inter.noise 2014
MELBOURNE AUSTRALIA
16-19 NOVEMBER

Narrow area control for individual sound image generation by combining NBSFC and liner loudspeaker array

Yumiko NAKAYAMA[1]; Yosuke TATEKURA

Shizuoka University, Japan

ABSTRACT

In this paper, we propose a method for reproducing different sound for multiple users by forming a narrow area, where the users can listen only the desired sounds. We have proposed a sound field control method for narrow area by combining Null-space Based Sound Field Control (NBSFC) and liner loudspeaker array. NBSFC filtering can suppress the sound pressure on the each control point and maintain the sound pressure around the user's ear. However, consideration of the effective arrangement of control points was insufficient. Therefore, we aim to improve the reproduce performance of individual sound image generation by optimizing the arrangement of control points. The numerical simulation showed that the averaged power difference between desired and undesired area is by about 2 dB by the proposed method.

Keywords: NBSFC, Inverse filter, Sound field control, Underdetermined condition
I-INCE Classification of Subjects Number(s): 74.3

1. INTRODUCTION

It is useful to achieve a system that can reproduce different sounds for multiple users in the same environment simultaneously. For example, such a system can be applied to the car audio system, in which a driver can listen only the sound of the car navigation system while other passengers can listen their favorite music. However, a sound reproduced by a typical single loudspeaker spreads in all directions. Therefore, when different sounds are reproduced from multiple loudspeakers simultaneously, each sounds are mixed in space, so that a listener hears the mixing sound.

To achieve above-mentioned system, any studies have been performed. One of them, for example, is a method of forming the directivity by using a ultrasonic parametric loudspeaker or a method to increase the sound pressure in the specified directivity by array signal processing. In the former case, it is necessary to emit sound waves in very large sound pressure compared with the effects on the human body is concerned as a problem in ultrasonic exposure (1). In the latter case, there is not only the main lobe to the desired direction but also side lobe to the undesired direction. A method of using inverse filtering, it can reproduce sound on only user's ear(2). However, reproduced condition of sound cannot be guaranteed in the area except the ears. Other method which reproduce a sound in wide area based on boundary sound field control has been represented, in which the controlled field is surrounded by the control points. Many large number of control points is required to ensure a wide listening area increases, the system becomes a large scale.

Therefore, to overcome these problems, we have proposed a method called as "individual sound image generation" (3)–(5). The proposed method is combined loudspeaker array that increase sound pressure in the specified directivity and null-space based sound field control (NBSFC) that suppresses sound pressure on any control points (6)–(7). This method provides "sound wall" which prevents undesired sound by arranging any control points, which has minimum sound pressure, between users, and listening area that user can listen only desired sound is formed. However, the side lobe of the reproduced sound form different position in each frequencies. The control points in the same positions are used at all frequencies, it is difficult to suppress the contamination of undesired sound especially at high frequencies. In this paper, we propose the method that suppress the contaminated sound pressure to undesired area more effectively by switching arrangement of control points in each frequencies, and evaluate performance of individual sound image generation. To achieve this objective, we extend the NBSFC filtering technique to underdetermined conditions which the number of loudspeakers is less than the number of control points by truncating minimum singular value of

[1] y-nakaya@spalab.eng.shizuoka.ac.jp

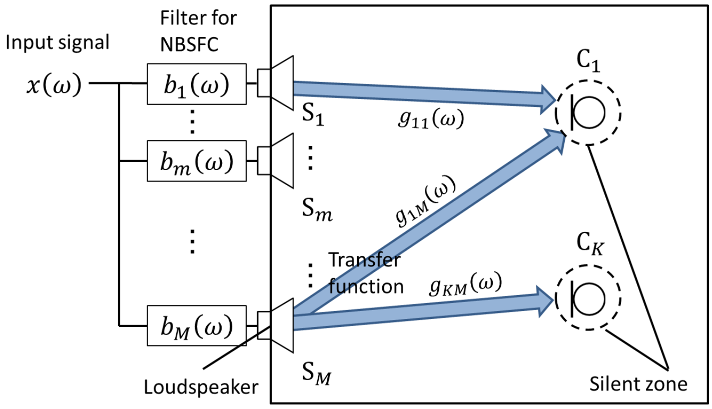


Figure 1 – Configuration of sound pressure suppression on each control points by NBSFC.

the room transfer function matrix.

