4 Spatial Audio Technology

4-1 Ultra-Realistic Audio by 3D Audio Display

KATSUMOTO Michiaki, YAMAKATA Yoko, and KIMURA Toshiyuki

We have been prototyping loudspeakers with radial directionality that differs in all directions (including longitudinally, vertically and laterally), and building systems that reproduce sound expressions of musical instruments in a 3D manner, as technology for faithfully reproducing a sound source according to physical laws as much as possible. In that process, we assumed that the presence of a sound source might be an important factor affecting how people perceive sound, in addition to the faithful reproducibility of a sound source by loudspeakers. We then conducted trial listening experiments to verify how much presence a sound source has, and evaluated performance of the prototyped loudspeakers. This paper introduces our prototyped loudspeakers and presents the results of those experiments.

Keywords
3D audio, Multi-radiated acoustic, Multi-radial loudspeaker, Presence

1 Introduction

Studies have been conducted in recent years to make ultra-realistic communication a reality by using 3D imagery and 3D audio[1]. If such an approach makes it possible to produce and appreciate more realistic 3D expressions of imagery and audio from all directions in a 3D space, then it should become possible to provide a presence that has not been possible with conventional imagery and audio media, or realize communication or service with presence. To that end, people will need a technology for faithfully reproducing imagery and physically reproducing audio more than ever before.

In the imagery field, studies on 3D imagery using electronic holography are being conducted to combine images in space[2][3]. A cube-type 3D display has also been developed, allowing several people to observe 3D images from all directions[4].

Then there is the need for acoustic technology that is suitable for these image systems and enables reproduction as physically faithful as possible. This paper focuses on wave field synthesis[5]-[8], which is one of the 3D sound field reproduction technologies, as a method of faithful acoustic expression, and describes the implementation of a 3D audio display offering different directionality in all directions, along with an evaluation of its performance.

2 Implementation of ultra-realistic 3D audio

Presenting 3D information in space means that the front and back, top and bottom, and left and right sides of a body must be provided. In the real world, a sounding body emits sound in all directions, and such matters as pertaining to musical instruments have already been documented[9]. Moreover, we consider that said body is recognized as a 3D body by our human visual ability, and that 3D information can be effectively used at an extremely short distance (within arm's reach).

It is therefore necessary to conduct research and development regarding an acoustic system with different radial directionality that enables expressions as physically faithful as possible in an environment near humans, and sound sources that emit sounds in all directions, unlike conventional stereo and surround systems.

Here, wave field synthesis, which is expected to express acoustics faithfully, is a 3D sound field reproduction technology that records sounds with a microphone array arranged on the boundary surface of the control area set in the original sound field, and faithfully reproduces sounds collected with a speaker array arranged on the boundary surface of the listening area set within the reproduction sound field, thereby faithfully synthesizing the wave surfaces of the control area in the listening area based on the Kirchhoff-Helmholtz integration equation.

We previously proposed, as an ultra-realistic 3D acoustic technology, a near-field 3D sound field reproduction technology that employs loudspeakers having different radial directionality in all directions (including longitudinally, vertically and laterally) and wave field synthesis, thereby allowing the listener to hear sounds around the sound source. We then considered the conditions under which wave surfaces can be synthesized faithfully by conducting computer simulations[10]. This paper introduces the simplest method of implementation, for which we actually built a microphone array and a speaker array in prototyping a system, and discusses whether the proposed technology can be implemented.

3 Overview of the proposed system

Figure 1 shows the composition of the previously proposed near-field 3D sound field reproduction system[11]. First, the original sound field is provided with a surrounding microphone array consisting of many omnidirectional microphones on the boundary surface around the sound source to record sound $x_i(t)$. Next, a 3D audio display consisting of the same number of directional loudspeakers as microphones is arranged in the reproduced sound field, and recorded sound $x_i(t)$ is reproduced as is. At that time, each loudspeaker unit should be arranged at the same location on the same axis as the microphones, and oriented toward the outer side of the boundary surface. Then, the wave surface is faithfully synthesized outside the boundary surface, so that listeners outside the boundary can hear the sound source as if sounding from within the boundary, regardless of where the listeners may be positioned outside.

