errors are −27% (0.5 m) and −29% (0.6 m), respectively.

Keywords: Wave field synthesis, front and rear loudspeaker array.

I. Introduction

Traditional loudspeaker stereophony is a widely-used sound field reproduction technique. While a multichannel surround reproduction technique based on stereo reproduction makes it possible to reproduce a more realistic sound field, it is not possible using only two loudspeakers. However, while sound field reproduction techniques based on multichannel surround have been created, they have drawbacks in that the localization region of the sound images is limited and its sweet spot is narrow. On the contrary, wave field synthesis (WFS) presents a consistent 3D sound image without restriction of the listener's position [1]-[4]. It can also render moving sound sources and makes it possible to localize sound images in the listening area. Because a theoretically perfect system requires loudspeaker arrays to be installed on all four sides of a listening area, when configuring a WFS array in a home, left and right side loudspeakers are very difficult to install in a practical manner. So, in this letter, we address the practical considerations of a home listening environment by introducing subjective listening experiments on a WFS reproduction system based on only front and rear arrays.

Manuscript received Sept. 7, 2010; revised Apr. 11, 2011; accepted May 19, 2011.

This work was supported by the IT R&D program of KCC [Development of Multiview 3D Compatible UHDTV Broadcasting Technology].

Jae-hyoun Yoo (phone: +82 42 860 1556, email: jh0079@etri.re.kr), Jeongil Seo (email: seoji@etri.re.kr), and Kyeongok Kang (email: kokang@etri.re.kr) are with the Broadcasting & Telecommunications Convergence Research Institute, ETRI, Daejeon, Rep. of Korea.

Hwan Shim (email: yum@acoustics.snu.ac.kr), Hyunjoo Chung (email: hjman@acoustics.snu.ac.kr), and Koeng-Mo Sung (email: msung@acoustics.snu.ac.kr) are with the Department of Electrical Engineering, Seoul National University, Seoul, Rep. of Korea.

http://dx.doi.org/10.4218/etrij.11.0210.0335

II. Wave Field Synthesis

1. Concept of WFS

Generally it is possible to consider a wavefront made by one point source as a continuous and infinite series of spherical wavefronts. Because one wavefront is made by the synthesis of

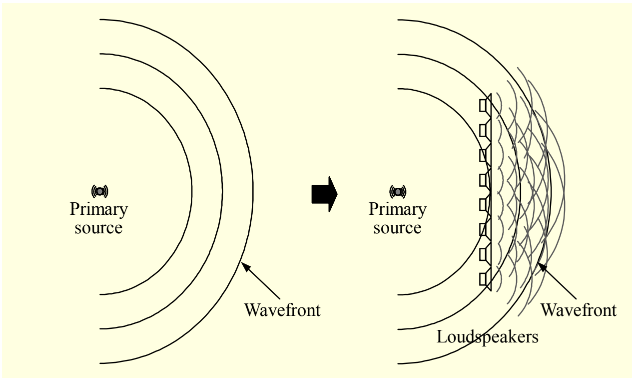

Fig. 1. WFS based on Huygens's principle.

another wavefront from every point of the foregoing wavefront, if we make loudspeakers as an array and each loudspeaker produces a certain wavefront to make one wavefront, we can make a wavefront of primary source. WFS theory is based on Huygens's principle, shown in Fig. 1, Rayleigh's representation theorem, and Kirchhoff-Helmholtz integral [1]-[4].

2. Problems of WFS Realization

There are some problems with a practical realization of theoretical WFS, for example, a spatial alias and truncation effect [1]-[4]. Because theoretically continuous secondary sources are replaced by discrete and distant loudspeakers, if the wavelength of sound is smaller than the distance of loudspeakers, the wavefront is not synthesized properly. This is a spatial alias. Because a theoretically infinite secondary wavefront is replaced by finite loudspeakers, the wavefront from the edge of the array is not synthesized properly. This is a truncation effect.

Another potential problem can arise from the installation of the loudspeaker array. To realize a theoretically perfect WFS system, a four-sided loudspeaker array enclosing the listener is required. Because of this, a number of loudspeakers need to be installed, so WFS is possible only in a large listening space, such as a theater or concert hall. In other words, it is not easy to perfectly install a WFS system in a narrow listening space found in a typical home living room.

