PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Determination of Optimal Parameters Using Metaheuristics for the Sound Zone Generation by the Least-Squares

Kazuya Yasueda[1], Daisuke Shinjo[1], Akitoshi Kataoka[1]

[1]Ryukoku University, Japan, t16d001@mail.ryukoku.ac.jp

Abstract

We propose a method to determine the optimal parameters by metaheuristics for sound zone generation using the least-square method. We achieved multiple sound zone reproduction with sound pressure differences in multiple areas using a loudspeaker array by controlling the sound field with an inverse filter based on transfer functions between each loudspeaker and control points. At that time, three parameters must be determined: the number of control points, the control point arrangement, and the regularization parameters. In the conventional method, these parameters are experimentally and empirically determined. The control points are often set with a regular form in the target control area. In this paper, we proposed determination the number of control points and their arrangement using metaheuristics, which is a solution of mathematical optimization. We studied applying such metaheuristics as the genetic algorithm to determine these parameters. We evaluated the performance with a computer simulation. As a result, we improved the performance with the relative sound pressure level, the frequency response, and the PESQ.

Keywords: sound field reproduction, loudspeaker array, multiple sound zone generation, genetic algorithm

1 INTRODUCTION

Sound reproduction in a limited zone is a very useful technology for guidance systems at museums and automatic guidance systems at train stations. Moreover, the technology of multi-zone reproduction is expected in applications, such as multi-lingual information and digital signage. Some methods that achieve sound reproduction in a limited zone have been proposed by inverse filters [1, 2]. These methods directly control the sound pressure at control points.

Multi-zone reproduction has been proposed with above methods [3, 4]. However, we must determine loudspeaker and control point arrangement it. The arrangement of the control points is critical. In the conventional method, these parameters are experimentally and empirically determined [5]. The control points are often set in a regular form in the target control area. In this paper, we propose a method that uses such metaheuristics as the genetic algorithm to determine the arrangement of control points and their contribution.

2 MULTIPLE SOUND ZONE GENERATION BY LEAST-SQUARES METHOD

2.1 Direction control by inverse filter

Inverse filter $\mathbf{W}$ is calculated from a transfer function between each control point and a loudspeaker:

$$\mathbf{W} = \mathbf{G}^{-1}\mathbf{D}, \tag{1}$$

where $\mathbf{G}$ is the matrix of the transfer functions between control position n and loudspeakers m and

$$\mathbf{G} = \begin{bmatrix} G_{11}(\omega) & \cdots & G_{M1}(\omega) \\ \vdots & \ddots & \vdots \\ G_{1N}(\omega) & \cdots & G_{MN}(\omega) \end{bmatrix}, \tag{2}$$

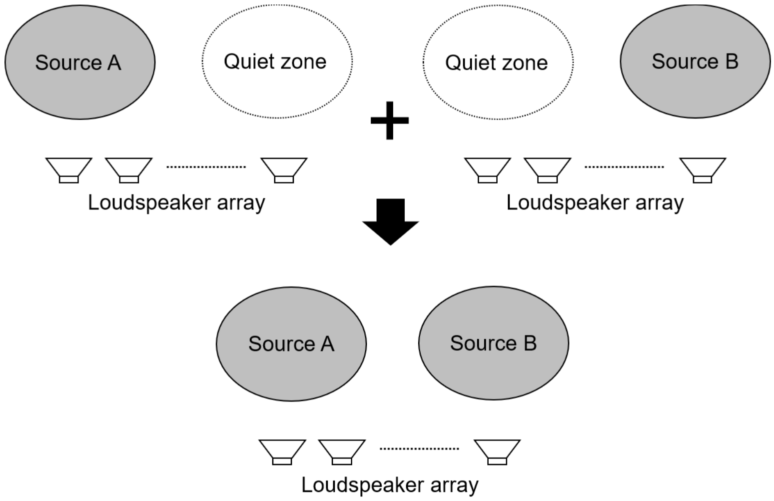


Figure 1. Concept of multiple sound zone generation using linear loudspeaker array

where ω is a frequency index. $\mathbf{D}$ is the vector of the sound pressure characteristic at control point n:

$$\mathbf{D} = \left[d_1(\omega), \quad d_2(\omega), \quad \cdots, \quad d_N(\omega) \right]^{\mathrm{T}}. \tag{3}$$

For generating reproduction and suppression zones, $d_n(\omega)$ is set to 1 or 0 respectively:

$$d_n(\omega) = \begin{cases} 1 & \text{reproduction point} \\ 0 & \text{suppression point} \end{cases}. \tag{4}$$