In the example given in Fig. 1, listeners close to the piano can hear the sounds as being nearby, as do the listeners close to the violin.

Then, in order to verify the effectiveness of the proposed acoustic system with different directionality, we prototyped a multi-radial 3D audio display that achieves different radial directionality.

4 Prototyping of multi-radial loudspeakers

4.1 Structure of multi-radial loudspeakers

The multi-radial loudspeakers are structured so that, as shown in Fig. 2, a multi-radial enclosure contains 26 loudspeaker units in order to express a single sound source (such as a violin), and each unit can produce different acoustic signals by using 26 amplifiers (custom-made products). The enclosure 17 m in

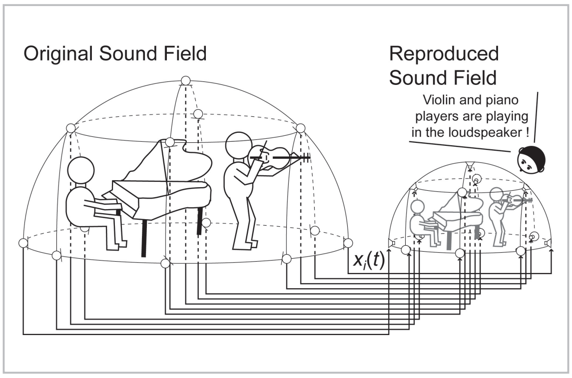


Fig.1 Conceptual diagram of 3D acoustic system based on loudspeakers with different directionality

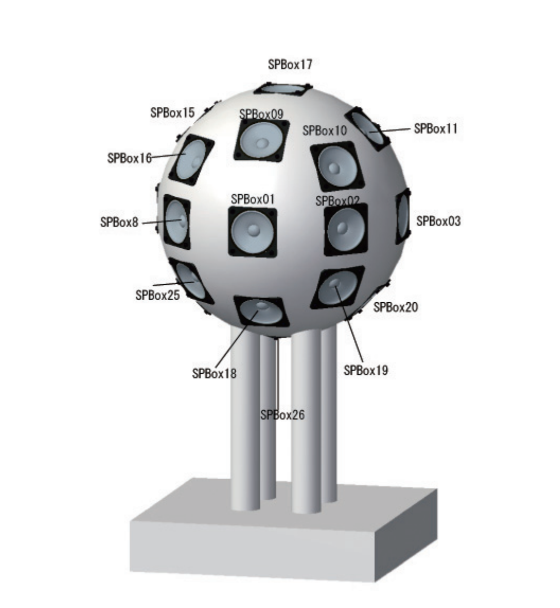


Fig.2 *Sketch of the multi-radial enclosure and loudspeaker units*

Table 1 *Arrangement information about loudspeakers in the multi-radial enclosure*

Horizontal plane				Northern hemisphereplane				Southern hemisphereplane			
SP BOX	r [cm]	θ [°]	φ [°]	SP BOX	r [cm]	θ [°]	φ [°]	SP BOX	r [cm]	θ [°]	φ [°]
1	8.5	0	0	9	8.5	0	45	18	8.5	0	-45
2	8.5	45	0	10	8.5	45	30	19	8.5	45	-30
3	8.5	90	0	11	8.5	90	45	20	8.5	90	-45
4	8.5	135	0	12	8.5	135	30	21	8.5	135	-30
5	8.5	180	0	13	8.5	180	45	22	8.5	180	-45
6	8.5	-135	0	14	8.5	-135	30	23	8.5	-135	-30
7	8.5	-90	0	15	8.5	-90	45	24	8.5	-90	-45
8	8.5	-45	0	16	8.5	-45	30	25	8.5	-45	-30

Top				Bottom			
SP BOX	r [cm]	θ [°]	φ [°]	SP BOX	r [cm]	θ [°]	φ [°]
17	8.5	---	90	26	8.5	---	-90

diameter was made of ABS resin integrally formed by using a 3D printer. Moreover, the 1-inch AURASOUND: NSW1-205-8A units or equivalent were used as the loudspeaker units. Each loudspeaker unit was arranged as shown in Table 1 to maximize its radial directionality.

