Proceedings of the 14th
KACC, Octorber 1999

Realization of Point-Listening Characteristics by Enclosed Microphone Array System with Optimal Complex Weighting

SHINJI OHYAMA*, YUKIFUMI SASAGAWA*, LI CAO** AND AKIRA KOBAYASHI*

* Faculty of Engineering, Tokyo Institute of Technology, Tokyo 152-8552, JAPAN
Tel: +813-5734-2325; Fax: +813-5734-3596; E-mail: sohyama@ctrl.titech.ac.jp
** Dept. of Automation, Tsinghua University, Beijing, CHINA

Abstract An electronically scannable microphone system is in the planning stage. For this purpose, a multiple microphone array with controllable delay is available. To achieve effective point-listening characteristics, we proposed an enclosed microphone array system with a complex weighting method. In this system, both the microphone arrangement and the value of the complex weighting are important. In this report, the construction of microphone array system and the signal-processing method are explained, and the calculation method for optimal complex weighting is also presented. A prototype experimental setup is designed and fabricated to verify the expected characteristics.

Keywords microphone array system, point-listening characteristics, DSP, complex weighting

1 Introduction

Many schemes had been proposed for a system that achieves excellent point-listening sensitivity characteristics by arranging a multiple microphone array[1]. If the pinpoint listening function aiming at a target point is realized by an electronic method, an electronically scannable point-listening microphone system should be constructed. This function is useful for an individual speech-detecting system in a meeting or a conference. However, most of the past studies were concerned with the far field problem, that is to say, the distance between the sound source and the microphone array are relatively long. In such a case, the analysis is comparatively easy because a sound wave has to spread and it can be handled as a plane wave. It is far more difficult to treat the near field case. In this study, the near field problem is considered assuming a rather narrow area(e.g., the meeting room).

In our previous report, the fundamental characteristics of this system is analyzed, especially the enclosed microphone array system was discussed which optimized the pinpoint sensitivity characteristics by adjusting both the amplitude and the phase shift on the output signal from each microphone[2]. In this report, both simulation and experimental results are also shown.

2 Construction of 2-dimensional microphone array system

Assume that the area of a quadrilateral room is $\ell \times \ell$ and N microphones are placed on the 4 latera of the ceiling and that there is a person at the position of the sound source. **Fig. 1** shows this situation. The distance between the sound source and the microphone plane is h.

The purpose of this system is to hear the voice of a specific person clearly; when the position of the person is known as (x_f, y_f), the sensitivity is adjusted so that it becomes maximum at the point (x_f, y_f).

Fig. 1 also shows the scheme of the signal processing. The operation for signals from each microphone involves both a multiplication W_i and a time delay D_i; that is, the complex weighting is multiplied. A single frequency sound(ω) derived from position (x, y) is radiated to each microphone, and the output signal of a microphone-i: $m_i(t)$ after operation of the complex weighting is expressed as:

$$m_i(t) = \frac{W_i}{L_i(x,y)} e^{-j(2\pi \frac{L_i(x,y)}{\lambda} + \omega D_i)} \cdot e^{j\omega t} \tag{1}$$

where

$$L_i(x,y) = \sqrt{(x-x_i)^2 + (y-y_i)^2 + h^2}, \tag{2}$$

and where λ is the wavelength of the sound, (x_i, y_i, h) is the position of the i-th microphone $(i = 1, 2, \cdots, N)$. The summation of these signals $S(t)$ is expressed as:

$$\begin{aligned} S(t) &= \sum_{i=1}^{N} m_i(t) \\ &= \sum_{i=1}^{N} \frac{W_i}{L_i(x,y)} e^{-j(2\pi \frac{L_i(x,y)}{\lambda} + \phi_i)} \cdot e^{j\omega t} \\ &= S_p(x,y) \cdot e^{j\omega t} \end{aligned} \tag{3}$$

In the original meaning, D_i denotes the time delay term of the complex weighting, but ϕ_i will be used for the sake of notation.

$$\phi_i = \omega D_i \tag{4}$$

In the following description, (W_i, ϕ_i) is the complex weighting.