2. CONVENTIONAL NULL-SPACE BASED SOUND FIELD CONTROL

Null-space Based Sound Field Control (NBSFC) is a sound field control technique according to the signal processing while maintaining the sound pressure outside of any control point and minimizing the sound pressure on the control points. In this study, by using NBSFC, we aim to reduce the interference sound between users with reproducing the desired sound for each user.

Fig. 1 illustrates the configuration of NBSFC with M loudspeakers and K control points, where S_m $(m = 1, \cdots, M)$ and C_k $(k = 1, \cdots, K)$ indicate loudspeaker and control point, respectively. Here, we assume that the number of loudspeakers is greater than the number of control point $(M > K)$ as in usual condition. We define the observed signal matrix representing the observed signal $y_k(\omega)$ at microphone C_k as $y(\omega) = [y_1(\omega), \cdots, y_K(\omega)]^T$, the original sound source signal as $x(\omega)$, the matrix representing m-th filter coefficient $b_m(\omega)$ as $b(\omega) = [b_1(\omega), \cdots, b_M(\omega)]^T$ and the matrix representing the room transfer function between m-th loudspeaker and k-th control point $G_{km}(\omega)$ as

$$G(\omega) = \begin{bmatrix} G_{11}(\omega) & \dots & G_{1M}(\omega) \\ \vdots & \ddots & \vdots \\ G_{K1}(\omega) & \dots & G_{KM}(\omega) \end{bmatrix}, \tag{1}$$

respectively. ω denotes the frequency, and $\{\cdot\}^{\mathrm{T}}$ indicate transposed matrix, respectively. We can represent the observed signal as

$$y(\omega) = G(\omega)b(\omega)x(\omega). \tag{2}$$

Here, in order to control so that $y(\omega) = 0$, $b(\omega)$ is obtained to satisfy the following condition;

$$G(\omega)b(\omega) = 0. \tag{3}$$

Note that the constraint condition $||b(\omega)|| = C$ is satisfied in order to avoid the output sound signal from the loudspeaker to be zero, where C is constant to adjust the gain.

To determine $b(\omega)$, the vector which spans the nullspace of $G(\omega)$ is extracted by the singular value decomposition (SVD). The singular value decomposition of $G(\omega)$ is expressed as

$$G(\omega) = U(\omega)[\Lambda(\omega) \mid O_{K,M-K}]V^{\mathrm{H}}(\omega), \tag{4}$$

where $O_{K,M-K}$ is a $K \times (M-K)$ null matrix, $\{\cdot\}^{\mathrm{H}}$ means the Hermitian transpose, $\Lambda(\omega)$ is a $K \times K$ diagonal matrix with the diagonal elements of the singular value, $U(\omega)$ and $V(\omega)$ are $K \times K$ and $M \times M$ orthogonal matrix, respectively. Here, when we assume R_ω is a rank of $G(\omega)$, the eigenvector of remaining after removing a column vector of R_ω from $V(\omega)$, $\left[v_{R_\omega+1}(\omega), ..., v_M(\omega)\right]$, can be extracted as follows:

$$W(\omega) = [v_{R_\omega+1}(\omega), ..., v_M(\omega)]. \tag{5}$$

However, because the pre-echo or post-echo are generated in this state, the reproduced sound is distorted significantly. In order to maintain a better sound quality, an impulse with equal amplitude and latency τ of

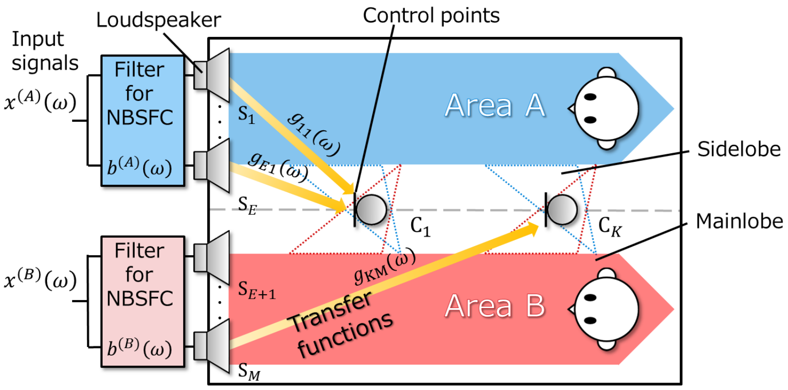


Figure 2 – Configuration of individual sound image generation for two users by the proposed method.

