Proceedings of 2008 IEEE International Conference on
Mechatronics and Automation

WD1-4

Noise reduction combining microphones and laser listening devices

Mitsuharu Matsumoto *Member, IEEE*, Tomomi Abe *Non-member, IEEE* and Shuji Hashimoto *Member, IEEE*

***Abstract*— This paper describes noise reduction combining microphones and laser listening devices. A microphone array is one of the useful approaches for reducing the noise. However, when the microphones are mounted on robot systems, the problem of internal noise from robots such as motors and gears arises. It is difficult to reduce the internal noise utilizing the conventional microphone array because the noise source is extremely close to the microphones. As the internal noise is not always stationary or sparse, some useful blind source separation (BSS) approaches such as Independent Component Analysis (ICA) and sparseness approach cannot be employed. Our aim is to reduce the internal noise by a sensor fusion of microphones and laser listening devices. In this paper, we describe the problem of the typical microphone array and show the design of the laser listening devices. Noise reduction combining microphones and laser listening devices is then described. Experimental results are also given to show the effectiveness of the proposed method.**

I. Introduction

NOISE reduction is an important aspect of human hearing and understanding of sounds. There are many studies concerning noise reduction. A microphone array is one of the useful approaches for reducing the noise. It can reduce the noise utilizing the phase difference or amplitude difference of each microphone [1], [2], [3]. Lots of methods based on the microphone array have been reported such as the delay-sum type microphone array [4], the adaptive microphone array [5], [6], [7] and Direction of Arrival (DOA) estimation [8], [9], [10]. However, when the conventional microphone array is mounted on the real system such as machines, vehicles, and robots, the internal noises from the systems often become problems. In those cases, the noise source is extremely close to the microphones. The number of noise cannot always be known. Moreover, the noise intensity is also often larger than the sound, because the main sound propagation path is not the air but the solid frame of the machine. Due to these reasons, it is difficult to reduce the internal noise from the systems such as motor noise and gear noise only by utilizing the conventional microphone array. Some useful blind source separation (BSS) algorithms have also been reported to detect the target signal. Independent Component Analysis (ICA) is one of the most promising BSS algorithms. ICA utilizes the statistical independency of received sounds and is based on the unitary diagonalization of the whitened data covariance matrix [11], [12], [13]. There also exist some BSS algorithms that utilize higher order statistics [14]. However, since the internal noise is not always stationary or statistically independent from other sources, it is difficult to employ ICA to reduce the internal noise. When the number of the noise sources is larger than that of the microphones, ICA also cannot be employed. Some approaches utilize sparseness to solve the BSS problem [15], [16], [17]. Sparseness means that most of the frequency components of a signal are zero, that is, the sources rarely overlap in the frequency domain. Under this assumption, it is possible to extract each signal using time-frequency binary masks. However, it is difficult to employ sparseness approach for reducing the internal noise because the internal noise is not always sparse.

To reduce the internal noise, the authors have already reported an acoustical array combining microphones and piezoelectric devices [18], [19]. However, it is sometimes difficult to attach the piezoelectric devices on the objects such as precision components and precision instruments. To solve the problems, we employ a laser listening device instead of a piezoelectric device. The laser listening device is a simple active sensor, which can directly detect the object vibration by utilizing the laser reflection without any contact on the objects. It is utilized for eavesdropping because it can directly detect the object vibration from the distant place ($\sim$ 1km). However, it is difficult to utilize the laser listening device as a general microphone to detect the sound wave because the laser beam has to point the vibrating object directly. Hence, there are few examples of laser listening devices in living space. Our aim is to reduce the internal noise of the microphones from the system such as machines, vehicles, and robots by utilizing the characteristics difference between microphones and laser listening devices. In Sec.II, we describe a general formulation and the problem of the conventional microphone array. We then give a description of the design concept of laser listening devices in Sec.III. Noise reduction combining microphones and laser listening devices is described in Sec.IV. The experimental results are shown to confirm the performance of the proposed method in Sec.V. Conclusions are given in Sec.VI.

II. Problem formulation

According to the survey by O'Grady *et al*, signal mixtures of an acoustical array are categorized as the instantaneous, anechoic and echoic mixings [20]. In this paper, we discuss an echoic mixing. Let us consider a sound with N kinds of noise recorded by J microphones in an echoic room.