4.2 Content recording and evaluation

To record a single musical instrument as content, it was decided that microphones would be installed around the musician in an anechoic chamber for recording the performance (Fig. 3). At this time, the installation location of each of the 26 microphones was determined to be within a radius of 1 m from the center of the musical instrument and thus similar to the arrangement of loudspeaker units. We employed 26 DPA 4060 microphones, Millennia HV-3D head amplifiers, and the Digidesign: ProTools HD as recording equipment. These units were used to record violin sounds (of the musical piece *El Choclo*) and acoustic guitar sounds (of an original musical piece) with a sampling frequency of 96 kHz and a quantization bit rate of 24 bits.

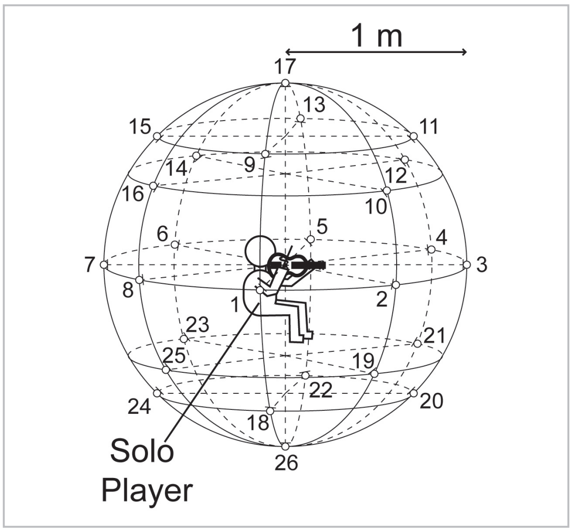


Fig.3 *Sketch of the multi-radial arrangement of microphones*

We used this system to reproduce these sound sources as they are, and conducted trial listening experiments involving musicians and others engaged and not engaged in acoustics, and our evaluation revealed that many did not know the shapes of sound sources or other details, but recognized that the sound sources had high presence[12].

5 Experimentation system and environment

5.1 Experimentation system

We conducted a subjective experiment on actual performances and multi-radial loudspeakers, in order to verify how much presence the system presented. At that time, we recorded real performances for trial listening and built a system to immediately reproduce those recordings, in order to enable the continuous