Note that eq. (3) is composed of both a time-invariant term $S_p(x,y)$ and a time-variant term $\exp(j\omega t)$. It is

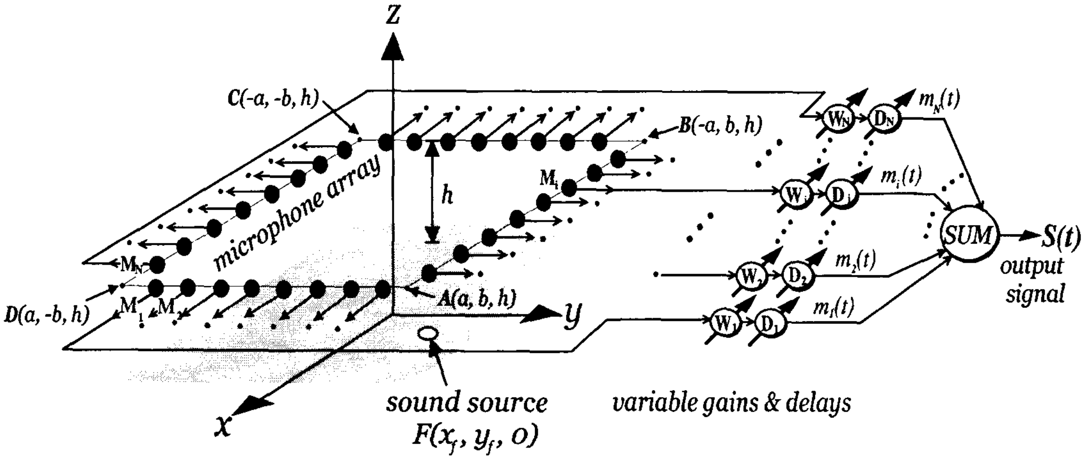


Fig. 1: Schematic construction of the enclosed microphone array system.

then important that the $S_p(x,y)$ term describes the amplitude characteristics. Calculate the square of $S_p(x,y)$ to consider the power amplitude characteristics.

$$\begin{aligned}|S_p(x,y)|^2 &= \Big(\sum_{i=1}^{N}\frac{W_i}{L_i(x,y)}\cos\big(2\pi\frac{L_i(x,y)}{\lambda}+\phi_i\big)\Big)^2\\ &+\Big(\sum_{i=1}^{N}\frac{W_i}{L_i(x,y)}\sin\big(2\pi\frac{L_i(x,y)}{\lambda}+\phi_i\big)\Big)^2\\ &=\sum_{i=1}^{N}\sum_{j=i}^{N}S_{P_{ij}}(x,y) \qquad (5)\end{aligned}$$

In eq.(5), the term $S_{pij}(x,y)$ is described as follows depending on the combination (i,j), that is:

- $i=j$

$$S_{P_{ij}}(x,y)=\left(\frac{W_i}{L_i(x,y)}\right)^2 \qquad (6)$$

- $i\neq j$

$$S_{P_{ij}}(x,y)=\frac{2W_iW_j}{L_i(x,y)L_j(x,y)}\cos\psi_{ij}(x,y) \qquad (7)$$

where

$$\psi_{ij}(x,y)=\frac{2\pi}{\lambda}\Big(L_i(x,y)-L_j(x,y)\Big)+(\phi_i-\phi_j) \qquad (8)$$

The next step to be considered is to find the combination (W_i,ϕ_i) so that the $|S_p(x,y)|^2$ is maximum around the focusing point (x_f,y_f) and smaller otherwise. The steepest descent method is applied to solve this optimization problem[2], and in the optimization process, the following evaluation function $I(x_f,y_f)$ is used.

$$I(x_f,y_f)=\frac{\int_{-b'}^{b'}\int_{-a'}^{a'}|S_p(x,y)|^2dxdy}{4a'b'|S_p(x_f,y_f)|^2} \qquad (9)$$

In eq.(9), a' and b' denote the consideration area for the sensitivity characteristics. The property of the function $I(x_f,y_f)$ is that the amplitude is the largest around the focusing point and least at the other points. The optimum complex weighting (W_i,ϕ_i) is given when $I(x_f,y_f)$ is minimized.

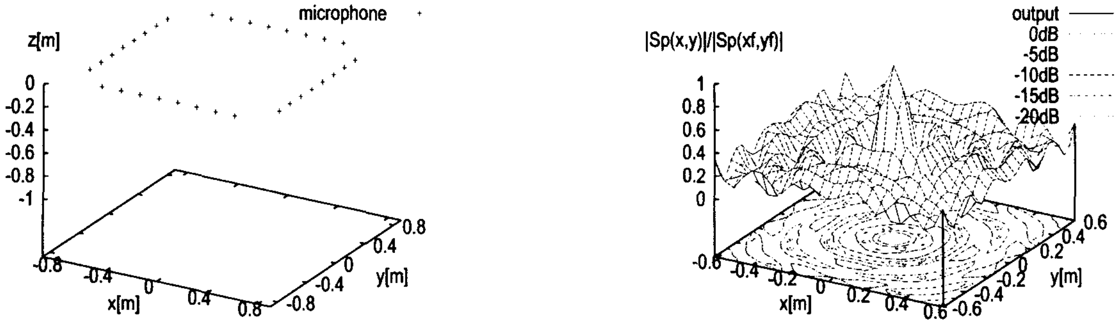
3 Listening sensitivity distribution characteristics by computer simulation

The listening sensitivity distribution characteristics are simulated using the microphone arrangement shown in **Fig. 2(a)**. First, the target point is set to $(0,0)$, and a set of the optimal complex weighting (W_i,ϕ_i) is then calculated using the steepest descent method based on the evaluation function $I(0,0)$. Next, the listening sensitivity distribution characteristics are calculated using this optimal complex weighting.

