7 - 12 SEPTEMBER 2014, KRAKÓW
FORUM ACUSTICUM

Acoustic feedback canceller for the three-dimensional sound field simulator using a "Sound Cask"

Syunya KOHNO, Hisato YOSHIDA, Yusuke IKEDA, Yuko WATANABE, Shiro ISE
School of Information Environment, Tokyo Denki University, Chiba, Japan/CREST, JST, Japan

Summary
We have developed a three-dimensional sound reproduction system based on the boundary-surface control principle. "Sound Cask" comprises a 96-channel loudspeaker system to reproduce the sound field. It has a space large enough to allow a small musical instrument such as a violin to be played. Using Sound Cask, we aim to construct a three-dimensional sound-field simulator. Furthermore, we also propose a sound-field-sharing system that enables telecommunication as if we were in the same place. However, acoustic feedback occurring in the system owing to the presence of microphones causes an echo and a degradation of simulated reverberation characteristics. In this study, we aim to suppress the acoustic feedback by applying additional control points called "null spaces" at the positions of the microphones.

PACS no. 43.55.Lb, 43.60.Dh, 43.60.Pt

1. Introduction

Recently, we have proposed a sound-field-sharing system using a three-dimensional sound reproduction technique based on the boundary-surface control (BoSC) principle[1] to realize a telecommunication system that can transmit another person's presence. The BoSC system consists of a recording microphone array (BoSC microphone array) and a sound-field reproduction system, referred to as "Sound Cask." Sound Cask, an immersive sound-field reproduction system with 96-channel loudspeakers, allows a listener to move his or her head freely.

A sound-field-sharing system has been introduced by connecting several Sound Casks through a network. In the sound-sharing system, the voice or musical performance is first recorded in one of the Sound Casks, and it is transmitted and reproduced in others. At the same time, the same recording and reproduction procedure is implemented for one of the other Sound Casks, providing the feeling of the shared-sound field to the listener.

In this case, two types of acoustic feedback occur owing to the installation of microphones inside the Sound Cask. This feedback causes an echo and leads to instability of the system, thereby degrading the accuracy of the reproduced sound field.

In this study, we introduce an acoustic-feedback-suppression method by manipulating the inverse system design algorithm, in which we introduce an additional control point, called a "null space," where summation of all signals fed from the speakers is equal to zero[2].

In this paper, we first outline the BoSC principle and the sound-sharing system and describe the method of acoustic feedback suppression. We then evaluate the relationship between the suppression level of the acoustic feedback and the position of the null space by computational calculations. Furthermore, the effect of the additional null space is also observed in terms of the accuracy of the reproduced sound field. Finally, we implement the feedback canceller into the BoSC system and examine its applicability.

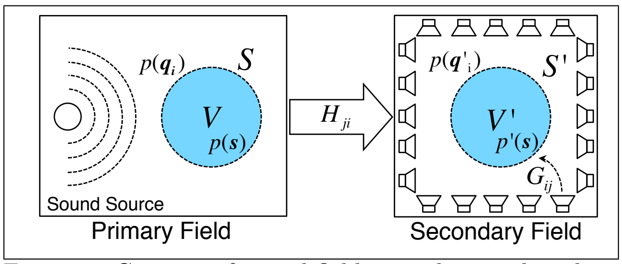

Figure 1. Concept of sound-field reproduction based on the boundary-surface control principle.

2. Sound Cask

2.1. BoSC principle

Figure 1 shows the basic concept of the sound-field reproduction system based on the BoSC method. First we consider volumes V and $V'(\equiv V)$ enclosed by boundary surfaces S and S' in the primary and secondary fields, respectively. Here, we assume that volume V and surface S are congruent with $V'(\equiv V)$ and S', respectively. According to the BoSC principle, if we measure the sound pressures and the particle velocities at surface S, and reproduce them at surface S', the sound field in volume V will be perfectly reproduced in volume V'[3, 4].