Fig.4 Photo of multi-radial arrangement of the microphones

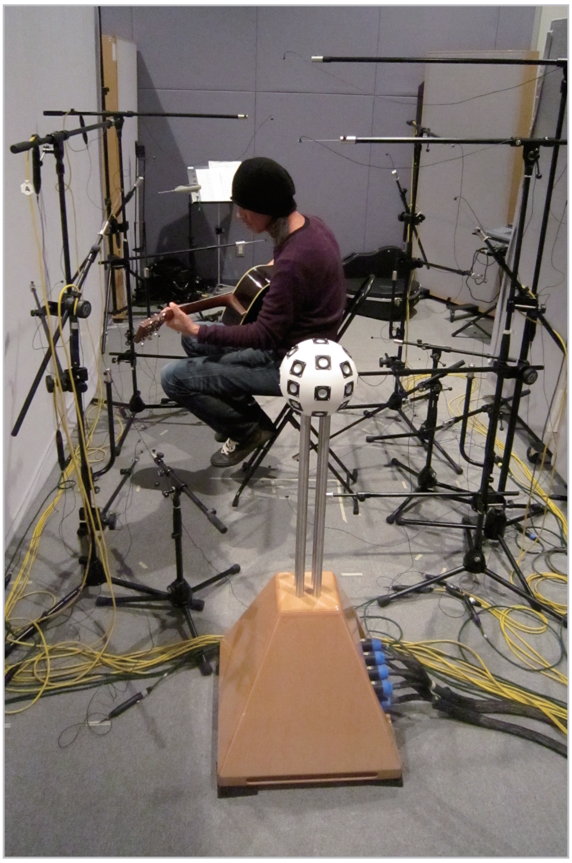


Fig.5 Arrangement of the musician and multi-radial loudspeakers

Fig.6 Photo of the actual experimentation environment

recording of real performances and continuous trial listening of multi-radial loudspeakers.

Music played by real musicians was recorded with the DPA miniature microphone 4060 arranged at a location similar to that of the multi-radial loudspeaker units at of 0.8 m from the acoustic center (Fig. 4). This state was recorded synchronously with 26 channels by using four Millennia HV-8D head amplifiers and with Digidesign: ProTools HD3 as output. The recording format was a PCM WAVE with a sampling frequency of 96 kHz and a quantization bit rate of 24 bits.

In reproduction, recorded data was adjusted to set its sound pressure to the same level as in the actual performance, and then synchronously reproduced with 26 channels by using multi-radial loudspeakers and custom-made digital amplifiers.

5.2 Experimentation environment

The main difference between multi-radial loudspeakers and conventional ones is that multi-radial loudspeakers radiate acoustic signals in all directions. We estimate that such radiation toward the rear side is particularly different from that in conventional loudspeakers, and largely responsible for the difference between both types of loudspeakers. For that reason, we arranged a reflection board and an absorbing board behind the musician and multi-radial loudspeakers, respectively, and built an environment enabling two patterns of experimentation, in order to check whether the present experiment could demonstrate the effects of that difference. At that time, the reverberation time was 228 ms for the reflection board and 224 ms for the absorbing board. Moreover, each board was placed at a distance of 1.2 m from both the musician and multi-radial loudspeakers.

Care should be taken to prevent a clear distinction between the musician and multi-radial loudspeakers, and any change in acoustic characteristics. To avoid noise stemming from microphone stand installation and movement by the musician, it was decided that both would be arranged as closely together as possible (Fig. 5). At that time, the distance between the musician and multi-radial loudspeakers was 1.2 m.

In order to avoid giving visual impressions to the subject, an acoustic penetration screen was also erected between the sound source subject to trial listening and the subject. This established an environment that would prevent people from visually distinguishing the musician and multi-radial loudspeakers (Fig. 6). The distance from each sound source to the screen was 1.2 m, with the subject to engage in trial listening at the position 2 m from the screen while seated in a chair midway between the actual performance and multi-radial speakers. Consequently, the distance between the subject and sound source was 2.2 m. The reverberation time on the subject's side at that time was 209 ms for the reflection board and 193 ms for the absorbing board.

6 Trial listening experiment

The present experiment involved the environment described in Section **5**, where the subjects would listen to an actual performance and the loudspeakers, and then describe personal impressions of the sounds, followed by a hearing survey based on interviews. At that time, the subjects were instructed to hear and compare sound sources A and B, sources D and C, and then describe their impressions. That is, the experiment was conducted by not informing the subjects that they were listening to a real musician or a set of loudspeakers. Moreover, since the experiment was conducted during actual recording, the subjects would hear the actual performance first, and then the multi-radial loudspeakers.

6.1 Preparations by the musicians

In this experiment, the musicians were requested to simply play music as usual and refrain from making noise by minimizing their presence when not playing a musical instrument. Some musicians were given a sufficient description of the experiment and those who agreed were asked to participate. Prior to the experiment, the musicians were asked to choose a familiar musical piece about three minutes long that they were able to play in a similar manner, and practice sufficiently for the experiment.

