Novel Application of Spherical Microphone Array Sensor with Three-Dimensional Directivity

KIKIWAKE 3D: Language Game for Children

Takahiro Nakadai, Tomoki Taguchi, Hiroshi Mizoguchi
Department of Mechanical Engineering
Tokyo University of Science
2641 Yamazaki, Noda-shi, Chiba, Japan
7514633@ed.tus.ac.jp

Ryohei Egusa, Etsuji Yamaguchi, Shigenori Inagaki, Yoshiaki Takeda
Graduate School of Human Development and Environment
Kobe University
Yayoigaoka 6, Sanda-shi, Hyogo, Japan
126d103d@stu.kobe-u.ac.jp

Miki Namatame
Faculty of Industrial Technology
Tsukuba University of Technology
4-3-15, Amakubo, Tsukuba, Ibaraki, Japan
miki@a.tsukuba-tech.ac.jp

Masanori Sugimoto
Graduate School of Information Science and Technology
Hokkaido University
Kita 15, Nishi 8, Kita-ku, Sapporo, Hokkaido, Japan
sugi@ist.hokudai.ac.jp

Fusako Kusunoki
Department of Information Design
Tama Art University
2-1723, Yarimizu, Hachioji, Tokyo, Japan
kusunoki@tamabi.ac.jp

Abstract—**Mixed sounds can be separated from multiple sound sources by utilizing a microphone array and signal processing. We believe that promotion of interest in this technique can lead to significant developments in future science and technology. Consequently, in order to experience this technique, we designed a language game for children called "KIKIWAKE 3D" that uses a sound source separation system with the aim of arousing children's interest in this technology. In this paper, we discuss the development of a spherical microphone array with three-dimensional directivity introduced to make the game more interesting and to facilitate a larger number of participants.**

Keywords—sound source separation; signal processing; voice recognition

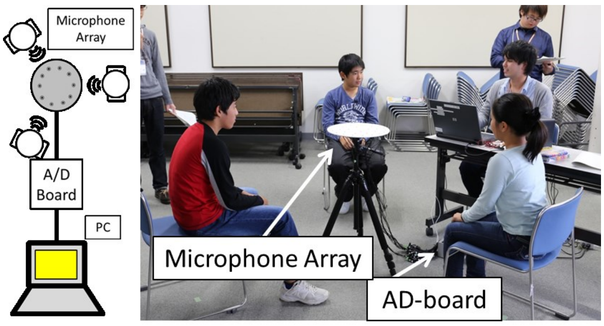


Figure 1. Participatory language game arrangement

I. INTRODUCTION

In this paper, we report on the development of a spherical microphone array sensor for a language game. Microphone array signal processing is used to separate an objective sound from mixed sounds. This technology is currently being actively researched for application in areas such as sound source identification apparatus and the environment recognition function of service robots [1].

The novel technology discussed in this paper was developed to spur interest in future science and technology. Thus, it is hoped that children will become interested in this technology. In previous research, a participatory language game was introduced to promote children's interest in the technology [2]. However, only three players were able to participate in that game because the microphone array used had only two-dimensional directivity control. We have since improved our design to a microphone array with three-dimensional directivity that enables as many as six players to participate.

We designed a game called "KIKIWAKE 3D" using language play in order to provide an opportunity to get children interested in this technology. This new design helps to promote greater interest by increasing the number of players in the game. We developed a spherical microphone array sensor. In this paper, we discuss the development of the sensor and examine its efficacy by means of a comparative experiment.

II. Sound Source Separation System

A. Optimal Microphone Array of KIKIWAKE 3D

There are several microphone array signal processing methods. In our research, we selected the Delay-and-Sum Beam Forming (DSBF) [3, 4] method because it is robust in real environments. The DSBF method can be used to capture a sound selectively and locally to form a high-sensitivity beam in an objective direction. In this method, the microphone arrangement significantly affects the shape of the beam. Therefore, we designed a microphone array with special focus on mainlobe and sidelobe.

Fig. 2 shows good and bad examples of beam-forming. If the mainlobe width is thick and the sidelobe gain is high, the noise source easily covers the high-sensitivity area. To effectively capture an intended sound, it is better to form a beam that is as narrow as the mainlobe width and minimize the sidelobe gain.