Now we consider reproducing signals recorded using M microphones on surface S by N loudspeakers installed in the secondary field. Let $[X_j](\in \mathbb{C}^{1\times M})$ be a recorded signal vector with the BoSC microphone array in the primary field. G_{ij} denotes a transfer function between the i-th loudspeaker and the j-th microphone on boundary surface S' in the secondary field, and $[G_{ij}](\in \mathbb{C}^{N\times M})$ is the transfer function matrix, whose inverse system matrix is $[H_{ji}](\in \mathbb{C}^{M\times N})$.

Based on the BoSC principle, a signal vector $[Y_j](\in \mathbb{C}^{1\times M})$ at surface S' is

$$[Y_j] = [X_j][H_{ji}][G_{ij}] \; (j = 1, \ldots, M, i = 1, \ldots, N). \quad (1)$$

as shown in Figure 2. To reproduce the sound pressure, which is measured at the boundary surface of the primary field, we must seek $[H_{ji}]$ to be $[Y_j] = [X_j]$ in the secondary field.

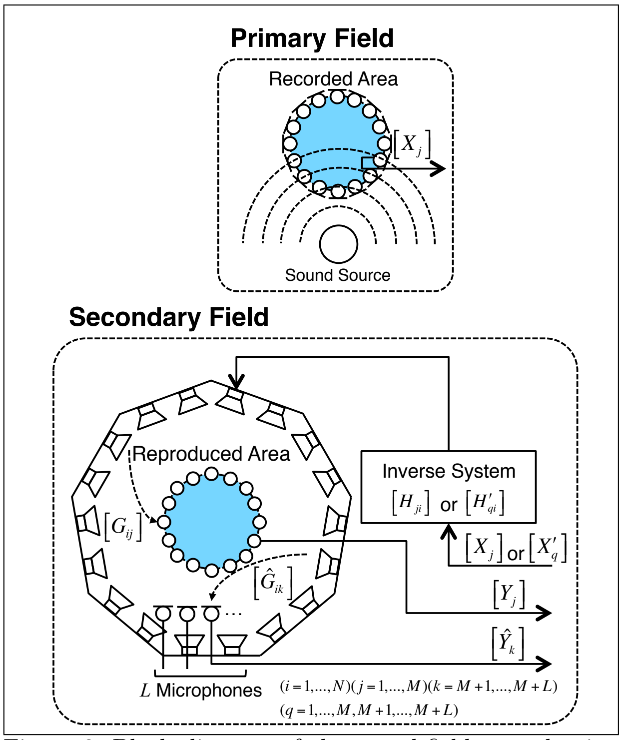

Figure 2. Block diagram of the sound-field reproduction system with L null spaces at the position of the source microphones.

2.2. BoSC system and feedback problem

A photograph of our 96-channel sound-field reproduction system based on the BoSC principle ("Sound Cask") is shown in Figure 3. Sound Cask has an inner space large enough to allow wind and stringed musical instruments to be played inside it.

Figure 3. Sound Cask.

We have also developed a BoSC microphone array used for measuring sound pressure on the boundary surface of the primary/secondary field as shown in Figure 4. The shape of the BoSC microphone array is designed to have the same configuration as C_{80} fullerene. 80 omnidirectional microphones (DPA 4060) are installed at the nodes of the fullerene. The diameter of the microphone array is about 46 cm and it is large enough to enclose a listener's head.

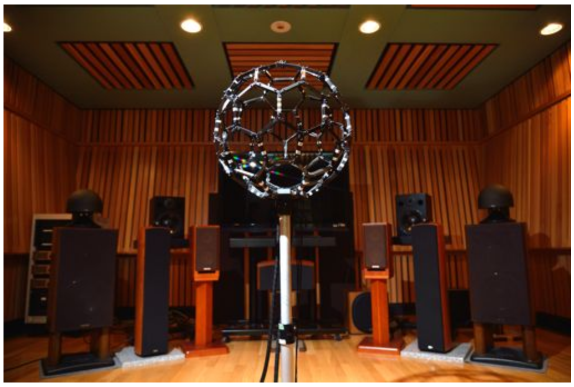

Figure 4. BoSC microphone array.