The simulated results are shown in **Fig. 2(b)**. The sound frequency is 1 kHz. The amplitude of the target point is the highest and lower on the other points. Some side-lobes like concentric circles can be observed, but the central peak is the highest. This result demonstrates that the point-listening characteristics are performed by the enclosed microphone array and proper complex weighting operation.

Table 1: Specifications of experimental system

number of microphones	N = 32
field constant	$a=b=0.75$[m]
measuring area	$a'=b'=0.6$[m]
sound frequency	$f=400\sim 4{,}000$[Hz]
quantize bit	12-bit
sampling frequency	40 [kHz]
complex weighting method	DSP(ADSP-21061)
data flow rate	10 [kHz]
resolution of W_i data	12-bit
resolution of ϕ_i data	25 [μs]



(a) Microphone arrangement

(b) Listening sensitivity distribution

Fig. 2: Microphone arrangement and listening sensitivity distribution in the case of an enclosed array (N=32)

4 Experiments

4.1 Experimental setup

A prototype measurement system was designed and fabricated for the purpose of this study. **Table 1** and **Fig. 3** show the specifications and the block diagram, respectively. The most important point is that the complex weightings $(W_i, \phi_i)(i = 1, \cdots, N)$ are operated on every microphone signal in real time and that these are flexibly changeable according to change in target points. This function is performed by a DSP(digital signal processor: Analog Devices, Inc. ADSP-21061) and by its software derived from PC.

To confirm the proposed theory, a 32-microphone array system is adopted[1]. Every analog signal from the microphones is converted to a digital signal with a 12-bit resolution. Thirty-two serial signals are converted to parallel data by FPGA(field-programmable gate array) and fed to a DSP where the complex weightings are operated. The output signal from the signal-processing unit is observed and recorded by a digital oscilloscope.

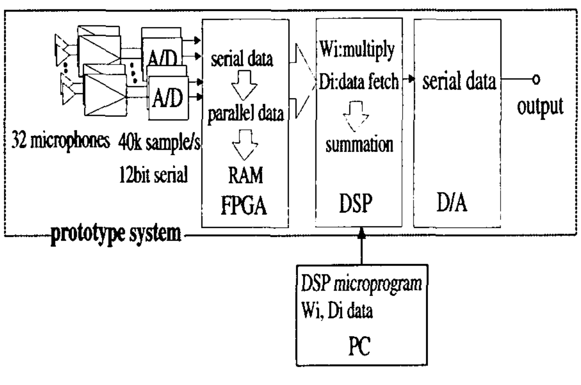
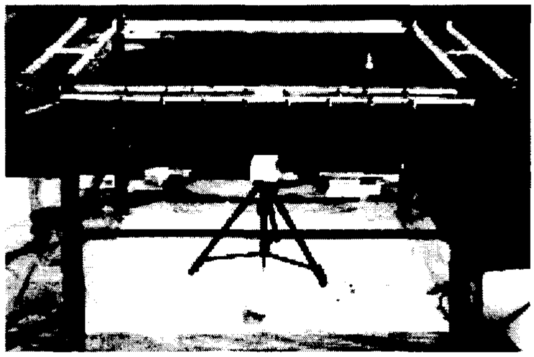


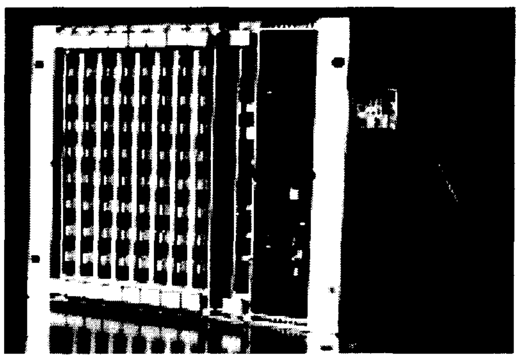
Fig. 3: Outline of signal processing system

[1]The whole system has 64 microphones, but only 32 microphones were used for this study.

Fig. 4 presents the photograph of the prototype experimental system; (a) shows the arrangement of the microphone array and (b) is the signal processing unit(9 boards of electric circuit). Thirty-two microphones and pre-amplifiers are placed on the four-latus; there is a speaker under the microphone array, and the sine wave is radiated from this. The coordinate of the speaker is (0,0), and the distance between the microphone array plane and the speaker is 0.36[m].