Calculating analytically the optimal microphone arrangement that is able to narrow the mainlobe and minimize the sidelobe is difficult. Therefore, we derived a number of candidate microphone arrangements, and evaluated the performance of each candidate to obtain an optimal microphone arrangement [5].

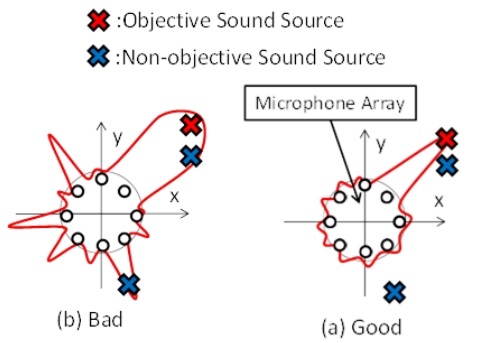
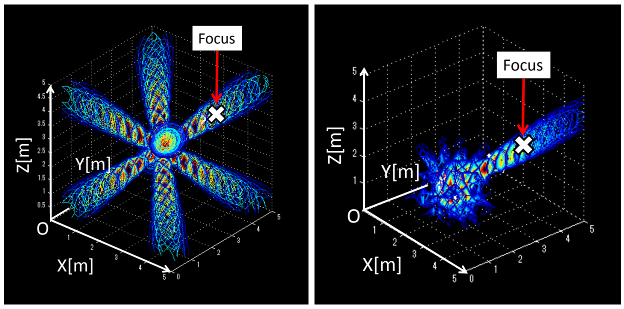


Figure 2. Shape of the beam formed by a microphone array

B. Realization of Three-Dimensional Directivity

The microphone array developed was designed by beam-forming simulation. Fig. 3(a) shows a sensitivity distribution map of cube isodensity placement microphone array with three-dimensional directivity. The white mark is called the "focus." The origin of the coordinate is the center of the cube isodensity placement microphone array. The distance between each microphone is 0.10 m. The number of microphone elements is 64. Fig. 3(b) shows a sensitivity distribution map of a spherical microphone array with three-dimensional directivity. The origin of the coordinate is the center of the spherical microphone array. The diameter of the spherical microphone array is 0.40 m. The number of microphone elements is 60.

We designed the optimal spherical microphone array depicted in Fig. 4 using this kind of simulation [6].



(a) (b)

Figure 3. Sensitivity distribution map

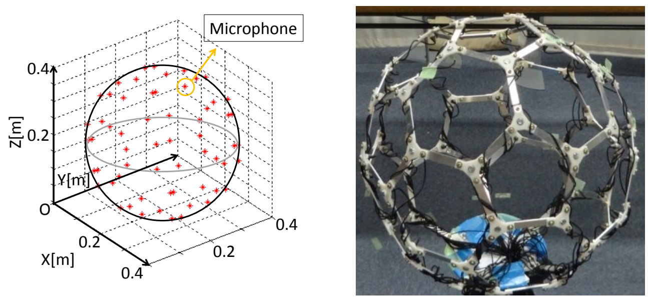


Figure 4. Spherical microphone array

III. Directivity Evaluation Experiment

A. Spherical Microphone Array

We conducted a directivity evaluation experiment using the spherical microphone array implemented in our research. Fig. 5 shows the experimental environment. We used two speakers in the experiment. The two speakers output sounds with different frequencies at the same time, which are absorbed by the spherical microphone array. By using the microphone array developed, non-objective sound is reduced and objective sound is amplified.

Speaker 1 output a sine wave of 1500 Hz, and Speaker 2 output a sine wave of 1000 Hz. In order to examine the focus formed for an objective speaker, we measured the relative sound pressure level at each speaker. On the right of Fig. 5, the center of the spherical microphone array is the origin (0, 0, 0). Speaker 1 is installed at (1000, 0, 700). Speaker 2 is installed at (1000, 0, -700). This state is called Situation 1 in this paper. Further, we configured Situations 2 and 3 with the parameters listed in Table 1.

Fig. 6 shows the power of the frequency component of the measured sound signals for each speaker. We conducted this experiment in Situations 2 and 3. In this experiment, we compared three results. First, no focus was formed by beam-forming. Second, a focus was formed for Speaker 1. Finally, a focus was formed for Speaker 2. Table 2 shows the power of the frequency component for each result. In the case where the focus was not formed, the power spectrum for 1500 Hz is approximately equal to that for 1000 Hz. In the case where the focus was formed for Speaker 1, the power spectrum for 1500 Hz is larger than that for 1000 Hz by 30 dB. In the case where the focus was formed for Speaker 2, the power spectrum for 1000 Hz is larger than that for 1500 Hz by 13 dB. These results confirm the efficacy of the spherical microphone array with three-dimensional directivity.