Additionally, a sound-field-sharing system using two or more BoSC (or Sound Cask) systems has been introduced. Figure 5 shows the concept of the system.

The sound-field-sharing system reproduces two types of sound through a shared sound field: 1. reverberation of one's own sound in the shared sound field and 2. the other's sound propagated through the shared sound field.

We first consider sound reproduction in system A when players A and B are inside their respective

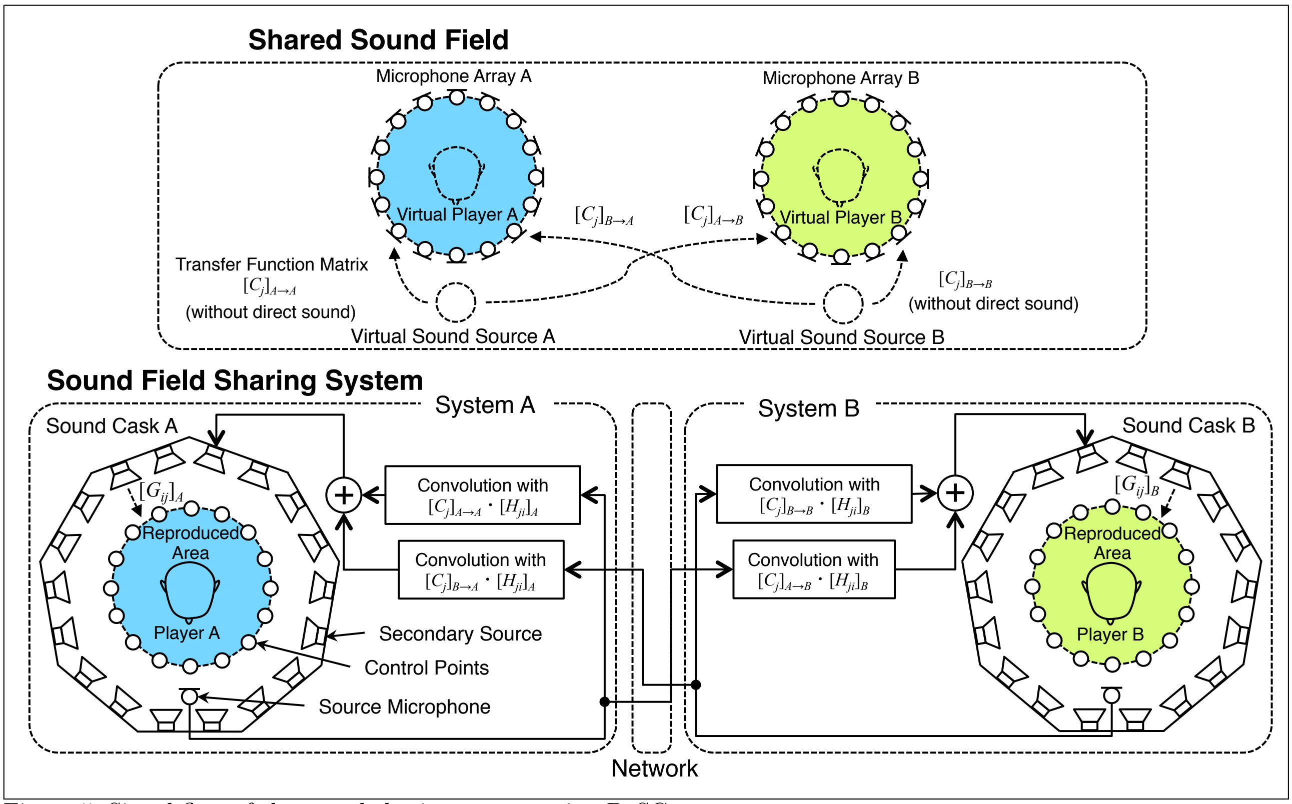

Figure 5. Signal flow of the sound-sharing system using BoSC systems.

Sound Casks A and B and they perform an ensemble. A musical signal from player A is recorded through a "source microphone," which is a microphone installed inside Sound Cask A.