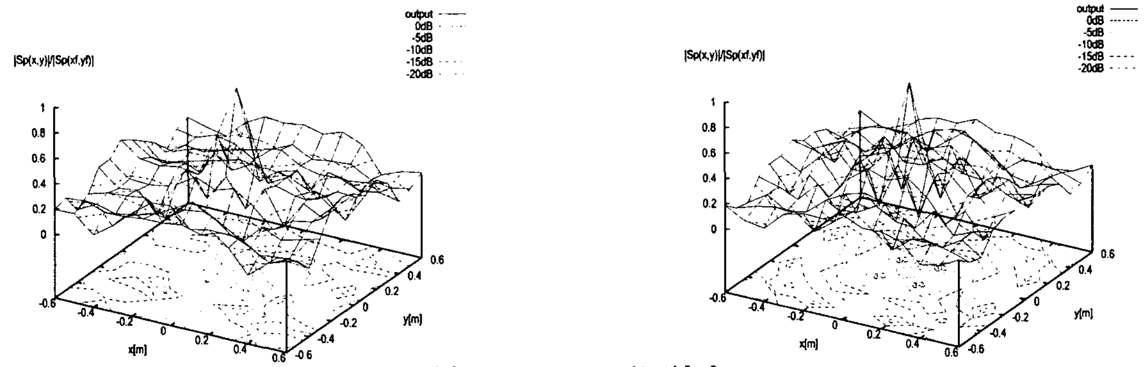
(a) Arrangement of microphone array and speaker.



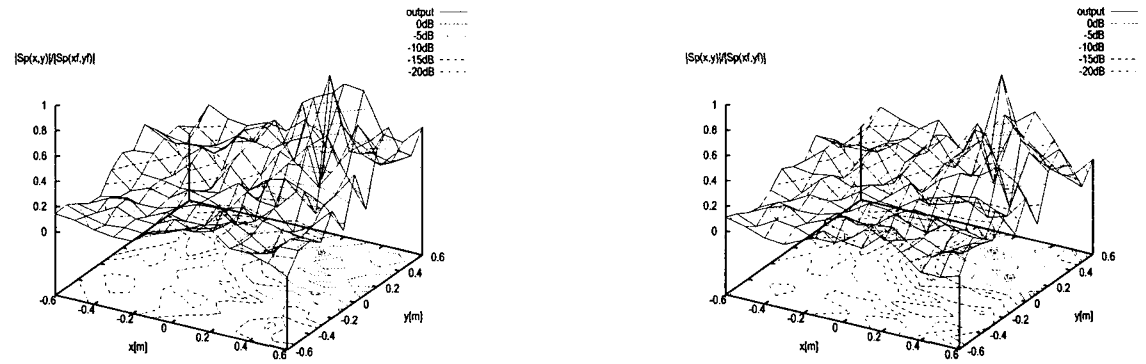
(b) Signal processing unit

Fig. 4: Photograph of prototype experimental system.

experimental | theoretical



(a) Target point: $(0, 0)$[m]



(b) Target point: $(0.3, 0.3)$[m]

Fig. 5: Experimental results of listening-sensitivity distribution compared to theoretical characteristics.

4.2 Experimental results

The experiment was carried out using the same microphone arrangement as Fig. 2(a) and the optimized complex weighting (W_i, ϕ_i) calculated for 1 kHz sound in section 3. Experimental results are shown in **Fig. 5** where the case of the target point $(0, 0)$[m] is (a), (0.3, 0.3)[m] is (b). Graphs on the left side are the experimental results; those on the right side are theoretical. The pair of graphs shows good accordance with respect to the target points. Note that the maximum amplitude point is displaced with the change in the target point and that the shapes of both contour lines are similar.

5 Conclusions

The microphone array system with both the enclosed microphone arrangement and the complex weighting was proposed for pinpoint-listening sensitivity characteristics and its validity examined by composing a complex weighting with a 32-microphone array. Experiments on the electronically scannable function are the subject of future study.

Acknowledgments

The authors wish to express their gratitude to the members of OKK, Inc., for support of the experimental system. This work was supported by a Grant-in-Aid for COE Research (Super Mechano-Systems) from the Ministry of Education, Science, Sports and Culture of Japan.

References

[1] C. L. Dolph: A current distribution for broadside arrays which optimizes the relationship between beam width and side-lobe level, Proc. of the I.R.E. and Waves and Electrons, **34**-6, 1946, pp.335–348

[2] S.Ohyama Y.Sasagawa, K.Takahashi and A.Kobayashi Electronically scannable pinpoint listening microphone array system, Proceedings of TITech COE/Super Mechano-Systems workshop '99, 1999 (in Tokyo), pp.179–184