M. Matsumoto is with the Department of Applied Physics, Faculty of Science and Engineering, Waseda University, 55N-4F-10A, 3-4-1, Okubo, Shinjuku-ku, Tokyo, Japan, (e-mail: matsu@shalab.phys.waseda.ac.jp)

T. Abe is with the Major in Pure and Applied Physics, Faculty of Science and Engineering, Waseda University, 55N-4F-10A, 3-4-1, Okubo, Shinjuku-ku, Tokyo, Japan, (e-mail: tomomi@shalab.phys.waseda.ac.jp)

S. Hashimoto is with the Department of Applied Physics, Faculty of Science and Engineering, Waseda University, 55N-4F-10A, 3-4-1, Okubo, Shinjuku-ku, Tokyo, Japan, (e-mail: shuji@waseda.jp)

978-1-4244-2632-4/08/$25.00 ©2008 IEEE

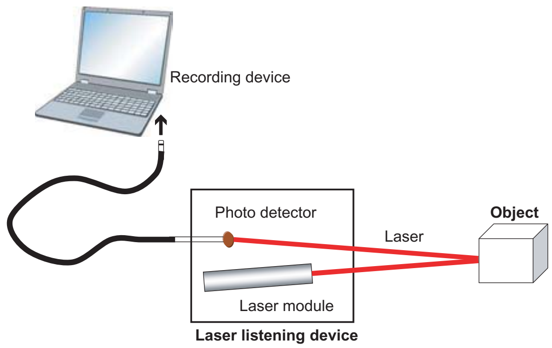

Fig. 1. Design concept of a laser listening device.

$s(t)$ and $n_j(t)$ represent the target signal and the jth noise, respectively. The signal obtained by the ith microphone $x_i(t)$ can be described as follows:

$$x_i(t) = m_i * s(t) + \sum_j m_{ij} * n_j(t) \quad (1)$$

where m_i denotes the acoustic impulse response from the desired signal to the ith microphone. m_{ij} denotes the acoustic impulse response from the jth noise to the ith microphone. The asterisk represents the convolution. In the frequency domain, we can express the J mixtures $\mathbf{X}(\tau,\omega)$ as follows:

$$\mathbf{X}(\tau,\omega) = \begin{bmatrix} X_1(\tau,\omega) \\ X_2(\tau,\omega) \\ \cdots \\ X_J(\tau,\omega) \end{bmatrix} = \mathbf{M}(\omega) \begin{bmatrix} S(\tau,\omega) \\ N_1(\tau,\omega) \\ N_2(\tau,\omega) \\ \cdots \\ N_N(\tau,\omega) \end{bmatrix} \quad (2)$$

where

$$\mathbf{M}(\omega) = \begin{bmatrix} M_1(\omega) & M_{11}(\omega) & \cdots & M_{1N}(\omega) \\ M_2(\omega) & M_{12}(\omega) & \cdots & M_{2N}(\omega) \\ \cdots & & & \cdots \\ M_J(\omega) & M_{1N}(\omega) & \cdots & M_{JN}(\omega) \end{bmatrix} \quad (3)$$

τ and ω represent the time frame and angular frequency, respectively. Let us consider the output vector of the microphone array $\mathbf{Y}(\tau,\omega)$ defined as follows:

$$\mathbf{Y}(\tau,\omega) = \mathbf{W}(\omega)\mathbf{X}(\tau,\omega) \quad (4)$$

where $\mathbf{W}(\omega)$ represents the $J \times (N+1)$ weight matrix. The aim of microphone array is to obtain an optimal $\mathbf{W}(\omega)$. However, in practical systems such as robot systems, the number of noise N cannot always be determined. The number of the noise is also often larger than the signal source. The noise is often nonstationary and not disjoint in the frequency domain. Hence in practical case, it is difficult to reduce the noise such as motor noise and gear noise by utilizing the typical microphone array.