Here, let $[C_j]_{A\to A}(\in \mathbb{C}^{1\times M})$ be transfer functions without direct sounds between virtual sound source A and microphone array A in the shared sound field. The recorded musical signal of player A with the source microphone is convolved with $[C_j]_{A\to A}\cdot[H_{ji}]_A$ to reproduce the reverberation of the shared sound field with the player's own sound. $[H_{ji}]_A(\in \mathbb{C}^{M\times N})$ is the inverse system of the transfer function matrix $[G_{ij}]_A$ between the secondary sources and the control points in Sound Cask A.

Next, we consider the reproduction of sound played by player B for Sound Cask A. A musical signal from player B is similarly recorded with a source microphone in Sound Cask B. Now, let $[C_j]_{B\to A}(\in \mathbb{C}^{1\times M})$ be a transfer function matrix from virtual sound source B to microphone array A. The recorded musical signal of player B is transmitted to system A via a network and then convolved with $[C_j]_{B\to A}\cdot[H_{ji}]_A$ to reproduce the sound of player B with the reverberation in the shared sound field. Finally, the summation of two kinds of convolved signals is played by the secondary sources in Sound Cask A.

In system B, the signals are reproduced in the same way as in system A. In Figure 5, $[C_j]_{B\to B}(\in \mathbb{C}^{1\times M})$ is the transfer function matrix without direct sounds between virtual sound source B and microphone array B in the shared sound field, $[H_{ji}]_B(\in \mathbb{C}^{M\times N})$ is the inverse system of the transfer function matrix $[G_{ij}]_B$ between the secondary sources and the control points in Sound Cask B, and $[C_j]_{A\to B}(\in \mathbb{C}^{1\times M})$ is the transfer function matrix from virtual sound source A to microphone array B.

Two kinds of feedback occur in this sound-field-sharing system. We consider the case of system A. The first feedback occurs through playing the convolution signal of player A's musical performance and $[C_j]_{A\to A}\cdot[H_{ji}]_A$ by the loudspeakers in system A. It causes the second feedback, in which player A's musical performance is transmitted to system B and then is returned via the sound field of system B with the musical performance of player B. Predicting the second feedback signal is more difficult because the path of feedback is more complicated.

Rokutanba et al. developed a sound-field simulation system using a six-channel recording and reproduction system and examined the acoustic feedback; however, the acoustic feedback was not a major problem because the reproduction field was an anechoic chamber[5].

In our study, loudspeakers and microphones are placed a short distance apart, and the reverberation within Sound Cask is not negligible; therefore, it is necessary to suppress the feedback to reproduce the sound field more accurately. These feedback sources occur with the inverse system. In this paper, we suppress them by improving the inverse system design.

3. Acoustic feedback canceller for Sound Cask

3.1. Sound-field control for feedback suppression

In this study, we achieve suppression of the feedback with additional control points called "null spaces"[2]. The sounds played by the loudspeaker system become zero at the null spaces. We set the null spaces at positions corresponding to all source microphones to suppress the feedback.

One well-known feedback canceller method entails using an adaptive filter. However, in the double-talk situation, the performance of the method is degraded owing to the increase of the estimation error[6]. Also, the coefficient of a multichannel feedback canceller using an adaptive filter does not converge owing to the large number of transmission paths[7]. Therefore, using an adaptive feedback canceller system with our sound-field-sharing system seems inadequate since the sound-field-sharing system would always be in the double-talk condition and have many transmission paths resulting from the 96-channel loudspeakers.

3.2. Inverse system design

In this section, we describe a method of an inverse system design with a feedback canceller.

First, let a reproduced signal matrix at the control points in the frequency domain using the inverse system without a null space be $[Y_j]$, which is given by equation 1.

Next, we consider the inverse system with L null spaces at the position of the source microphones. Figure 2 shows a block diagram of the sound reproduction system with the null spaces. Let $[\widehat{G}_{ik}](\in \mathbb{C}^{N\times(M+L)})$ be a transfer function matrix between the secondary sources and the source microphones.