the target filter is defined as

$$l(\omega) = e^{j\omega\tau}[\underbrace{1,\cdots,1}_{M}]^T. \tag{6}$$

The output sound distortion can be reduced by using a filter coefficient close to the impulse because it has an all-pass and liner phase characteristics. Based on this procedure, the desired filter coefficient $b(\omega)$ can be expressed as follows:

$$b(\omega) = C\frac{W(\omega)W^{\mathrm{H}}(\omega)l(\omega)}{\sqrt{l^{\mathrm{H}}(\omega)W(\omega)W^{\mathrm{H}}(\omega)l(\omega)}}. \tag{7}$$

3. INDIVIDUAL SOUND IMAGE GENERATION WITH NBSFC

3.1 Basic strategy

In this section, we explain the method of achieving the individual sound image generation for multiple users by loudspeaker array and NBSFC. Figure 2 shows the configuration of the proposed individual sound image generation. Generally, when the same sound wave is reproduced from each loudspeaker constituting a loudspeaker array which is arranged in a straight line, side lobe is generated in the left and right as well as the main lobe of the sound pressure to the front of the loudspeaker array. In reference (4), although the sound pressure is increased in listening area by suppressing the side lobe using NBSFC, the approach is to form a desired listening area for only one user.

In this study, to form desired listening areas for multiple users, we try to reduce the interference in the listening area of each sound by arranging the control points, in which the sound pressure minimizes, so as to separate the users. Namely, each side lobes formed by the loudspeaker arrays is suppressed by each system. Therefore, we can expect to achieve the individual sound image generation for multiple users with relaxing the movement restrictions of the user.

3.2 Formulation

Hereafter, we assume that the enclosed region where exist two users is divided into two areas; Area A and Area B as shown in Fig. 2, in which the number of loudspeakers and the number of microphones are M and K, respectively. When the number of loudspeakers used in Area A is E, the number of loudspeakers used in Area B can be represented as $M-E$.

The filter coefficients by the proposed method by NBSFC in Area A and Area B are represented as follows:

$$b^{(A)}(\omega) = [b_1^{(A)}(\omega),\cdots,b_E^{(A)}(\omega)]^{\mathrm{T}} \tag{8}$$

$$b^{(B)}(\omega) = [b_1^{(B)}(\omega),\cdots,b_{M-E}^{(B)}(\omega)]^{\mathrm{T}}, \tag{9}$$

respectively, where $\cdot^{(A)}$ denotes an index for Area A and $\cdot^{(B)}$ denotes that for Area B. In addition, the transfer functions for designing NBSFC filters of each area are defined as follows;

$$G^{(A)}(\omega) = \begin{bmatrix} G_{11}(\omega) & \dots & G_{1E}(\omega) \\ \vdots & \ddots & \vdots \\ G_{K1}(\omega) & \dots & G_{KE}(\omega) \end{bmatrix}, \tag{10}$$

$$G^{(B)}(\omega) = \begin{bmatrix} G_{E+1}(\omega) & \dots & G_{1M}(\omega) \\ \vdots & \ddots & \vdots \\ G_{K(E+1)}(\omega) & \dots & G_{KM}(\omega) \end{bmatrix}, \tag{11}$$

respectively. Thus, the observed signal $\hat{X}_P(\omega)$ on arbitrary position P in the enclosed region is obtained as

$$\hat{X}_P(\omega) = G_P(\omega)\{b^{(A)}(\omega)X^{(A)}(\omega) + b^{(B)}(\omega)X^{(B)}(\omega)\} \tag{12}$$

by using these NBSFC filters, where $X^{(A)}(\omega)$ and $X^{(B)}(\omega)$ is input signal for each area, and $G_P(\omega)$ represent the room transfer function matrix between loudspeakers and an arbitrary position P.

3.3 Extension to the underdetermined condition of NBSFC filter design method

The conventional NBSFC filter is assumed overdetermined condition which means the number of loudspeakers is greater than that of control points. Therefore, there is a limit of the number of control points which suppress sound pressure from the loudspeaker array. So that, it is difficult to suppress all side lobe especially at high frequency band. In this section, we propose the method to design NBSFC filter in underdetermined condition which means the number of control points is greater than that of the loudspeakers.

