Development of a 22.2 Multichannel Sound System

1. Introduction

The NHK Science and Technical Research Laboratories is conducting research into an ultra-high definition video system to serve as a next-generation broadcast medium, capable of providing not only realism, but also a vivid sense of immersion and fidelity. In pursuing this research, we developed a prototype 4,000-scan line image capture and display system (Super Hi-Vision), and a 22.2 multichannel sound system that realizes 3-D audio. The main features of the 22.2 multichannel sound system are that the loudspeakers are positioned not only at ear-height, but also above and below the viewer. In contrast to conventional theater sound systems, in which loudspeakers are placed only at ear level, with sound coming from the front and rear, and left and right, the sound in the 22.2 multichannel system approaches the listener from above and below as well.

In the paper, we outline the features of the 22.2 multichannel sound system, and describe how such a system offers superior sound field reproduction compared with conventional theater sound systems. Furthermore, we investigate the optimum loudspeaker positioning for this system.

2. Proposal for a Suitable Sound System for Super Hi-Vision

2.1 Conventional theater sound systems

A multichannel sound system is an audio reproduction method that makes use of multiple sound playback channels. Various such systems are in use in theaters, cinemas, and the like. The loudspeaker configurations for the main multichannel sound formats are shown in Fig. 1. Of these, the most widespread type that is used in cinemas and home theaters is 5.1 surround sound. In this setup, there are three channels in front of the viewer, two channels behind, and one LFE (low frequency effects) channel. Other sound formats include 6.1, which adds one loudspeaker directly behind the viewer, and 7.1, which positions five channels in front of the viewer. In all these various formats, the loudspeakers are positioned at approximately ear level. This limitation presents a number of issues that make these sound system formats unsuitable for Super Hi-Vision, as explained below.

- Difficulty in matching on-screen picture to sound for wide vertical viewing angle

Since current sound system formats place loudspeakers only at the height of the center of the screen, sound image movement in the up/down direction is impossible. At the same time, the horizontal and vertical viewing angles are far greater in Super Hi-Vision than in Hi-Vision. For this reason, when an object image on-screen moves up and down, or left and right, the sound image synchronized to this visual movement must similarly be moved not only horizontally, but also vertically.

- Difficulty of creating realism when viewing range is wide

When reproducing the kinds of spatial impression obtained in a venue such as a concert hall using conventional multichannel sound systems, reflected sound is recreated only by positioning loudspeakers at the side and rear walls. However, as the listener approaches the walls, the sound from these loudspeakers becomes louder, thereby destroying the playback sound field. Since the viewing angle of Super Hi-Vision is large even at the rear of the audience area, the sound system must be capable of maintaining a high level of realism over a wide range.

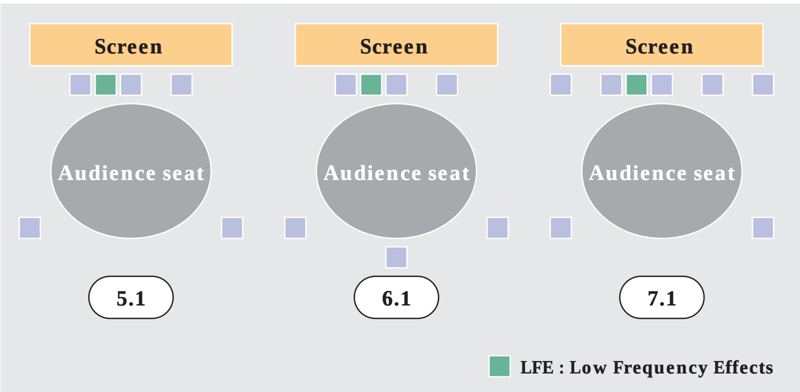

Figure 1: Loudspeaker positions in popular cinema theater sound systems

2.2 Outline of the 22.2 multichannel sound system

The above problems with current theater audio systems led us to develop a 22.2 multichannel sound system that would be suitable as a sound format for Super Hi-Vision systems. Figure 2 offers a general illustration of a 22.2 multichannel sound system. In this system, the loudspeakers are positioned in layers, at three heights. In addition to a middle layer at approximately ear level, there is an upper layer and lower layer. Four features of this system are described below.