III. Device design

In this section, we briefly describe the device design of the laser listening device. Fig.1 illustrates the design concept of the laser listening device. As shown in Fig.1, the laser listening device is composed of a laser module and a photo detector. The laser module is set apart from the target object and the laser beam aims at the object. The photo detector detects the reflected beam. The laser module is active and fixed during the use. The output of the photo detector is connected to the recording device such as a personal computer (PC). When the object vibrates, the reflected beam also vibrates depending on the object vibration. The photo detector catches the reflective light from the object and directly transforms the light intensity into the voltage. Although the design is simple, we can record the sound by utilizing the above laser listening device as well as typical microphones. Fig.2 illustrates the concrete examples of the laser listening device.

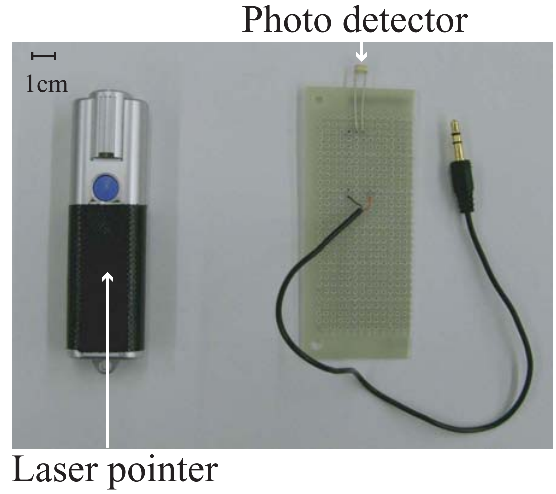

Fig. 2. Experimental tools.

IV. Noise reduction combining microphones and laser listening devices

When we consider the noise in the robot system, the noise can be categorized into two types. One is *external noise* coming from environment. Another is *internal noise* coming from the inside of the robots such as motors. Although the external noise can usually be regarded as the point source of sound, it is difficult to know the number and location of the noise in advance. The external noise sources also often move in the living space. On the other hand, although the number and location of the internal noise can often be known in advance, the noise source is not always regarded as the point source of sound. We pay attention to the features of these two kinds of noise, and aim at directly reducing the internal noise by utilizing microphones and laser listening devices. The process of the algorithm is as follows:

We employ not only typical microphones but also laser listening devices for detecting the noise source vibration such as motor noise and gear noise. Let us consider J microphones, K kinds of internal noise and N kinds of external noise. To reduce the internal noise, we also prepared K laser listening devices corresponding to each internal noise. We set the laser listening devices so that it can detect the vibration of the noise sources. The signal obtained by the ith microphone $x_i^M(t)$ can be described as follows:

$$x_i^M(t) = m_i * s(t) + \sum_j m_{ij}^I * n_j^I(t) + \sum_k m_{ik}^E * n_k^E(t) \quad (5)$$

where $n_j^I(t)$ and $n_k^E(t)$ represent the jth internal noise and kth external noise, respectively. m_i denotes the acoustic impulse responses from the target signal $s(t)$ to the ith microphone. m_{ij}^I denotes the acoustic impulse responses from the jth internal noise to the ith microphone. m_{ik}^E denotes the acoustic impulse responses from the kth external noise to the ith microphone. The signal obtained by the ith laser listening device $x_i^L(t)$ can also be described as follows:

$$x_i^L(t) = l_i * s(t) + \sum_j l_{ij}^I * n_j^I(t) + \sum_k l_{ik}^E * n_k^E(t) \quad (6)$$

where l_i denotes the acoustic impulse responses from the target signal to the ith laser listening device. l_{ij}^I denotes the acoustic impulse responses from the jth internal noise to the ith laser listening device. l_{ik}^E denotes the acoustic impulse responses from the kth external noise to the ith laser listening device. The aim of setting laser listening devices is to detect only the internal noise information by measuring the physical vibration of the noise sources as the effects of the target sound and the external noises to the robot body vibrations. In other words, we can assume the following equations:

$$l_i * s(t) = 0 \quad (7)$$

$$l_{ik}^E * n_k^E(t) = 0 \quad (8)$$

We can also assume that each internal noise can be separately detected by setting the laser beam point adequately. Hence, we can assume the following equations:

$$l_{ij}^I * n_j^I(t) = 0 \qquad (i \neq j) \quad (9)$$

It should be noted that a internal noise source may include multiple causes for the object vibration such as multiple motors inside of the robot body. Here, we consider the problem in the frequency domain. Let us define the output vector of the microphones and laser listening devices $\mathbf{X^M}(\tau,\omega)$ and $\mathbf{X^L}(\tau,\omega)$ as follows:

$$\mathbf{X^M}(\tau,\omega) = [X_1^M(\tau,\omega), X_2^M(\tau,\omega), \cdots, X_J^M(\tau,\omega)]^T \quad (10)$$

$$\mathbf{X^L}(\tau,\omega) = [X_1^L(\tau,\omega), X_2^L(\tau,\omega), \cdots, X_K^L(\tau,\omega)]^T \quad (11)$$