At first, $G(\omega)$ is decomposed by SVD as

$$G(\omega) = U(\omega) \begin{bmatrix} \Lambda(\omega) \\ O_{K-M,M} \end{bmatrix} V^{\mathrm{H}}(\omega). \tag{13}$$

When the rank of $G(\omega)$, R_ω, is equal to M, we can not extract column vector $\left\{v_{R_\omega+1}(\omega), \cdots, v_M(\omega)\right\}$ of $V(\omega)$ as null-space. Therefore, we truncate the minimum singular value of the transfer function matrix $G(\omega)$ to be zero for extracting null-space. We assume that t is defined as the number of truncations of singular value, null-space $W(\omega)$ is extracted as

$$W(\omega) = [v_{R_{\omega-t+1}}(\omega), \cdots, v_M(\omega)]. \tag{14}$$

The multiplication of $W(\omega)$ and $v_s(\omega)$ $(s = R_{\omega-t+1}, \cdots, M)$ is expressed as

$$\begin{aligned} G(\omega)v_s(\omega) &= U(\omega) \begin{bmatrix} \Lambda(\omega) \\ O_{K-M,M} \end{bmatrix} V^{\mathrm{H}}(\omega)v_s(\omega) \\ &= U(\omega) \begin{bmatrix} \Lambda(\omega) \\ O_{K-M,M} \end{bmatrix} \cdot [\underbrace{0,\dots,0}_{s-1}, 1, \underbrace{0,\dots,0}_{M-s}]^{\mathrm{T}} \\ &= U(\omega)[\underbrace{0,\dots,0}_{s-1}, \lambda_s, \underbrace{0,\dots,0}_{K-s}]^{\mathrm{T}}. \end{aligned} \tag{15}$$

From this, $V^{\mathrm{H}}(\omega)v_s(\omega)$ is not equal to 0 because a component of the singular value remains in $[0,\dots,0,\lambda_s,0,\dots,0]^{\mathrm{T}}$. However, the components of $G(\omega)v_s(\omega)$ is approximated to zero elements when the truncated singular value can be approximated. Therefore, we can design NBSFC filter in underdetermined conditions.

4. NUMERICAL EVALUATION

To evaluate the performance of the proposed individual sound generation, numerical simulations were carried out. This section expresses the results of the simulations.

4.1 Experimental conditions

Figure 3 shows the arrangement of apparatus for the numerical simulations. In the simulations, we assume the enclosed region 2.00 m in length and 1.50 m horizontal, in which the floor and ceiling are not considered. Here, the left side of the region (Width: 0.00–0.75 m) is regarded as Area A, and right side of the region (Width: 0.75–1.50 m) is regarded as Area B. Eight loudspeakers are arranged in each area in a straight line at intervals of 0.10 m.

Figure 4 shows the arrangement patterns of the NBSFC control points. Using above, we have compared four patterns; (Arr. I) NBSFC for seven points control and interspace of control point is 0.15 m, (Arr. II) NBSFC for ten points control and interspace of control point is 0.10 m, (Arr. III) NBSFC for nineteen points

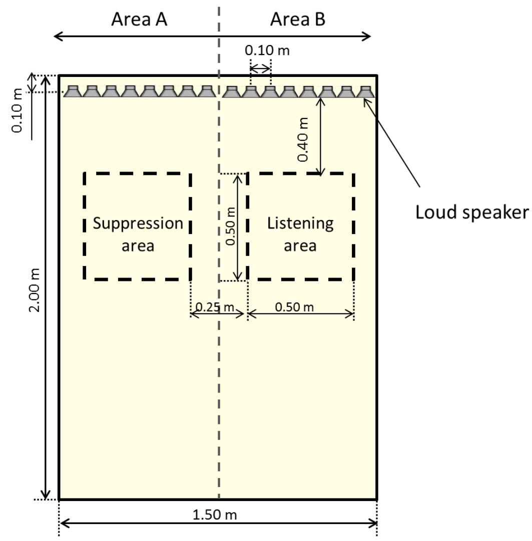


Figure 3 – Arrangement of the apparatus for numerical evaluation of the proposed method.