We deal with fewer loudspeakers than control points. Since this problem is ill-conditioned, we use the least-squares Eq. (5) for calculating the inverse filters [6]:

$$\mathbf{W} = (\mathbf{G}^{\mathrm{H}}\mathbf{G} + \delta(\omega)\,\mathbf{I})^{-1}\mathbf{G}^{\mathrm{H}}\mathbf{D}, \tag{5}$$

where $(\cdot)^{\mathrm{H}}$ is a complex conjugate transpose, $\delta(\omega)$ is a regularization parameter, and $\mathbf{I}$ is a unit matrix.

2.2 Weighted least-squares

Obtaining sufficient performance with few control points is important. The weighted least-squares method is very effective to improve the performance. The filter coefficient using it is expressed in Eq. (6):

$$\mathbf{W} = (\mathbf{G}^{\mathrm{H}}\mathbf{A}\mathbf{G} + \delta(\omega)\,\mathbf{I})^{-1}\mathbf{G}^{\mathrm{H}}\mathbf{A}\mathbf{D}, \tag{6}$$

where $\mathbf{A}$ is a weighting coefficient that is expressed as the next diagonal matrix,

$$\mathbf{A} = \mathrm{diag}\left[a_1(\omega), \quad a_2(\omega), \quad \cdots, \quad a_N(\omega) \right]. \tag{7}$$

The directional characteristic is controlled efficiently by weighting each control point. However, determining the weighting coefficients is difficult. In the conventional method, these coefficients are sometimes determined experimentally. We treat this coefficient as a parameter just like the arrangement of the control points.

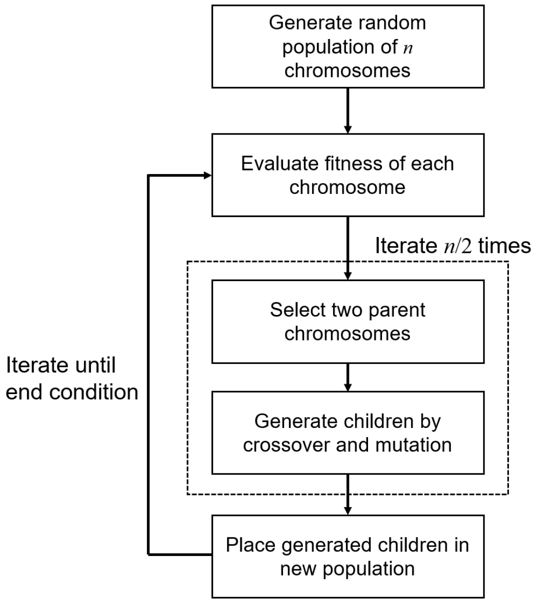


Figure 2. Process of determining arrangement of control points using GA

2.3 Multiple sound zone generation

Figure 1 shows the concept of multiple sound zone generation using a loudspeaker array. Its upper left reproduces the sound to the left side and suppresses it at the right side. On the other hand, the upper right of Figure 1 reproduces the sound to right side and suppresses it at the left side. By adding these filters that reproduces each area, we can reproduce sound to multiple zones. In this paper, we reproduce the sound to the left side in the area. Moreover, to simplify the problem, we define only 1 reproduction point: $D_n = 1$ in the reproduction zone. We discuss the suppression control points by their number, their arrangement, and their weighting coefficients.

3 DETERMINATION OF CONTROL POINTS BY GENETIC ALGORITHM

The genetic algorithm (GA) is one metaheuristic technique that calculates optimal solutions. GA was inspired by the process of natural selection, including crossover and mutation. A simple model in the GA uses a roulette wheel selection based on fitness [7]. In this paper, each control point and weighting coefficient are imaged as genes and these sets are imaged as chromosomes. By evolving a chromosome, we get an optimal arrangement of the control points and the weighting coefficients for sound reproduction. Figure 2 shows the GA process. First, 50 chromosomes are generated randomly. Second, we calculate the fitness of each one. Third, we select some good individuals and by crossover and mutation, the next generation is created. By iterating these operations, we can optimize the arrangement of the control points and the weighting coefficients.