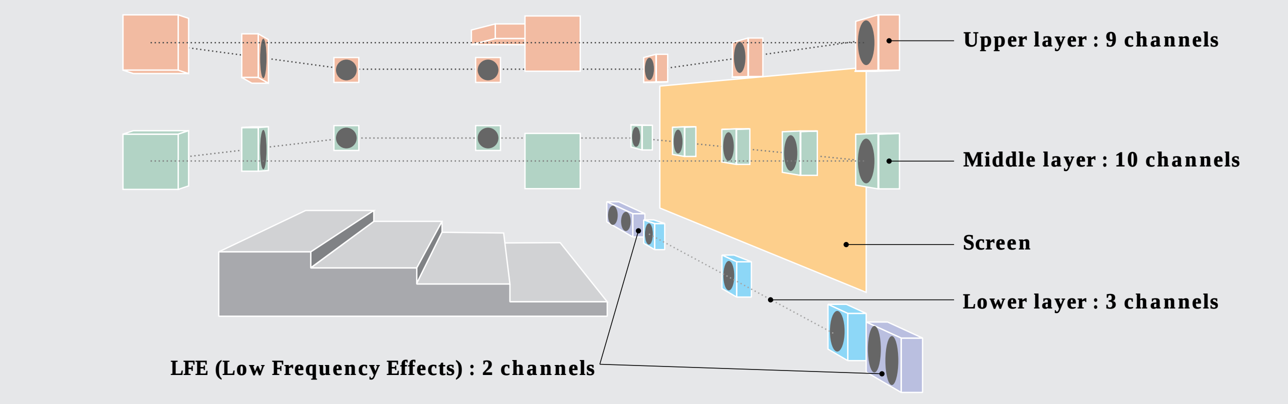

Figure 2: Outline of 22.2 multichannel sound system

(1) Matching of sound with on-screen picture

The placement of loudspeakers above and below the screen enables the matching of sound with images even for wide vertical viewing angles. In addition, five channels of audio in front of the viewer in the middle layer ensure good sound reproduction for wide horizontal viewing angles. (At present, Super Hi-Vision screens do not transmit sound, so loudspeakers cannot be placed behind the screen. To address this limitation, three of the five channels in the front, middle layer (excluding the far left and far right ones) are set above or below the screen. The combination of these loudspeakers creates virtual sound images on the screen.

(2) Maintaining realism over a wide viewing range

By reproducing reflected and reverberant sound from the upper layer loudspeakers, it is possible to distribute the approach direction of reflected sound, and thereby maintain realism over a wide viewing range.

(3) Achieving a sense of sound approaching from above or below

By positioning multiple loudspeakers above the viewer, it is possible to reproduce sound approaching from above. For example, it would be possible to re-create a scene of angels swooping down over the audience from above, and in combination with the middle and lower layer loudspeakers realizing a sense of vertical movement in the direction of the approaching sound.

(4) Ensuring compatibility with current theater sound formats

The middle layer of this system embraces all the sound formats shown in Fig. 1, thus offering compatibility with existing formats.

3. Evaluating the Realism of the 22.2 Multichannel Sound System

We conducted an experiment to compare the realism achievable with a 22.2 multichannel sound system with that of a conventional 5.1 surround sound theater audio system. To create the sound source for this experiment, we recorded an orchestra performance at NHK Hall. The composition was Tchaikovsky's Symphony No. 6. To pick up direct and indirect sound as independently as possible, we positioned multiple microphones having different directionalities on stage and above the listener, and

Table 1: Playback conditions of the experiment. Loudspeaker numbers are as in Fig. 3

Playback conditions	Used loudspeakers	Explanation
All used	1-19	All except lower layer used Conditions for achieving realism with a 22.2 multichannel sound system
5.1 compliant	2-4, 8,9	ITU-R BS.7775.1 compliant Conditions based on conventional theater audio methods
Front 3 ch	2-4, 6-19	3 channels for front, middle layer Conditions for verifying effectiveness of 5 front channels
Only middle layer	1-10	Middle layer only (no upper layer) Conditions for verifying effectiveness of upper layer