A transfer function matrix $[G'_{iq}]$ is defined by

$$[G'_{iq}] = [G_{ij}\ \widehat{G}_{ik}](\in \mathbb{C}^{N\times(M+L)}), \tag{2}$$

where $k = 1, \ldots, L,\quad q = 1, \ldots, M, M+1, \ldots, M+L$. In the same manner, the recorded signal vector in the primary field and the output signal vector in the secondary field are defined by

$$[X'_q] = [X_j\ \widehat{X}_k]\ (\in \mathbb{C}^{1\times(M+L)}) \tag{3}$$

$$[Y'_q] = [Y_j\ \widehat{Y}_k]\ (\in \mathbb{C}^{1\times(M+L)}) \tag{4}$$

where $[\widehat{X}_k]$ is a signal vector on the additional control points in the primary field and $[\widehat{Y}_k]$ is a signal vector of the additional control points in the secondary field.

To reproduce the primary sound field and generate the null space, the following equation must be hold:

$$[E_{qq}] = [H'_{qi}][G'_{iq}], \tag{5}$$

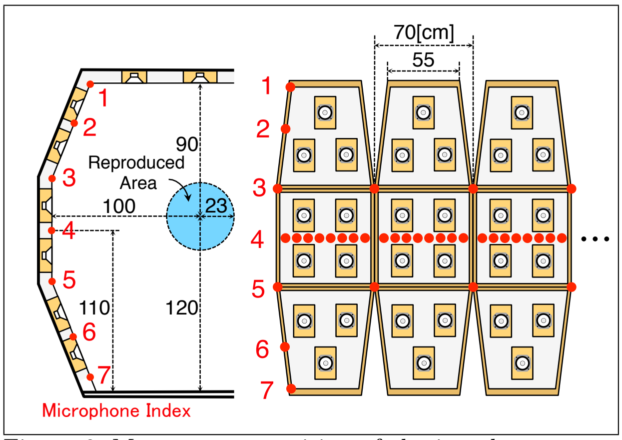

Figure 6. Measurement position of the impulse response (left: cross-section view; right: development view). The positions of the microphones are shown as red circles.

where $[E_{qq}]$ is

$$[E_{qq}] = \begin{cases} 1\ (q \le M) \\ 0\ (q > M). \end{cases} \tag{6}$$

That is to say, $[H'_{qi}](\in \mathbb{C}^{(M+L)\times N})$ is the inverse system with L null spaces.

The signal matrix $[Y'_q]$ at the control points in the secondary field with null spaces is

$$[Y'_q] = [X'_q][H'_{qi}][G'_{iq}]. \tag{7}$$

In equations 1–7, the solution of the inverse system is indeterminate, since $M < N$ and $M < N + L$. However, it is possible to obtain the inverse system by using the Moore-Penrose pseudo-inverse method, which gives a least norm solution[8].

4. Optimum position of the null space

4.1. Effect of the position of the null space on the suppression level

The source microphones should be installed near the inside wall of Sound Cask to avoid possible collisions with the player from body movement. However, the physical features of Sound Cask such as its acoustic modes cause differences in the suppression level achieved by the position of the null space. In this section, we discuss the optimum position of the null space in Sound Cask. We evaluate the relationship between the suppression level of the feedback and the position of a single null space by simulation.

First, we measure $[G_{ij}]$, the transfer function matrix between the 96 loudspeakers and the 80 control points in Sound Cask. Figure 6 shows a part of the measurement position. We choose 92 positions on the wall of Sound Cask, except on the door, as the position of the null space, and we install a microphone (DPA 4060) at each point. Then, we measure $[\widehat{G}_{i1}]$, which is the transfer function matrix between the 96 loudspeakers and each microphone.

We use a time-stretched-pulse (TSP) signal whose length is 65536 points as a measurement signal. The measurement sampling rate is 48 kHz and the synchronized averaging number of implies responses is 10.

Second, we choose a source point from 92 points on the wall and design $[H'_{qi}]$ to follow equations 2-6 for each source point. We use the minimum error relaxation algorithm[9] to design the inverse system.