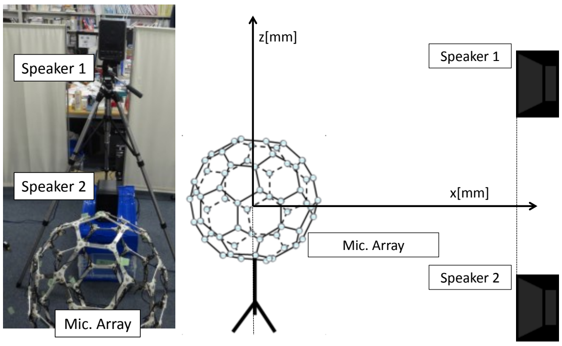
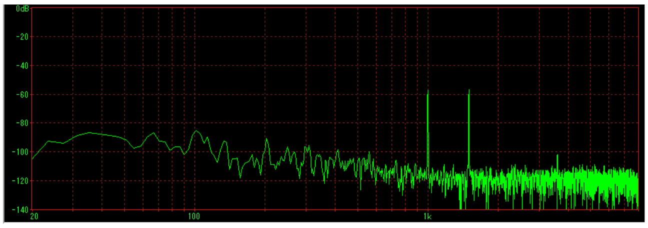
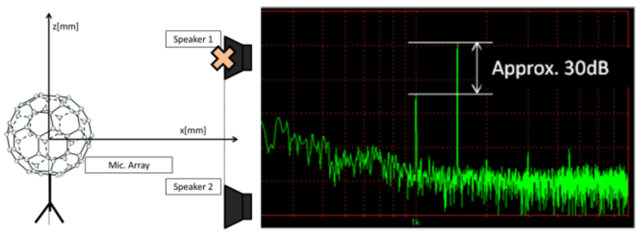


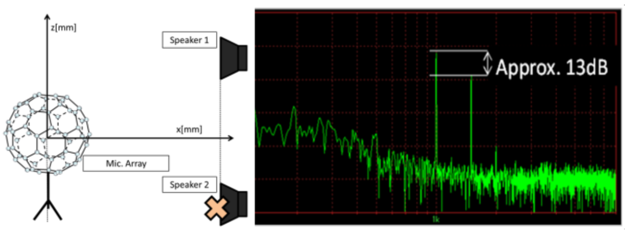
Figure 5. Experimental setup of spherical microphone array



(a)



(b)



(c)

Figure 6. Directivity evaluation experiment results

TABLE I. COORDINATES OF EACH SITUATION

	Speaker 1	Speaker 2
Situation 1	(1000, 0, 700)	(1000, 0, -700)
Situation 2	(1000, 0, 500)	(1000, 0, -500)
Situation 3	(1000, 0, 300)	(1000, 0, -300)

TABLE II. POWER OF FREQUENCY COMPONENT

	SPL of 1500 Hz (dB)	SPL of 1000 Hz (dB)
Before DSBF	-57.03	-56.84
Focus Speaker 1	-48.85	-18.82
Focus Speaker 2	-24.30	-35.54

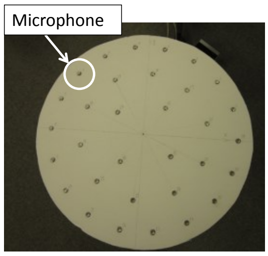


Figure 7. Circular microphone array

B. *Comparison Experiment*

In order to compare directivity, we conducted the same experiment using the circular microphone array. Fig. 7 shows the circular microphone array used in the experiment. The diameter of the circular microphone array is 0.3 m. The number of microphone elements is 32. Fig. 8 depicts the layout of the experiment. We conducted the experiment in Situations 1, 2, and 3.

Fig. 9 shows the difference in power spectrum of each frequency component versus situation number. The vertical axis shows power spectrum difference. In all the situations, the spherical microphone array sensor has the larger power spectrum difference, confirming that it is more effective than the circular microphone array.