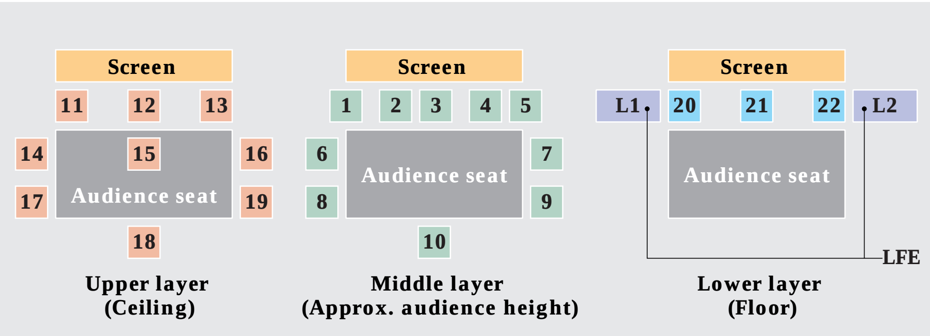

Figure 3: Loudspeaker arrangement of 22.2 multichannel sound system

recorded these separately on a multitrack recorder. This raw recording was then mixed by a recording engineer seated in the audience center of (listening position C2 in Fig. 4) of the 22.2 multichannel sound system, based on the four reproduction conditions described in Table 1. From this mix, the engineer created a listening test piece (continuous duration of 25 seconds). The mixing was done so as to ensure constant perceived loudness and reverberation levels. Since, the recording contained no suitable audio source for reproduction from the lower layer loudspeakers, the lower layer channels were not used in this experiment.

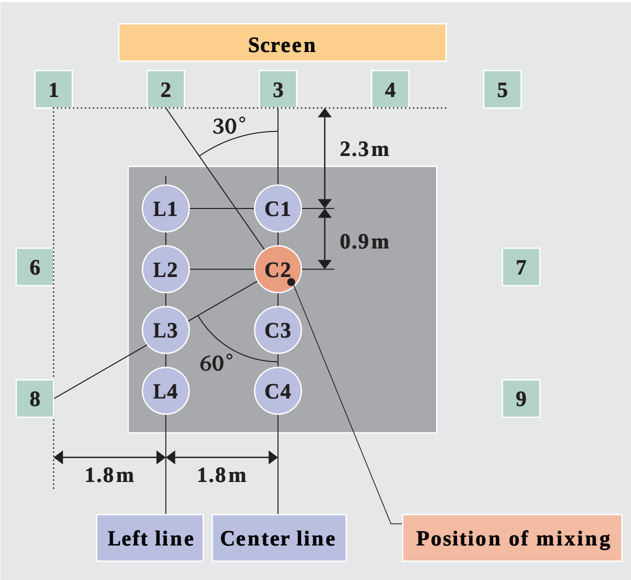

Figure 4: Audience positions in experiment

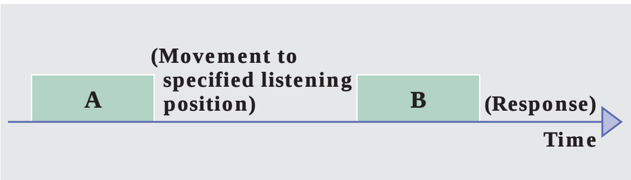

Figure 5: Method of presenting stimulus to subject

To determine if the level of realism can be preserved over a wide range of positions with a 22.2 multichannel sound system, the experiment was repeated at eight listening positions, four on the central line (C1-C4 in Fig. 4) and the four on the left line (L1-L4). Because of the left-right symmetry of the arrangement, evaluations to the right of center were unnecessary.

The magnitude estimation method enabled a quantitative comparison of the 22.2 channel and 5.1 channel systems. Figure 5 shows the method of presentation to the subjects. "A" indicates a reference sound that was presented under playback condition "5.1" to the subject, who listened to it at position C2 in Fig. 4. After this, the subject moved to one of the other listening positions, and listened to the evaluation sound "B." "B" was presented under one of the four conditions listed in Table 1. The subject was then asked to assign a score to the perceived realness of "B" based on a score of 100 for "A." A total of 21 subjects (17 audio specialists, and four music university students) participated in the experiment.