The cost function that evaluates the sound reproduction is shown in Eq. (8), which expresses the sound pressure level ratio of the reproduction and quiet zones. The cost function is the fitness to use in the GA. $\sum P_j$ is a summation of the sound pressure level in the reproduction zone. $\sum P_k$ is a summation of the sound pressure level in the quiet zone. When the fitness is high, the performance of multiple zone sound reproduction is good:

$$F = 10\log_{10}\frac{\sum_{\in j} P_j}{\sum_{\in k} P_k}, \tag{8}$$

where j is the measurement points in the reproduction zone, k is the measurement points in the quiet zone, and

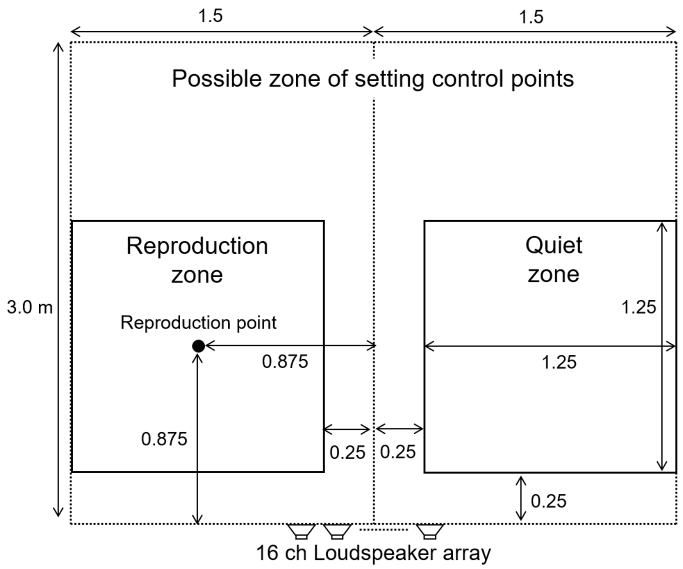


Figure 3. Position of each zone and loudspeaker array

P is the sound pressure level at each measurement point.

We used the roulette wheel selection to select the chromosomes and did a uniform crossover. This crossover create a bit array randomly, that length is the number of control points. Each bit is chosen from either parent with equal probability. The first child selects the position of the control point from first parent if a bit is 0 and does so from second parent if a bit is 1. The second child use inversion this bit array. The mutaion changes position of one control point randomly. This operation is at one time per a mutation. The crossover ratio was 0.6 and the mutation ratio was 0.01. The end condition is that no replacement continuously exists on the maximum of fitness 15 times. The weighting coefficient range is from one to ten.

4 EVALUATION BY COMPUTER SIMULATIONS

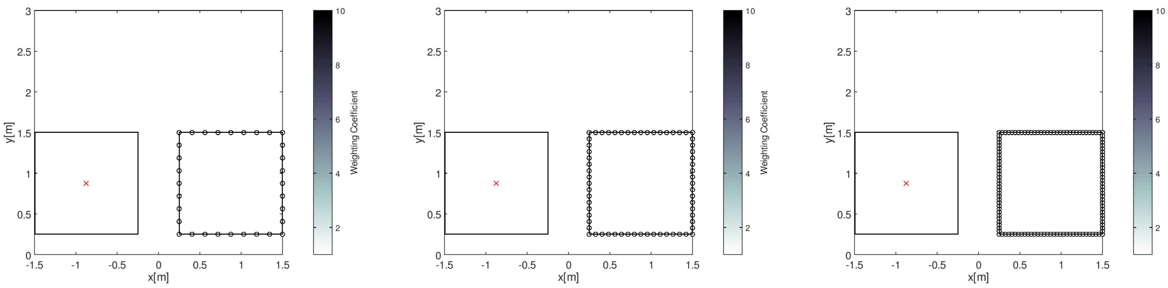
4.1 Experimental conditions

In this computer simulation, we assumed a free-field sound field. The free-field Green function to be the transfer function is given by Eq. (9):

$$G_{mn}(\omega) = \frac{1}{4\pi r_{mn}} \exp(-j\omega \frac{r_{mn}}{c}), \tag{9}$$