The signal matrix on the control points using the inverse system with the null space is given by Equation 7.

In contrast, the signal matrix $[Y_q](\in \mathbb{C}^{1\times(M+L)})$ on the control points using the inverse system without the null space is given by

$$[Y_q] = [X_j][H_{ji}][G'_{iq}]. \tag{8}$$

In this section, let $[X'_q] = [G'_{iq}], [X_j] = [G_{ij}]$ to calculate $[Y'_q]$ and $[Y_q]$.

We evaluate the suppression level with each inverse system that has the null space. The suppression level E is defined by

$$E = 10\log_{10}\frac{\sum_{t=0}^{T}|y_{src}[t]|^2}{\sum_{t=0}^{T}|y'_{src}[t]|^2}[\mathrm{dB}], \tag{9}$$

where $y_{src}[t]$ and $y'_{src}[t]$ are the impulse responses of the output signal of a source microphone using the each inverse system without and with the null space, respectively, and t is a time index.

We conduct a simulation experiment using the impulse responses measured in Sound Cask. The inverse system is calculated in the frequency domain with a Fourier transform of length 8192. Then, we use 2048-sample-length impulse responses to calculate the inverse system. In the time domain, the inverse system is multiplied by the Hanning function to shorten it to a sample length of 4096.

Figure 7 shows the results of the suppression level at the leftmost positions (microphone indices 1–7) in Figure 6. Figure 7 also shows the maximum and minimum values of the suppression level on three horizontal planes that are at the same height as the third, fourth, and fifth microphones in Figure 6.

As Figure 7 shows, the inverse system achieves more than 17 dB suppression over all null space positions. Variations of the suppression level are caused by the difference of the positions of the null space only on the vertical plane.

4.2. Signal-to-noise ratio of the reproduced sound field

In general, in a multichannel sound reproduction system with a particular number of secondary sources, it is well known that, as the number of control points increases, the accuracy of the reproduced sound field decreases. Our BoSC system with the additional null space for feedback cancellation may thus suffer from degradation of its reproduction accuracy.

Therefore, we introduce the signal-to-noise ratio (SNR) as a means to evaluate the reproduced accuracy. The SNR in the reproduced area in the case of no null space is calculated by using the following equation:

$$\mathrm{SNR} = 10\log_{10}\frac{\sum_{q=1}^{80}\sum_{t=0}^{T}|x'_q[t]|^2}{\sum_{q=1}^{80}\sum_{t=0}^{T}|y_q[t]-x'_q[t]|^2}[\mathrm{dB}]. \tag{10}$$

where $x'_q[t]$ and $y_q[t]$ are the impulse responses of $[X'_q]$ and $[Y_q]$, respectively.

If we let $y'_q[t]$ be the impulse response of $[Y'_q]$, the SNR in the case of null space addition is calculated by substituting $y'_q[t]$ for $y_q[t]$ in equation 10.

From a numerical analysis, we obtain SNR about 22.9 dB both with and without a null space, regardless of the null space position. These results demonstrate that the accuracy of the reproduced sound field is not greatly affected by a single additional control point.

5. Suppression level to the actual sound source

We measure the suppression level in Sound Cask by playing sounds convolved with the inverse filter matrix including the null space.

We use two types of measurement signals. The first signal is simulated pink noise in the free-field condition. The BoSC microphone array is faced toward the point source located 1.5 m away. The height of point source is the same as the center of the BoSC microphone array. The second signal is the recorded signal of an orchestra on a stage using the BoSC microphone array.