Figure 6 shows the geometric mean of the scores. The results for the center row (right side of Fig. 6) for the playback condition "All used" show a score of 129 at C2. In other words, the perceived realism was approximately 1.3 times greater than that of "5.1." At other listening positions too, excluding L1 (see left side Fig. 6), a score of 100, more or less, was maintained. These results make it clear that compared with a 5.1 system, a 22.2 multichannel sound system can maintain a higher level of realism over a wider range of listening positions.

The score for the playback condition "Forward 3ch," in which only three front channels of the middle layer were played, was greatly inferior to the "All used" condition, in which five front channels were used. This shows that the use of five front channels was very effective in conveying realistic sensations. Furthermore, the score for the "Middle layer only" condition, in which no upper layer loudspeakers were used, was much less than "All used," except at C4 (see right side of Fig. 6), proving that the added upper layer loudspeakers are very effective.

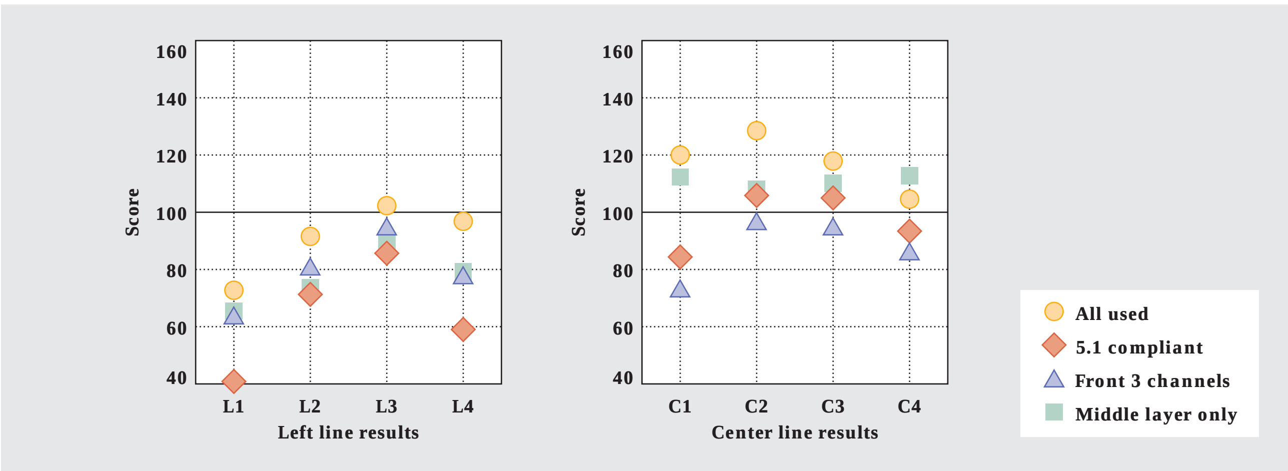

Figure 6: Scores (geometric mean)

4. Investigation of Optimum Height of Upper Layer Loudspeakers

One of the characteristics of the 22.2 multichannel sound system is that it can provide a sense of sound approaching from above or below. The height of the upper layer loudspeakers is an important element in achieving this effect. In a study on the perception of sound approaching from above or below, conducted by Kurosawa et al., white noise was used to measure stereotaxic discrimination limits. According to the report, the discrimination limit is approximately 4 to 5° in front of, and directly behind the viewer, and approximately 10 to 20° directly above the listener. The ability to discriminate between angles of approach becomes duller as the position gets higher. Based on this previous research, we would predict that the perception of sound moving higher increases as the vertical angle of the upper layer loudspeakers increases, but that after a certain angle a state of saturation is reached. To test this, we conducted an evaluation experiment to assess the degree of perception of sound images approaching from above and below (subjective ratings) in relation to the angle of playback loudspeakers.