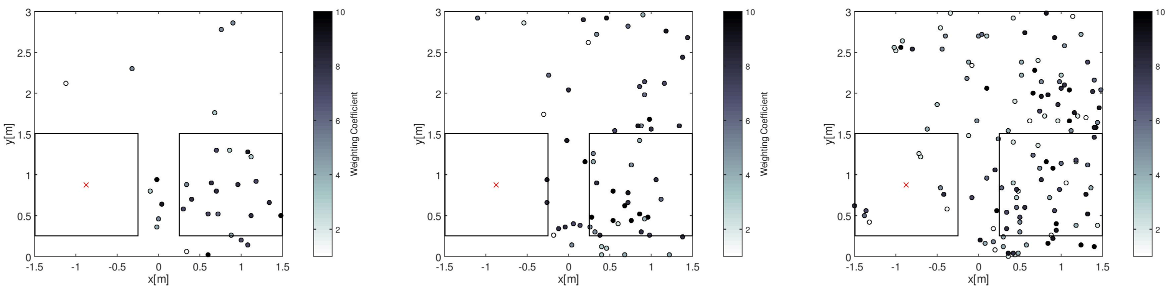
where r_{mn} is the distance from loudspeaker m to control point n and c is the speed of sound.

Figure 3 shows the position of the reproduction zone, the quiet zone and the possible zone for setting the control points. The reproduction and quiet zone regions are square of 1.25 m $\times$ 1.25 m. The control points are set to 3 m $\times$ 3 m in front of the loudspeaker array. Reproduction point $D_n = 1$ is set to $x = -0.875$ and $y = 0.875$. Measurement point j, k of Eq. (8), are set overall in each zone with lattice-like discrete point. The interval of the measurement point is set 0.25 m. Therefore, each zone has 36 measurement points. Figure 4 shows the conventional arrangement of the control points. $\times$ is the reproduction control point and $\circ$ is the suppression control point in this figure. We used the arrangement of the suppression control point setting on the boundary of the quiet zone as a conventional arrangement, and the all weighting coefficients are set to one.



(a) 32 points (b) 64 points (c) 128 points

Figure 4. Conventional arrangement of control points



(a) 32 points (b) 64 points (c) 128 points

Figure 5. Arrangement of control points by GA arrangement

Table 1. Mean fitness with each number of control points and mean of generations

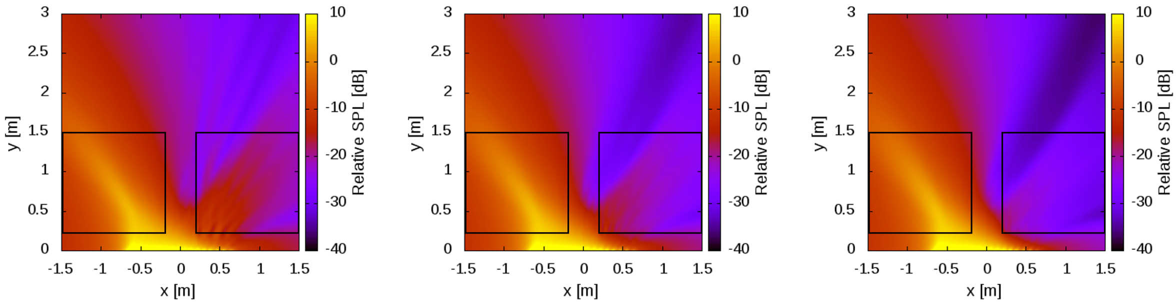
Number of control point	16	32	64	96	128	151
Mean fitness [dB]	15.89	21.32	25.99	28.83	27.28	26.75
Standard deviation of fitness [dB]	1.65	0.76	1.76	1.07	1.36	1.97
Mean number of generations	28.2	37.0	51.8	60.8	63.0	62.2

4.2 Results of experiment

Table 1 shows the mean fitness, the standard deviation, and the mean number of generations by GA. The color of a circle shows weighting coefficient. It is small if a circle is the white. These values are the means of five trials and the number of control points, including the reproduction point. The fitness tends to be high with an increase in the number of control points. However, the fitness is not so different when the number of control points exceeds 96. The variation of the fitness of each trial is small from the standard deviation. The mean number of generations become large when the number of control points is large.

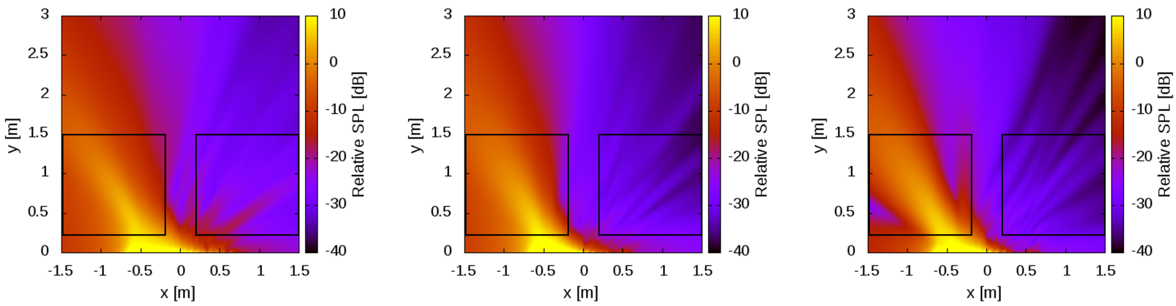
Figure 5 shows the arrangement of the control points by GA. The suppression control points are set in the quiet zone side and near the loudspeaker array. However, as shown by Figure 5(c), the control points are set into the reproduction zone when using a large number of control points. There is a tendency to set a large weighting coefficient to the control points near the loudspeaker array.