T represents transpose. We can also express the noise vector of the internal noise and external noise $\mathbf{N^I}(\tau,\omega)$ and $\mathbf{N^E}(\tau,\omega)$ as follows:

$$\mathbf{N^I}(\omega) = [N_1^I(\tau,\omega), N_2^I(\tau,\omega), \cdots, N_K^I(\tau,\omega)]^T \quad (12)$$

$$\mathbf{N^E}(\omega) = [N_1^E(\tau,\omega), N_2^E(\tau,\omega), \cdots, N_N^E(\tau,\omega)]^T \quad (13)$$

The mixing vector $\mathbf{M}(\omega)$ regarding signal of the microphones is defined as follows:

$$\mathbf{M}(\omega) = [M_1(\omega), M_2(\omega), \cdots, M_J(\omega)]^T \quad (14)$$

The mixing matrices $\mathbf{M^I}(\omega)$ and $\mathbf{M^E}(\omega)$ regarding both the internal noise and the external noise of the microphones are defined as follows:

$$\mathbf{M^I}(\omega) = \begin{bmatrix} M_{11}^I(\omega) & \cdots & M_{1K}^I(\omega) \\ M_{21}^I(\omega) & \cdots & M_{2K}^I(\omega) \\ \cdots & & \cdots \\ M_{J1}^I(\omega) & \cdots & M_{JK}^I(\omega) \end{bmatrix} \quad (15)$$

$$\mathbf{M^E}(\omega) = \begin{bmatrix} M_{11}^E(\omega) & \cdots & M_{1N}^E(\omega) \\ M_{21}^E(\omega) & \cdots & M_{2N}^E(\omega) \\ \cdots & & \cdots \\ M_{J1}^E(\omega) & \cdots & M_{JN}^E(\omega) \end{bmatrix} \quad (16)$$

In the same way, the mixing vector $\mathbf{L}(\omega)$ regarding signal of the laser listening device is defined as follows:

$$\mathbf{L}(\omega) = [L_1(\omega), L_2(\omega), \cdots, L_K(\omega)]^T \quad (17)$$

The mixing matrices $\mathbf{L^I}(\omega)$ and $\mathbf{L^E}(\omega)$ regarding the internal noise and the external noise of the laser listening devices are expressed as follows:

$$\mathbf{L^I}(\omega) = \begin{bmatrix} L_{11}^I(\omega) & \cdots & L_{1K}^I(\omega) \\ L_{21}^I(\omega) & \cdots & L_{2K}^I(\omega) \\ \cdots & & \cdots \\ L_{K1}^I(\omega) & \cdots & L_{KK}^I(\omega) \end{bmatrix} \quad (18)$$

$$\mathbf{L^E}(\omega) = \begin{bmatrix} L_{11}^E(\omega) & \cdots & L_{1N}^E(\omega) \\ L_{21}^E(\omega) & \cdots & L_{2N}^E(\omega) \\ \cdots & & \cdots \\ L_{K1}^E(\omega) & \cdots & L_{KN}^E(\omega) \end{bmatrix} \quad (19)$$

In frequency domain, we can express the L mixtures of microphones as follows:

$$\mathbf{X}(\tau,\omega) = \mathbf{H}(\omega)\mathbf{S}(\tau,\omega) \quad (20)$$

where

$$\mathbf{X}(\tau,\omega) = \begin{bmatrix} \mathbf{X^M}(\tau,\omega) \\ \mathbf{X^L}(\tau,\omega) \end{bmatrix} \quad (21)$$

$$\mathbf{H}(\omega) = \begin{bmatrix} \mathbf{M}(\omega) & \mathbf{M^E}(\omega) & \mathbf{M^I}(\omega) \\ \mathbf{L}(\omega) & \mathbf{L^E}(\omega) & \mathbf{L^I}(\omega) \end{bmatrix} \quad (22)$$

$$\mathbf{S}(\tau,\omega) = \begin{bmatrix} S(\tau,\omega) \\ \mathbf{N^E}(\tau,\omega) \\ \mathbf{N^I}(\tau,\omega) \end{bmatrix} \quad (23)$$