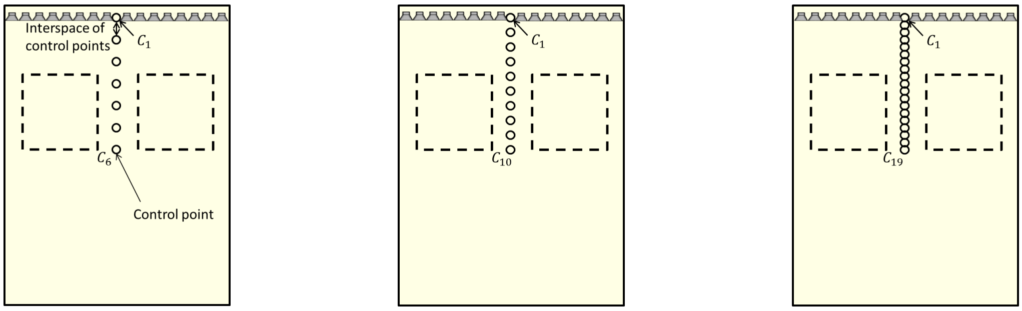


Figure 4 – The arrangement patterns of the NBSFC control points.

control and interspace of control point is 0.50 m. In addition, we used the pattern of switching arrangement of control points each frequency band as Arr. IV: four control points at 0.30 m intervals in 150–1000 Hz, Arr. I in 1000–1500 Hz, Arr. II in 1500-2000 Hz, and Arr. III in 2000-3850 Hz. As the input signals, three kinds of male and female speeches are reproduced in Area B and zero signal is reproduced in Area A. Therefore, we define Area A as suppression area and Area B as listening area in the simulations, and other conditions between areas are omitted in this paper. If sound pressure is not zero in Area A, it is due to the reproduced sound from Area B. In addition, we define the evaluation areas inside the dashed line box as shown in the Fig. 3.

The simulation conditions are listed as follows: the sampling frequency, the impulse response, the tap length of NBSFC filter and the passband range are 8000 Hz, 256 points, 4096 points and 150–3850 Hz, respectively.

Next, we describe a procedure of calculating the observed signals. We have divided in grid of 0.01 m of the assumed space shown in Fig. 3 and have calculated the impulse response from each loudspeakers to each grid points by image method(8), where the reflection of sound by wall is considered only primary reflection and there is no attenuation of sound by reflection. Then, NBSFC filters are designed based on the impulse responses between the control points and the loudspeakers, and the observed signal $\hat{X}_P(\omega)$ on arbitrary grid point P by Eq. (12) was calculated.

4.2 Evaluation score

We evaluate suppression effectiveness of the undesired sound by result of the sound pressure distribution. First, the power distribution is obtained from the observed signals on each grid points. The reproduced power on the grid points in the listening area and suppression area are regarded as $P_{lis}^{(i)}$ and $P_{sup}^{(j)}$, respectively, where i and j denote index of the grid point. Then, to evaluate the performance of each method, we introduce the averaged power difference between listening area (Area B) and suppression area (Area A) as evaluation score,

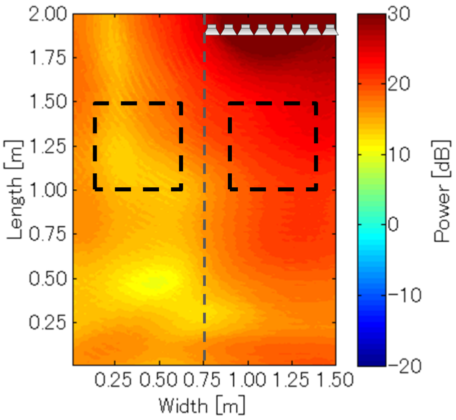


Figure 5 – Sound pressure distribution in unprocessed case.

which is defined by

$$D = \frac{1}{L}\sum_{i=1}^{L} P_{lis}^{(i)} - \frac{1}{S}\sum_{j=1}^{S} P_{sup}^{(j)} \text{ [dB]}, \tag{16}$$

where L and S are the number of grid points in the listening area and the suppression area, respectively. When the evaluation score D is large, we can consider that mixing of a speech sound for listening to the suppression area can be prevented and the desired speech sound can be listened with agreeableness in the listening area because the power difference between each area is large.