Figure 6 and 7 show the distribution maps of the relative sound pressure level with each of the control



(a) 32 points (b) 64 points (c) 128 points

Figure 6. Distribution maps of relative sound pressure level using conventional arrangement

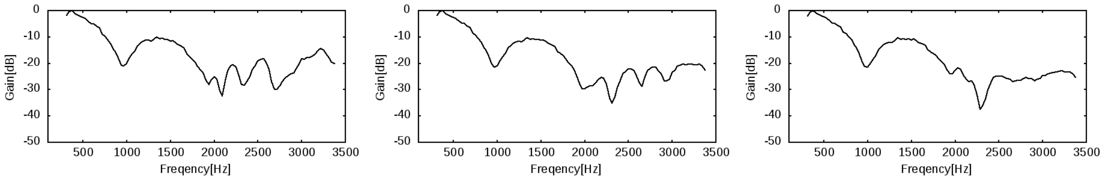


(a) 32 points (b) 64 points (c) 128 points

Figure 7. Distribution maps of relative sound pressure level of GA arrangement

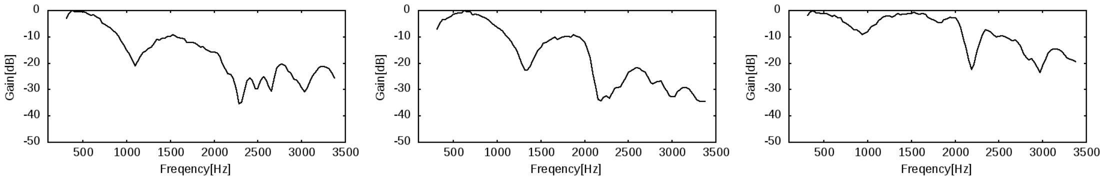
points. These figures show the best one in each of five trials. The reproduction point is a reference point (0 dB) when calculating the relative sound pressure level identically as the fitness. From Figure 7, the arrangement using GA can reproduce the sound to the reproduction zone and reduce it leakage to the quiet zone even if fewer control points are used. From Figure 7(c), when the control points are set into the reproduction zone, the directional characteristic to the reproduction point is stronger. Therefore, the beamform to the reproduction point narrows. It is not desirable if the reproduction zone is large. Moreover, the GA arrangement decreases the number of control points compared to the conventional arrangement based on experience.

Figure 8 and 9 show the frequency response at the reproduction point. The GA arrangement is compared with the conventional one, and the frequency characteristics are better especially in 128 points. However, they are slightly worse in the middle and high frequency bands. Figure 10 and 11 show the distribution maps of the Perceptual Evaluation of Speech Quality (PESQ) [8], which is an automated assessment of speech quality simulated by a user. It is standardized as ITU-T recommendation P.862. Its score is principally a model mean opinion score (MOS) that covers a scale from 1 (bad) to 5 (excellent). The score is calculated using speech signals with 300 $\sim$ 3400 Hz. These figures are calculated by the mean of two people's speech. In all the figures, the PESQ score is higher than three in the reproduction zone. Therefore, there are no complications with listening to speech.



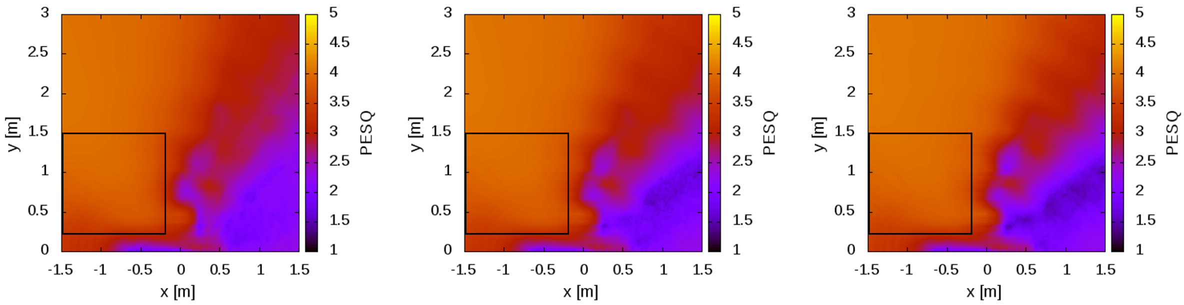
(a) 32 points (b) 64 points (c) 128 points