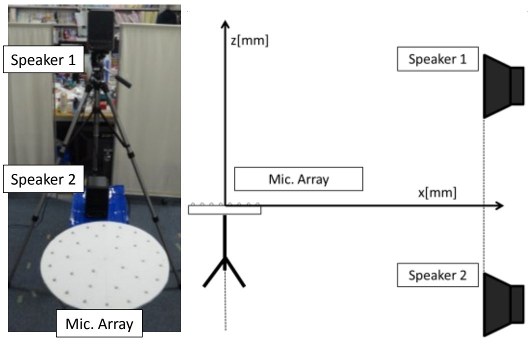


Figure 8. Experimental setup of circular microphone array

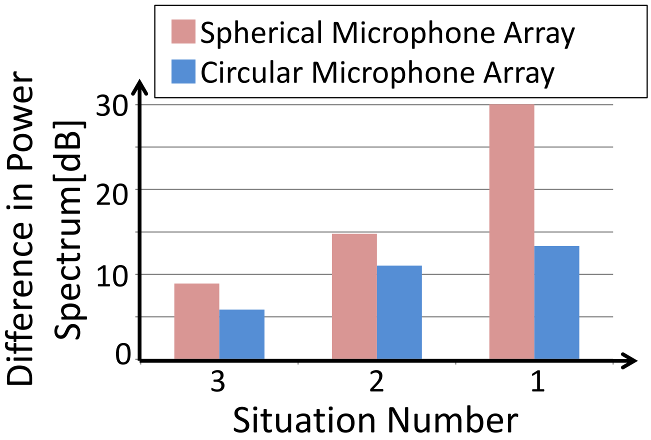


Figure 9. Comparison experiment results

IV. CONCLUSION

In this paper, we reported on the development of a spherical microphone array sensor to use in a language game. In our study, in an effort to get elementary school children interested in sound source separation technology, which is considered very important for the future development of science and technology, we designed a language game called “KIKIWAKE 3D” through which the children can experience the sound source separation technology using their own voice.

We developed the microphone array with three-dimensional directivity in order to make the game more interesting and to increase the number of participants. In the game, it is essential that the same directivity be formed in every direction in the space. Therefore, a spherical microphone array was decided on and developed for the game by simulation. It was designed as an optimal microphone array arrangement that is able to capture sound effectively. The results of the directivity evaluation experiment conducted indicate that the spherical microphone array sensor is valid and effective for the game.

In the future, we plan to let children play using KIKIWAKE 3D and investigate how they feel about the technology.

ACKNOWLEDGMENT

This work was supported in part by Grants-in-Aid for Scientific Research (B) (24300290 and 23300303).

REFERENCES

[1] M. Goseki, M. Ding, H. Takemura, and H. Mizoguchi, “Combination of microphone array and camera image processing for visualizing sound pressure distribution.” SMC2011, pp.139-143, 2011.

[2] T. Taguchi, M. Goseki, R. Egusa, M. Namatame, F. Kusunoki, M. Sugimoto, E. Yamaguchi, S. Inagaki, Y. Takeda, and H. Mizoguchi, “KIKIWAKE: Sound source separation system for children-computer interaction.” CHI 2013 Extended Abstracts, pp. 757-762, April 27–May 2, 2013.

[3] N.-V. Vu, H. Ye, J. Whittington, J. Devlin, and M. Mason, “Small footprint implementation of dual-microphone delay-and-sum beamforming for in-car speech enhancement.” International Conference on Acoustics, Speech, and Signal Processing. USA, pp. 1482-1485, March 2010.

[4] M. Fuchs, T. Haulick, and G. Schmidt, “Noise suppression for automotive applications based on directional information.” International Conference on Acoustics, Speech, and Signal Processing. Canada, vol. 1, pp. I – 237-40, May 2004.

[5] T. Fujihara, S. Kagami, Y. Sasaki, and H. Mizoguchi, “Arrangement optimization for narrow directivity and high S/N ratio beam forming microphone array.” IEEE SENSORS, Italia, pp. 450–453, October 2008.

[6] T. Taguchi, T. Nakadai, R. Egusa, M. Namatame, F. Kusunoki, M. Sugimoto, E. Yamaguchi, S. Inagaki, Y. Takeda, and H. Mizoguchi, “Investigation on optimal microphone arrangement of spherecal microphone array to achieve shape beamforming.” Intelligent Systems, Modelling and Simulation, pp. 330-333, 27–29 January, 2014.