In the actual experiment, all the pieces were identical. Given the live nature of the performances, slight differences arose in each experiment, but the subjects had virtually the same impressions.

6.2 Subjects

The subjects were six people (men and women) in their 20s and 30s who were able to make appropriate expressions of sounds—the familiar and actual sounds of musical instruments—and who were engaged in music or were musicians. Moreover, these subjects knew that multi-radial loudspeakers existed, though not through actual experience. This meant that the subjects participated in the experiment, thinking that it was a trial listening experiment involving multi-radial loudspeakers.

6.3 Experiment

The experiment was conducted as follows: The subjects listened to sound source A (real performance about three minutes long, using a reflection board), followed three seconds later by sound source B (performance by multi-radial speakers for the same period of time, using a reflection board) as the first instance of trial listening. After that, 10-minute interview 1 was conducted. Then, the subjects listened to sound source C (real performance about three seconds long, using an absorbing board), followed three seconds later by sound source D (performance by multi-radial loudspeakers for the same period of time, using an absorbing board) as the second instance of trial listening. After that, 10-minute interview 2 was conducted. Last but not least, after the subjects were informed of the actual method of experimentation, interview 3 was conducted for an arbitrary amount of time. The procedure for the experiment is described below.

1. The subjects were given a description of the experiment in the waiting room (a normal living room).
2. First instance of trial listening: The subjects went to the laboratory and participated in trial listening.
3. The subjects returned to the waiting room, described their impressions, and then were interviewed.
4. Second instance of trial listening: The subjects went to the laboratory and participated in trial listening.
5. The subjects returned to the waiting room, described their impressions, and then were interviewed.
6. Final description: At this time, the subjects were informed that one performance was an actual performance. Based on that result, interviews were conducted.

With the above steps combined, each subject participated in a 30-minute experiment. A researcher accompanied each subject to the laboratory, and the researcher inspected the musician's state of mind and listened for unnecessary noises.

7 Experiment results and observations

7.1 Results obtained immediately after first instance of trial listening

Immediately after the first instance of trial listening to sound sources using a reflection board (i.e., sound sources A and B for trial listening), four of the six subjects said that the multi-radial loudspeakers gave a better sound impression. At that time, all six subjects said that the real performance was louder. In other words, the subjects apparently heard the performances from the tall line array of loudspeakers. The subjects were therefore expected to say that one of the sound sources was better due to the perceived loudness of the sound source.

7.2 Results obtained immediately after second instance of trial listening

Immediately after the second instance of trial listening to sound sources using an absorbing board, four of the six subjects who listened to sound sources D and C said that the multi-radial loudspeakers were better as expected. It can be estimated that the difference from the first instance of listening was attributed to a feeling of the sound image deviating toward the center, that is, the second instance of trial listening was conducted with the loudspeakers being repositioned.

7.3 Results obtained after the final interview

Upon having been told that one performance was a real performance, one subject who deemed the real performance to be better had also felt "a human presence" all along, and stated that the real performance was better both in terms of hearing and presence. Another subject had suspected that "it might have been a real performance" and, after being thus informed, recognized that it was a real performance and apparently had the feeling that it made sense.

The remaining four subjects did not find that it apparently made sense, even after being informed of the truth. They said that sound quality was definitely better in the real performance, but felt somewhat uncomfortable with the magnitude of the sound image.

7.4 Discussions

The point clarified as the result of the trial listening experiment was that the multi-radial loudspeakers have high reproduction capacity. There was actually a great challenge posed by reproducing the magnitude and location of the sound image. However, it remains unclear why the subjects failed to notice differences between the real performance and multi-radial loudspeakers or the reasoning behind their impressions.