Figure 8. Frequency response of conventional arrangment at reproduction point



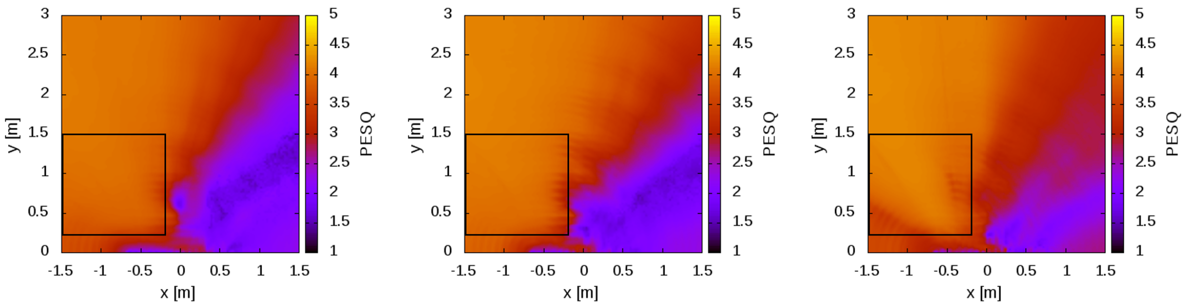
(a) 32 points (b) 64 points (c) 128 points

Figure 9. Frequency response of GA arrangment at reproduction point



(a) 32 points (b) 64 points (c) 128 points

Figure 10. Distribution maps of PESQ score using conventional arrangement



(a) 32 points (b) 64 points (c) 128 points

Figure 11. Distribution maps of PESQ score of GA arrangement

5 CONCLUSIONS

Multiple zone sound zone generation by inverse filters with a linear loudspeaker array is very useful technology. In this paper, we determined the arrangement of suppression control points using GA to optimize each parameter in the sound reproduction by treating the control points and their weighting coefficient as chromosomes. In an evaluation with a computer simulation, the arrangement by our proposed method is more useful than the conventional one. We achieved multiple zone generation by a small number of control points in the proposed method. However, the control point by GA is a halfway position. Since measuring impulse responses from loudspeaker to each control point is very difficult, we need to study that problem in the future.

REFERENCES

[1] O. Kirkeby and P. Nelson, “Reproduction of plane wave sound fields,” *J. Acoust. Soc. Am.*, vol. 94, no. 5, pp. 2992-3000, Nov. 1993.

[2] O. Kirkeby, P. Nelson, F. Orduna-Bustamante, and H. Hamada “Local sound field reproduction using digital signal processing,” *J. Acoust. Soc. Am.*, vol. 100, no. 3, pp. 1584-1593, Nov. 1996.

[3] J.-W. Choi and Y.-H. Kim, “Generation of an acoustically bright zone with an illuminated region using multiple sources,” *J. Acoust. Soc. Am.*, vol. 111, no. 4, pp. 1695-1700, Apr. 2002.

[4] M. Shin, S. Q. Lee, F. M. Fazi, P. A. Nelson, D. Kim, S. Wang, K. H. Park, and J. Seo, “Maximization of acoustic energy difference between two spaces,” *J. Acoust. Soc. Am.*,vol. 128, no. 1, pp. 121-131, July 2010.

[5] K. Yasueda, D. Shinjo, and A. Kataoka, "Arrangement of control points for sound spot generation in the multi-point control method,"*5th Joint Meeting of the ASA and ASJ*, *J. Acoust. Soc.. Am.*, vol. 140, no. 4, Pt. 2, p. 3140, 2016

[6] H. Tokuno, O. Kirkeby, P.A. Nelson, and H. Hamada, "Inverse Filter of Sound Reproduction System Using Regularization," *IEICE Trans. Fandamentals.*, E-80A, No. 5, pp. 809-820, 1997.

[7] D. E. Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning," Addison-Wesley, 1989

[8] Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs, ITU-T P.862, 2002.