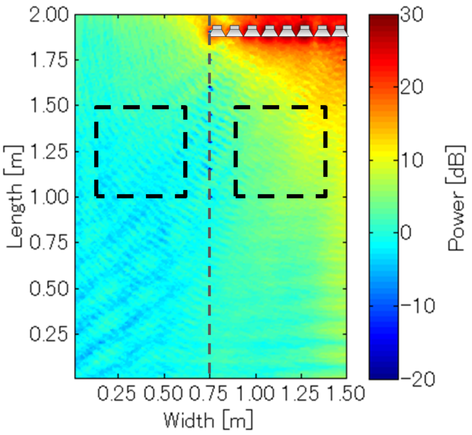
4.3 Results

Figure 5 and 6 show the sound pressure distribution of unprocessed case and each control point arrangements cases, in which a male speech is reproduced. In unprocessed case shown in Fig. 5, high power value can be shown not only throughout the Area B but also in Area A. This reveals that the sound reproduced to Area B is mixed even in Area A. On the other hand, Fig. 6 shows that the reproduced power is reduced in Area A by using NBSFC filtering in each cases. In particular, in case of Arr. III, the sound suppression performance in Area A is improved by arranging control points in closely. However, the sound pressure is suppressed not only Area A, but also Area B. Moreover, we can confirm that large sound pressure area is formed towards the side wall rather than in front of loudspeaker array. Therefore, it is concerned that the observed sound is uncomfortable in evaluation area, because sound pressure in the evaluation area is uneven. In case of arr. IV, the sound pressure in Area A is large compared to arr. I–III, but the sound pressure in Area B is also large by switching control points in each frequencies band. And, the sound pressure distribution in Area B also becomes uniform.

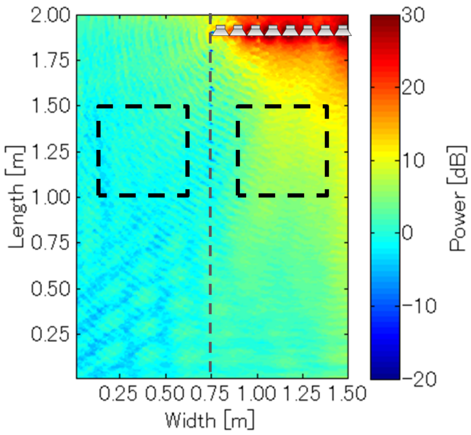
Figure 7 shows the evaluation score D in each arrangement of control points. The evaluation score D is almost the same as the unprocessed in case of Arr. I–III. On the other hand, the evaluation score D is improved by about 2 dB in case of Arr. IV. There are because the side lobe is effectively suppressed by the proposed control. In high frequency band, although many side lobes is formed, closely arranged control points can suppress undesired sound in the suppression area. However, if the number of control points is large in low frequency band, sound pressure is suppressed in excess not only suppression area but also listening area. So, by switching arrangement of the control points, the number of control points can be reduced in low frequency band that has few side lobe. Inversely, the number of control points are is needed in high frequency band which has a lot side lobes. Therefore, it is considered that we can improve evaluation score D in arr. IV by suppressing the sound pressure in the suppression area and ensuring the sound pressure in the listening area.

5. CONCLUSIONS

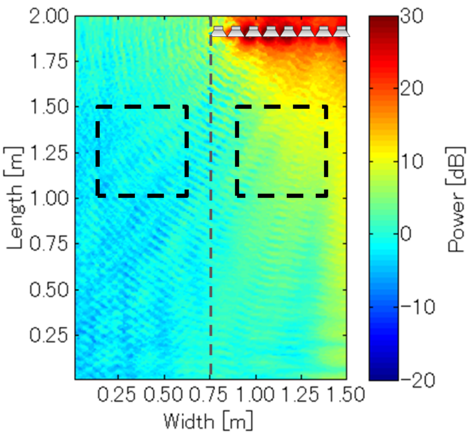
In this paper, we have proposed the NBSFC filter designing method in underdetermined condition by truncating singular value of the transfer function matrix to achieve individual sound image generation. Moreover, we evaluated the arrangement of control points, and proposed the method to switching arrangement of control points in each frequency band to improve the evaluation score D. By result of numerical simulations, the evaluation score D is almost regardless between overdetermined and underdetermined condition in case of arranging the same number control points in all frequency band. On the other hand, the evaluation score D is improved about 2 dB by switching arrangement and number of control points in each frequency band. In