Upon entering the laboratory, one subject felt a human presence, stating that, "I felt a little afraid." As a paradoxical concept, one could also surmise that the multi-radial loudspeakers had low reproduction capacity had the other subjects not noticed it on their own, but felt a human presence and unconsciously felt "afraid" of the real performance, thereby deciding that the multi-radial loudspeakers perceived as harmless were better.

It is also estimated regarding the effects of radiation onto the rear side in the present experiment that reflection from the rear side simply disappeared, and that the subjects felt sound pressure more intensely from the front side and a more deviated location.

8 Conclusion

In the present experiment, we considered it impossible to draw a conclusion that sufficiently shows the high performance of multi-radial loudspeakers by merely viewing the evaluation results. The reason was revealed in discussions held after interviewing the subject who had felt a human presence all along. Since that actual subject was the final one, we were unable to change the method of experimentation. Based on that result, an additional experiment is thus scheduled to be conducted.

We also plan to enhance and improve the multi-radial loudspeakers, and establish a method of experimentation for the reflection of a sound source based on the results regarding the magnitude of the sound image.

References

1 K. Enami, "Research and Developments of Ultra-Realistc Coomunication at NICT and URCF," Image Labo, 18, 9, pp. 41–14, 2007.

2 R. Ooi, T. Mishina, M. Okui, Y. Nojiri, and F. Okano, "A Fast Hologram Calculation Method for Real Objects," ITE, Vol. 61, No. 2, pp. 198–203, 2007.

3 K. Yamamoto, T. Mishina, R. Ooi, T. Senoho, and M. Okui, "A Real-Time Color Holography System for Live Scene," Proc. SPIE, Vol. 7233, No. 723310, pp. 1–9, 2009.

4 S. Yoshida, R. Lopez-Gulliver, and S. Yano, "gCubik: A Graspable Cubic 3D Display for Multiuser Environments," 3D Forum, 22, 3, pp. 25–28, 2008.

5 H. Fletcher, "Symposium on wire transmission of symphonic music and its reproduction on auditory perspective: Basic requirement," Bell System Technical Journal, 13, 2, pp. 239–244, 1934.

6 T. Itho, "Theory of Sound Engineering," pp. 587–589, CORONA PUBLISHING CO., LTD., 1973.

7 M. Camras, "Approach to recreating a sound field," J. Acoust. Soc. Am., 43, 6, pp. 1425–1431, 1968.

8 A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," J. Acoust. Soc. Am., 93, 5, pp. 2764–2778, 1993.

9 N. H. Fletcher and T. D. Rossing, "The Physics of Musical Instruments, 2nd edition," Springer, pp. 76–78, 2005.

10 T. Kimura, "Research on volume of information reduction technique in solid sound place reproduction by wave field synthesis method," Dr thesis Nagoya University, 2005.

11 T. Kimura, Y. Yamakata, and M. Katsumoto, "Theory of multi-dimensional speaker with frequency dependence of directivity," EA2008-4, IEICE, 2008.

12 M. Katsumoto, Y. Yamakata, and T. Kimura, "Realization of multi-dimensional speaker with frequency dependence of directivity," Acoust Sci. & Tech., 37, 10, pp. 837–842, 2007.

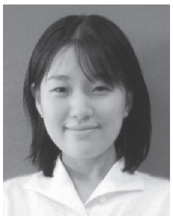
13 T. Okamoto, R. Nishimura, and Y. Iwaya, "Estimation of sound source positions using a surrounding microphone array," Acoust. Sci. & Tech., 28, 3, pp. 181–189, 2007.

(Accepted Sept. 9, 2010)

KATSUMOTO Michiaki, *Ph.D.*

Senior Researcher, 3D Spatial Image and Sound Group, Universal Media Research Center

3D Audio



YAMAKATA Yoko, *Ph.D.*

Kyoto University

3D Audio

KIMURA Toshiyuki, *Ph.D.*

Expert Researcher, 3D Spatial Image and Sound Group, Universal Media Research Center

3D Audio