III. Front and Rear Loudspeaker Array WFS

When configuring an WFS array in a small and narrow listening space such as a home, left and right side loudspeakers are very difficult to install in a practical manner. As a matter of fact, in a typical living room, a window is located on one side, while the other side opens up to the kitchen or another room. Therefore, it is almost impossible to install a loudspeaker array on those two sides. Also, a TV is generally installed in the front area of the living room, and in the rear, there may be a sofa against the wall.

Therefore, it can be very simple to install a system in front of the listener, in which most of the sound images come from the display device. In this case, however, there are problems in that the rear sound field is difficult to present; therefore, the sound image localization area is reduced. Hence, an additional loudspeaker array has to be installed in the rear region. In a traditional rear sound image localization scheme using a power panning method, there are problems of coloration and elevation. By installing a loudspeaker in the rear, these problems can be solved using WFS. Also, it is partially possible to localize sound images in the listening area using a front and rear loudspeaker array structure, and sound images on the sides can be presented using surround panning with loudspeakers on the edges of two arrays.

A useful way to use the system is when the sound image is on one side or the other. If there are sound images in the back of the front array, only the front loudspeaker array is driven to reproduce the sound field. If there are sound images in the back of the rear array, only the rear loudspeaker array is driven to reproduce the sound field. Also, if there are sound images between the front and back array, the front loudspeaker array and rear loudspeaker array are driven together to reproduce the sound field.

IV. Experiments & Results

1. Experiments

To build a front and rear WFS system, 16 loudspeakers are installed at each end, 32 loudspeakers in all. The distance between them is 20 cm, placed in parallel with the opposite side as depicted in Fig. 2.

The loudspeaker arrays are driven by 6-channel power amplifiers, ROTEL RMB-1066. Audio signals are generated using a MATLAB program and played from Max/MSP patches through four 8-channel audio interfaces, MOTU 896HD.

Subjective tests of sound image localization and distance localization are executed to inspect the performance of the system. Experiments were executed in an anechoic room at Seoul National University, and 12 subjects participated. The test sound signal was a burst of train noise containing 300 ms of white noise and 250 ms of silence repeated 8 times. The distance between the loudspeaker and listener was 1.5 m, measured from the midpoint between the front and rear loudspeaker arrays. There were two listening points: one at the

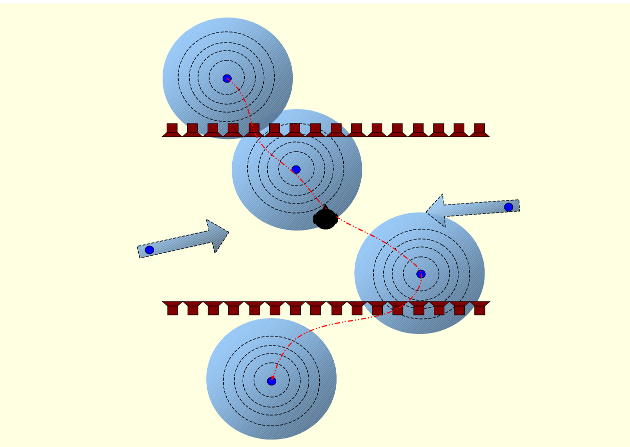

Fig. 2. Sound field reproduction using front and rear arrays.

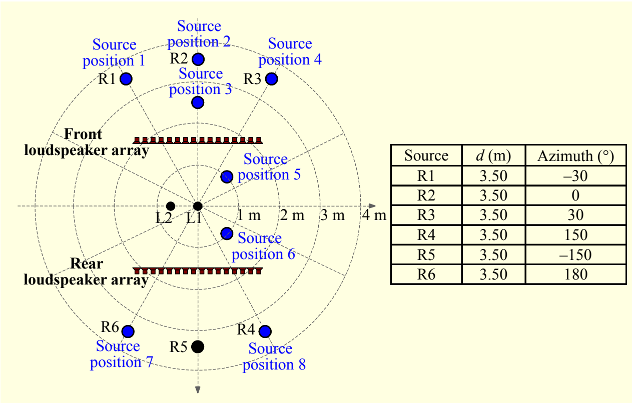

Fig. 3. Positions of test signals and references.

Table 1. Position and distance information of test signals.

Listening position	L1		L2	
Source	*d* (m)	Azimuth (°)	*d* (m)	Azimuth (°)
Position 1	3.50	−30	3.21	−18.15
Position 2	3.50	0	3.58	12.09
Position 3	2.50	0	2.61	16.70
Position 4	3.50	30	3.95	39.45
Position 5	1.00	50	1.60	61.19
Position 6	1.00	140	1.58	117.86
Position 7	3.50	−150	3.20	−161.79
Position 8	3.50	150	3.94	140.45

center and the other 0.75 m left of the center. Each loudspeaker array is visible to the listener.