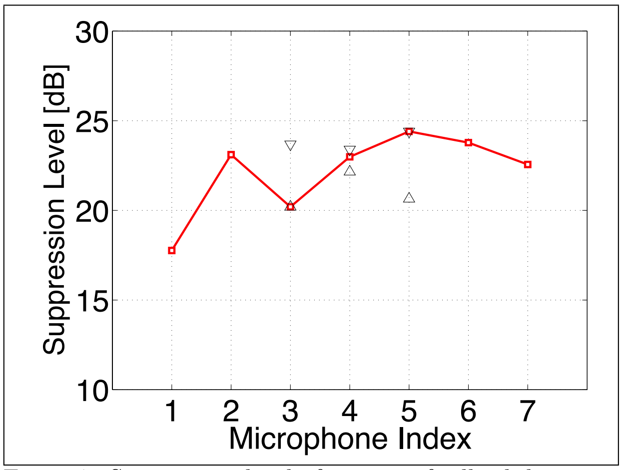

Figure 7. Suppression level of acoustic feedback by simulation. The minimum values of the results from each horizontal plane are plotted as △'s, and the maximum values are plotted as ▽'s.

The results of the simulation in the preceding section show that significant differences between the suppression levels of measurement points occur only in the vertical plane. Therefore, in this section, the measurement positions are the leftmost microphones in Figure 6. The suppression level is calculated by using equation 9. The durations of the pink noise and the musical signal are 10 and 25 s, respectively. We place a sound level meter at the center of the reproduced area in Sound Cask, and the average sound pressure levels of the sound reproduction using all inverse systems are 80 dB.

Figure 8 shows the suppression levels of the feedback in Sound Cask using the pink noise and the musical signal. The figure shows that the suppression levels with pink noise are more than 10 dB at any position. The figure also shows that the suppression levels with the musical signal are about 20 dB at any position. Therefore, feedback with musical signals could be more suppressed than with pink noise.

Next, Figures 9 and 10 show the results of an octave band analysis at the fourth microphone using the pink noise and the musical signal, respectively. The signal can be suppressed over all ranges of frequency bands. Specifically, the suppression level of the center frequency of 500 Hz is about 30 dB in both measurement signals. However, the suppression level is low in the higher frequency band.

Figure 8 shows that the suppression levels with the musical signal are higher than those with the pink noise, because the pink noise has more energy in the high-frequency bands than does the musical signal. Therefore, when we use this method for ensemble performances with musical instruments, feedback is adequately suppressed.

6. Conclusion

In this paper, we introduced an acoustic feedback method in which an additional control point was applied to create a silent space for the source microphone. To evaluate our proposed method, we observed the suppression level of feedback through both computer simulation and actual measurement. With reference to our discussion, we confirmed that our proposed method could provide a signal suppression level of more than 17 dB at any position of the null space. While the suppression level varied with corresponding to the position in the vertical plane.

We also considered the effect of adding control points on the accuracy of sound-field reproduction by using the SNR; the results showed that degradation in the accuracy resulting from the additional null spaces was negligible.

Actual measurements using pink noise and music were taken, and the suppression efficiency was found to depend on the frequency components of the sound source, though this slight dependence would not pose a major problem in practical use.

Further investigation of both the sound-field simulator system and the sound-sharing system should be pursued. Systems with multiple null spaces should also be considered for cooperative research aimed at directivity reproduction of a sound source[10].

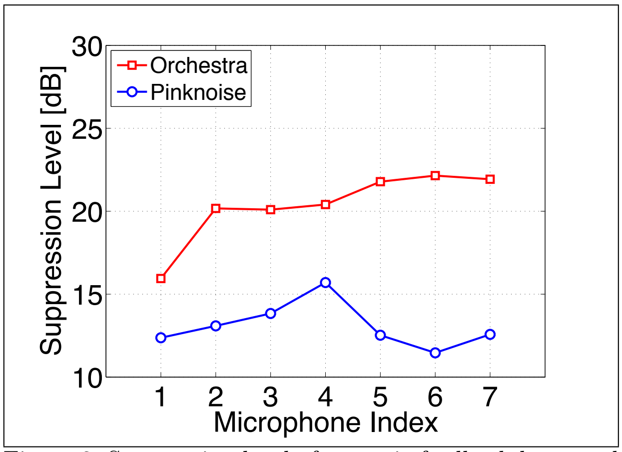

Figure 8. Suppression level of acoustic feedback by actual measurement. Pink noise is plotted as blue markers, and the orchestra is plotted as red markers.