The experiment was conducted in a sound proof room (reverberation time at 500 Hz = 0.18 sec.) Figure 7 shows the positions of the loudspeakers in the experiment. Seven loudspeakers were positioned along the circumference of a vertical circle of radius 2 m at intervals of 15° (with the listener at the circle center). Each loudspeaker faces the center of the circle. The group of loudspeakers is covered so that the subjects cannot see them. The heads of the subjects are directed in three ways, as shown in Fig. 8: to the front, to the rear, and sideways toward the loudspeaker group. Pink noise was used as the stimulus (continuous for 2 seconds), and the loudness level was 70 dB. Scheffe's Paired Comparison was used as the evaluation method. Figure 9 illustrates the presentation method. Sound "A" was played back from one of the seven loudspeakers shown in Fig. 7. Sound "B" was then played back from one of the other six loudspeakers (the loudspeaker used to play sound "A" was excluded). Note that the pairs of sounds were presented to the subjects only once, in random order. The subjects were asked to judge the amount by which sound "B" was higher or lower than sound "A" using a scale of seven levels, separated by intervals of 0.5, as shown in Table 2. Prior to the experiment, the subjects were shown Fig. 8 and instructed to make a judgment as to whether the vertical angle was higher or lower, ignoring all other elements, such as the tone and horizontal distance of the sound image. A total of 15 subjects (13 audio specialists,

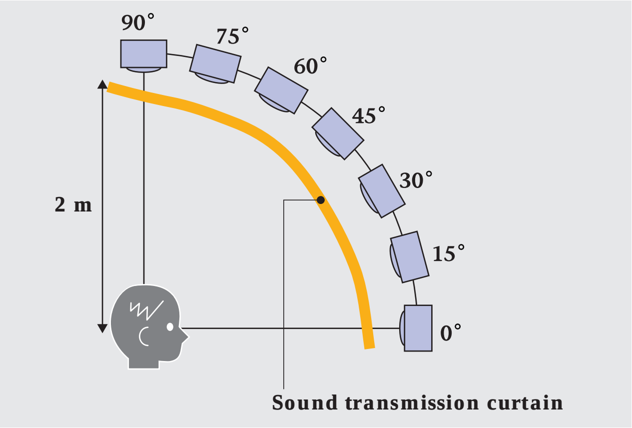

Figure 7: Loudspeaker positions used in experiment

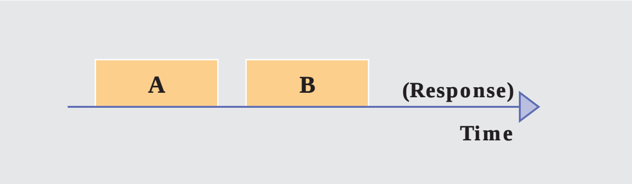

Figure 9: Presentation method for stimulus pairs for subjects

Table 2: Seven score levels

3	Much higher	-1	Slightly lower
2	Somewhat higher	-2	Somewhat lower
1	Slightly higher	-3	Much lower
0	Not higher or lower		

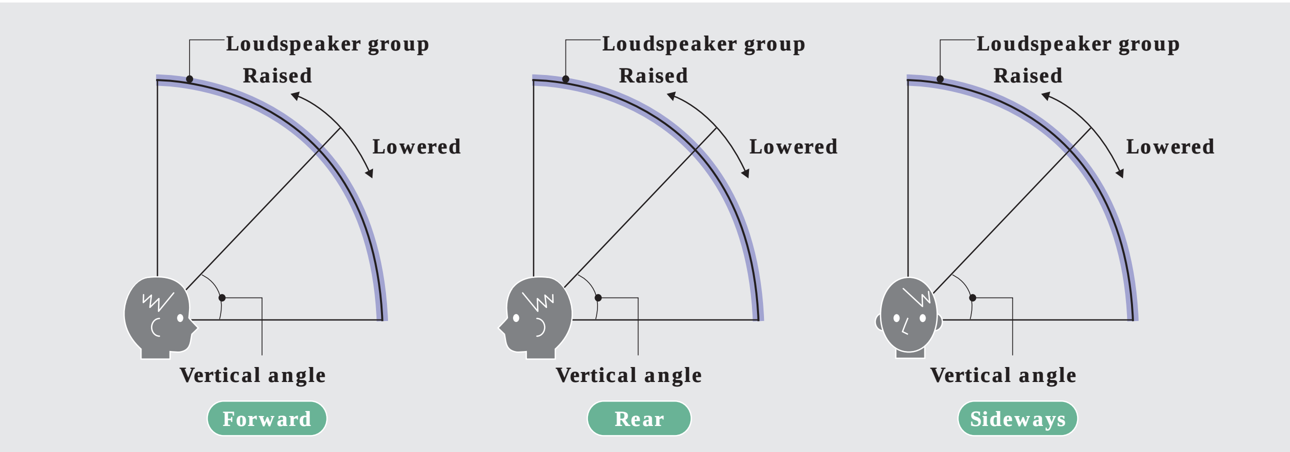

Figure 8: Audience conditions. Head orientation of subjects for loudspeaker groups