The layout of the loudspeakers is represented in Fig. 3. L1 and L2 represent the listener positions, while R1 through R6 represent the reference loudspeaker positions where physical sound images were presented to the listener for referencing the localization and distance of the image sources (see Table 1). At the start of each test, these reference sound images were reproduced for comparing the localization and distance. A subject test answer sheet (see Fig. 3) was also provided to the listeners. The test sound signals were reproduced randomly from the positions represented in Fig. 3.

2. Results

Tables 2 and 3 show the results of sound image localization and distance localization at the L1 and L2 listening positions, respectively. In the distance column, we present a maximum error rate in a 95% confidence interval (CI) as referencing information of how far the sound image is from the source. It is represented by the rate of difference between the presented distance and replied distance of an averaged 95% CI. For example, −30% indicates that the test image source is perceived in the region at a distance 30% away from the intended position of the listener. Also, in the azimuth column, we present a maximum error azimuth in a 95% CI as referencing information.

Table 2. Results from L1 position.

L1	*d* (m)		Azimuth (°)	
Source	Average	95% CI*	Average	95% CI**
Position 1	2.73	−30%	−35.00	8.27
Position 2	2.77	−29%	2.90	6.65
Position 3	2.07	−29%	1.30	3.29
Position 4	2.81	−28%	32.30	5.77
Position 5	1.08	−35%	54.30	7.56
Position 6	1.27	−9%	123.20	14.64
Position 7	2.77	−29%	−147.30	0.44
Position 8	2.96	−28%	143.00	2.22

*Max. error rate in 95% confidence interval.
**Max. error azimuth in 95% confidence interval.

Table 3. Results from L2 position.

L1	*d* (m)		Azimuth (°)	
Source	Average	95% CI*	Average	95% CI**
position 1	2.65	−26%	−31.70	15.10
position 2	2.43	−41%	5.70	0.43
position 3	1.83	−44%	12.50	6.49
position 4	2.86	−36%	33.30	3.35
position 5	1.65	−14%	61.70	3.71
position 6	1.77	−17%	132.50	22.37
position 7	2.67	−26%	−145.00	13.30
position 8	3.04	−32%	142.10	8.65

*Max. error rate in 95% confidence interval.
**Max. error azimuth in 95% confidence interval.

For localization recognition at the L1 listening position, the average error distance and average error azimuth are −27% (0.5 m) and 6.1° in a 95% CI, respectively. For azimuth recognition, from the results of position 6, the sound image inside the listening area, which is the region between two loudspeaker arrays, is more difficult to hear than outside the listening area. Also, for distance recognition, the sound image of position 6 is likely to be perceived farther away from that position. This system does not use left-side and right-side loudspeaker arrays, and sound image positioning inside the listening area is somewhat distorted. We guess that because the subjects are conscious of the existence of the loudspeaker array, their responses at position 6 are that the sound image is close to

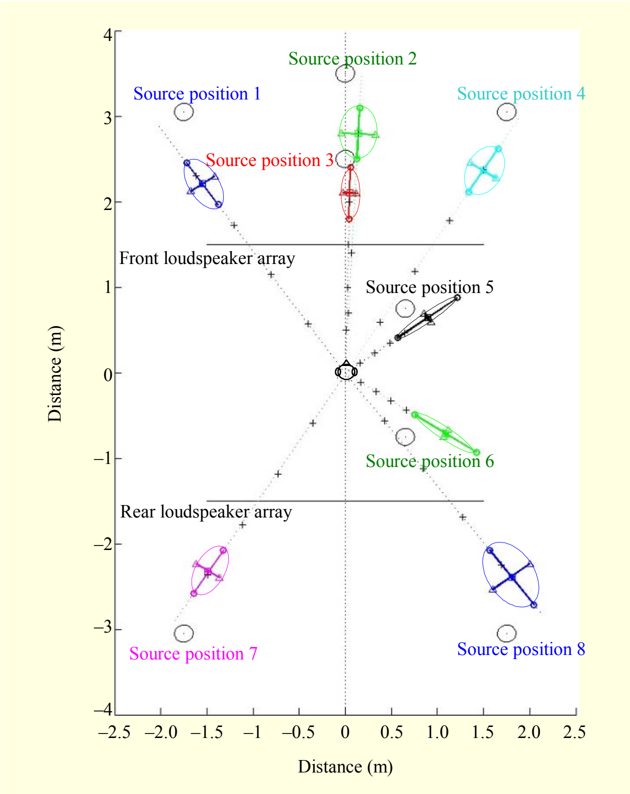

Fig. 4. Results from L1 position.