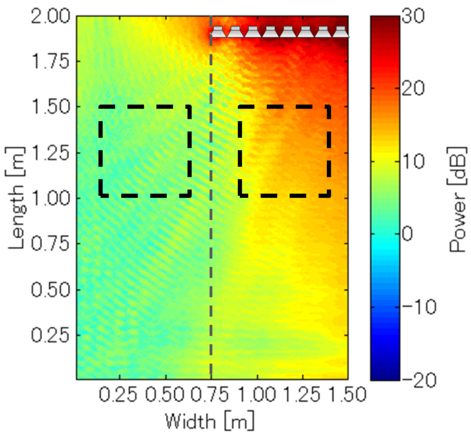
[1]Arr. I



[2]Arr. II



[3]Arr. III



[4]Arr. IV

Figure 6 – Sound power distribution by each arrangement of control points.

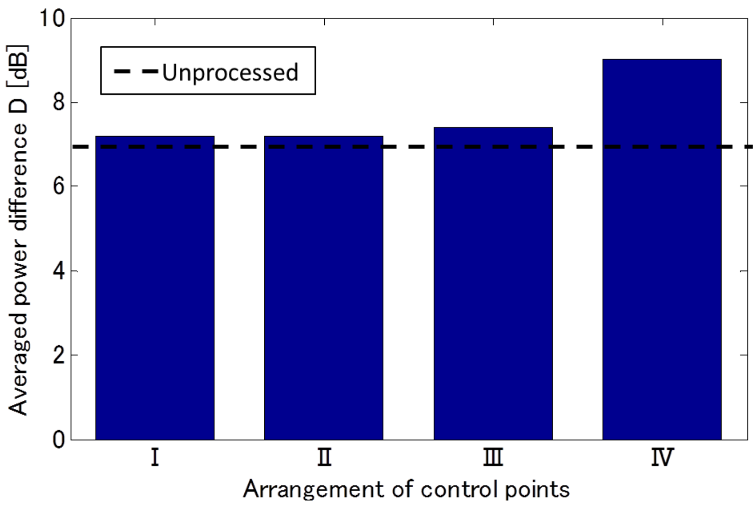


Figure 7 – Evaluation score *D* by each arrangement patterns.

future work, we try to the reproduced sound quality evaluation and improvement in listening area by modifying NBSFC filtering. Moreover, we aim to present a pleasant sound image to the listener and the extension of the listening area.

ACKNOWLEDGEMENTS

The present study is supported in part by Ono Acoustics Research Fund in Japan.

REFERENCES

1. Kamakura T, Ito M, Nomura H. Survey of ultrasound exposure: Through topic of parametric loudspeaker. J Acoust Soc Am. 2011;67(5):200-203.

2. Bauck J, Cooper DH. Head-related two–channel stereophony with loudspeaker reproduction. J Acoust Soc Am. 1996;50(4):683–705.

3. Miyabe S, Saruwatari H, Shikano K, Tatekura Y. Interface for Barge-in Free Spoken Dialogue System Using Nullspace Based Sound Field Control and Beamforming. IEICE Trans. Fundamentals, 2006;E89-A(3):716-726.

4. Hara Y, and Tatekura Y. Narrow Listening Area Formation Around One User Based on Speaker Array with NBSFC. IEICE Trans. Fundamentals, 2012; J95-A(7):645–649. (in Japanese)

5. Nakayama Y, Adachi M, Ishimoto K, Tatekura Y. Individual Sound Image Generation for Multiple Users Based on Loudspeaker Array with NBSFC. Proc 18th International Conference on Digital Signal Processing; 1–3 July 2013; Santorini, Greece 2013. DOI: 10.1109/ICDSP.2013.6622740

6. Uematsu H, Haneda Y, Kataoka A. Localization sound reproduction method that takes into consideration desired sound characteristics in a localized area. J Acoust Soc Am. 2006;62(2):89–97.

7. Wu YJ, Thushara D. Spatial Multizone Soundfield Reproduction: Theory and Design. IEEE Transactions on audio, speech, and language processing; August 2011;19(6):1711–1720.

8. Allen JB, Berkley DA. Image method for efficiently simulating small-room acoustics. J Acoust Soc Am. 1979;65(4):943–950.