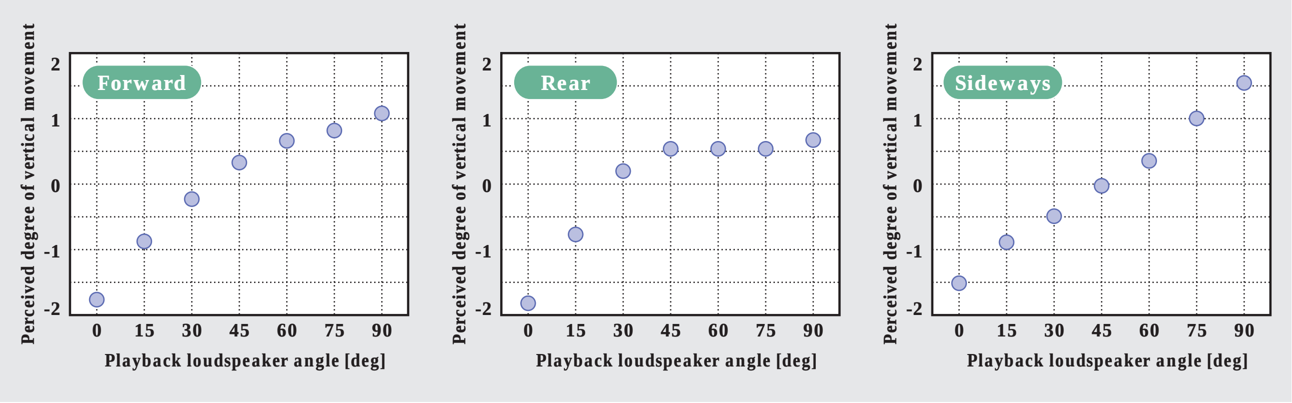

Figure 10: Perceived vertical displacement of sound image (subjective rating) for vertical angle of playback loudspeakers (physical quantity)

and two music university students) participated in the experiment.

Figure 10 presents the evaluation results. When the head is facing to the front or rear, the changes in subjective rating become small after the range of approximately 45 to 60°. Particularly, when the head is facing backward, there is no significant difference in subjective rating within 45 to 90°. On the other hand, when the head is facing sideways, the subjective rating increases more or less constantly right up to 90° (directly above). From these findings, we can draw the following conclusions:

- With head facing directly to front or rear: The perception of sound rising increases as the playback loudspeakers' height increases, but reaches saturation in the range of 45 to 60°.
- With head facing sideways: The perception of sound rising increases in proportion to the playback loudspeakers' height and does not reach saturation until sound comes from directly above. This shows that an effective way to create a perception of sound moving above from the side is to position a loudspeaker directly above the listener.

Based on these findings, we decided to set the angle of elevation of the upper layer loudspeakers in the 22.2 multichannel sound system at approximately 45°, relative to the center of the audience seats, with loudspeakers set directly above the audience.

5. Conclusion

This paper presented an outline of the 22.2 multichannel sound system that offers high fidelity over a wide range of audience seating positions, thereby making it suitable for Super Hi-Vision systems. We conducted experiments to evaluate the sense of realism that can be achieved with a 22.2 multichannel sound system, and found that such a system delivers greater realism over a wider listening range compared with conventional 5.1 channel theater sound systems.

In future research, we will investigate the synergistic effect of Super Hi-Vision images and 22.2 multichannel sound effects. In addition, we will conduct tests to standardize this sound format, by optimizing room sound characteristics for theater effects, loudspeaker frequency characteristics, etc. Furthermore, we hope to develop optimal sound pickup production methods, such as microphone technique for recording, addition of reverberation in production, etc. This 22.2 multichannel sound system was demonstrated at EXPO 2005, Aichi, Japan, as a sound system for Super Hi-Vision.

(Kimio HAMASAKI, Koichiro HIYAMA)