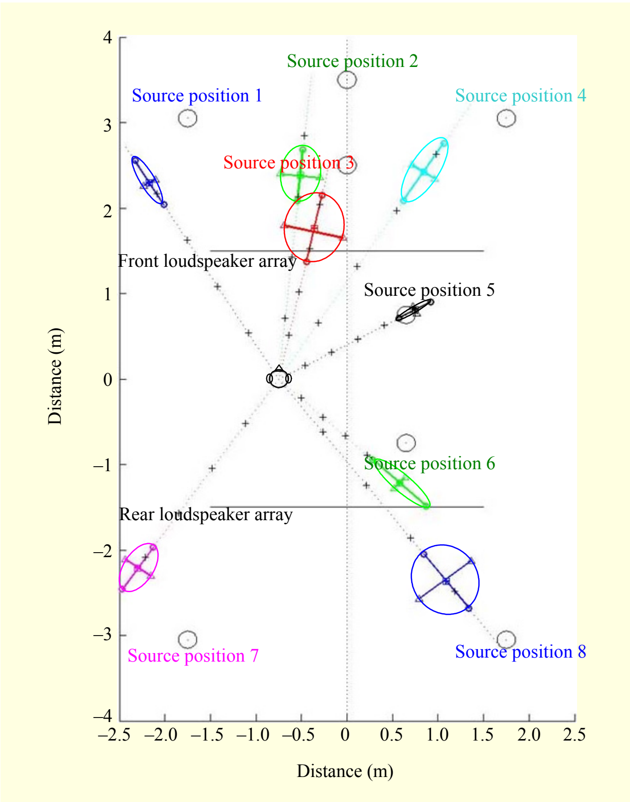

Fig. 5. Results from L2 position.

the loudspeaker array. This is same as the position 5 sound image. Sound images apart from the array, such as positions 1, 2, 4, 7, and 8, are perceived close to the array. We assume that because the subjects are conscious of the existence of the loudspeaker array, their responses at those positions are that the sound image is close to the loudspeaker array. However, most of the subjects distinguished the difference between positions 2 and 3.

For localization recognition at the L2 listening point, the average error distance in a 95% CI is −29% (0.6 m) and the average error azimuth in a 95% CI is 9.18º. The L2 listening point is selected in order to check the wide sweet spot of the WFS. Although we need to conduct further experiments at several locations, the results from L2 are somewhat low but similar to the results from L1. The properties of the results from L2 are also almost the same as those in L1. Thus, we could say that this system provides a wide sweet spot for several listeners.

Figures 4 and 5 show the results of the subjective listening experiments from L1 and L2, respectively. Each colored-ellipse represents the region of an averaged 95% CI error distance and azimuth for each virtual image.

V. Conclusion

The front and rear array-based WFS system may be a practical and useful setup in a home environment. Based on the results of this experiment, at the L1 and L2 listening positions, the average localization errors are 6.1º and 9.18º, while the average distance errors are −27% (0.5 m) and −29% (0.6 m), respectively. This can be considered a useful WFS system when the installation of a side loudspeaker setup is not possible. For a more accurate evaluation of this system, i) a comparison with four-sided loudspeaker installed systems is needed, which we are planning to conduct, and ii) experiments in a listening room having similar reverberation to an actual living room need to be conducted.

References

[1] A.J. Berkhout, D. de Vries, and P. Vogel, "Acoustic Control by Wave Field Synthesis," *J. Acoust. Soc. Am.*, vol. 93, no. 5, 1993, pp. 2764-2778.

[2] E. Verheijen, "Sound Reproduction by Wave Field Synthesis," PhD Thesis, TU Delft, 1998.

[3] G. Theile, "Wave Field Synthesis – A Promising Spatial Audio Rendering Concept," *Proc. 7th Int. Conf. Digital Audio Effects*, Oct. 2004, pp. 125-132.

[4] H. Wittek, "Perception of Spatially Synthesized Sound Fields," 2003. http://www.hauptmikrofon.de/wittek.htm