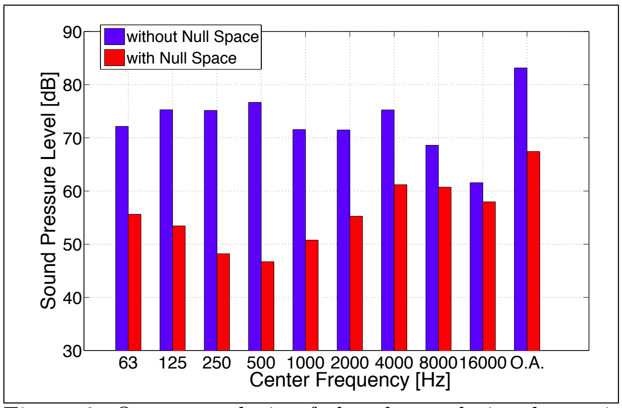

Figure 9. Octave analysis of the observed signal at microphone index 4 when using pink noise as the primary signal.

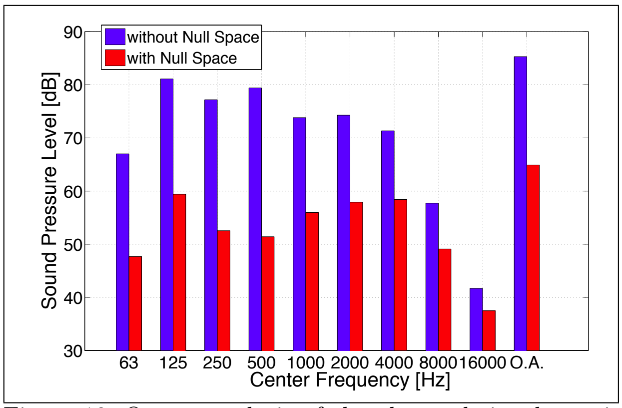

Figure 10. Octave analysis of the observed signal at microphone index 4 when using the orchestra as the primary signal.

References

[1] Yusuke IKEDA, Shiro ISE: Sound Cask: Music and voice communications system with three-dimensional sound reproduction based on boundary surface control principle. The 21st International Congress on Acoustics (ICA2013), 2013.

[2] Yoichi HINAMOTO: Interface for barge-in free spoken dialogue system based on sound field reproduction and microphone array. International Conference on Acoustics, Speech and Signal Processing, 2003.

[3] Shiro ISE: A study on the sound field reproduction in a wide area(1) -based on Kirchhoff-Helmholtz integral equation (in Japanese). Proc. Meeting of Acoustical Society of Japan., 479–480, 1993

[4] Shiro ISE: A principle of sound field control based on the Kirchhoff-Helmholtz integral equation and the theory of inverse systems. Acustica 85, 78–87, 1997.

[5] Motoko ROKUTANDA et al.: A sound field simulation system for the study of ensemble performance on a concert hall stage, Acoustical Science and Technology, Vol. 25, No. 5, 373–378, 2004.

[6] Eberhard HÄNSLER, Gerhard SCHMIDT: Topics in Acoustic Echo and Noise Control. Springer, Berlin, Heidelberg, 2006.

[7] Akihiko SUGIYAMA: Multichannnel echo cancellation: Problems and challenges to a solution (in Japanese). The IEICE Transactions of Fundamentals of Electronics, Communications and Computer Sciences, Vol. 81, No. 3, 266–274, 1998.

[8] Atsunobu KAMINUMA, Shiro ISE, Kiyohiro SHIKANO: A method of designing inverse system for multichannel sound reproduction-system using least-norm-solution. Active 1999, Florida, US, 1999.12.

[9] Yusuke KAIHARA et al.: Minimum error relaxation algorithm of inverse filter in multichannel sound reproduction system (in Japanese). IEICE Technical Report, EA, 105(556), 7–12, 2006.

[10] Yusuke IKEDA and et al.,:Design of inverse filters for multi-channel directivity reproduction of a sound source -Application to telecommunication system using sound casks-, Proc. Forum Acusticum, Krakow, Poland, 2014.