$M_i(\omega)$ and $M_{ij}^E(\omega)$ will change in robot systems because the positional relationship among the external noise, the microphones and laser listening devices often changes. Conversely, $M_{ij}^I(\omega)$ and $L_{ij}^I(\omega)$ will not change because the positional relations among the internal noise, the microphones and laser listening devices are fixed. Therefore, $M_{ij}^I(\omega)$ and $L_{ij}^I(\omega)$ can be measured in advance, while the other transfer functions cannot be measured in advance. We pay attention to the above features and define $\mathbf{W^E}(\omega)$ and $\mathbf{W^I}(\omega)$ as the weight matrices for reducing the external noise and the internal noise, respectively. The output vector $\mathbf{Y}(\tau,\omega)$ of microphones and laser listening devices is described as follows:

$$\mathbf{Y}(\tau,\omega) = \mathbf{W^E}(\omega)\mathbf{W^I}(\omega)\mathbf{X}(\tau,\omega) \quad (24)$$

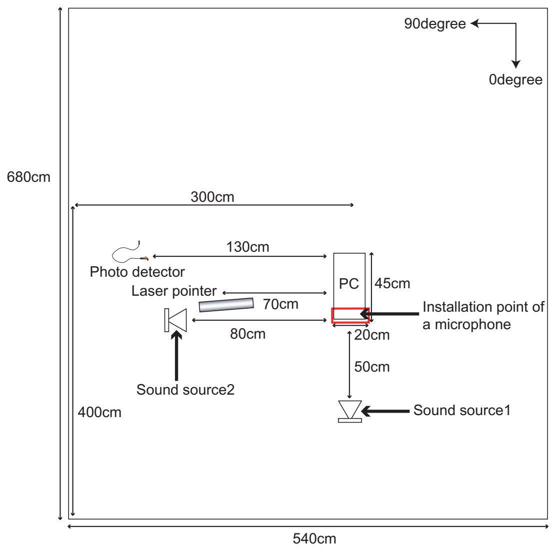

Fig. 3. Experimental setup. (Top view)

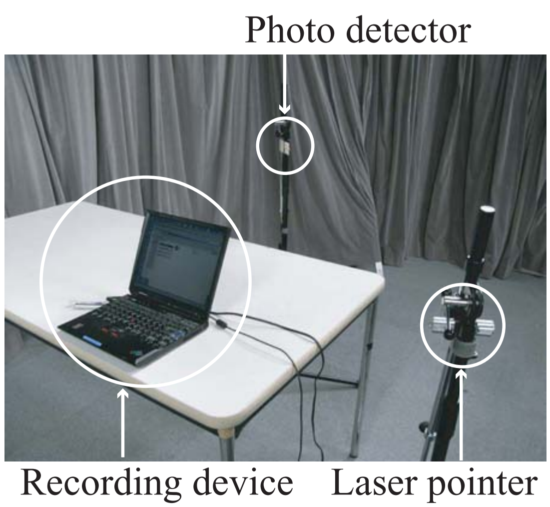

Fig. 4. The photograph of the experimental setup.

Hence, we firstly find the matrix $\mathbf{W^I}(\omega)$ as follows:

$$\mathbf{H}'(\omega) = \mathbf{W^I}(\omega)\mathbf{H}(\omega) = \begin{bmatrix} \mathbf{M}(\omega) & \mathbf{M^E}(\omega) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (25)$$

We then optimize $\mathbf{W^E}(\omega)$ by employing various kind of method based on the microphone array such as beamforming and adaptive microphone array, ICA and sparseness approaches.

V. Experiment

We conducted the experiments to evaluate the performance of the proposed method. Throughout the experiments, we aim to reduce the noise from the PC. Fig.3 shows the experimental setup. As shown in Fig.3, we set a laser pointer 70 cm apart from PC. We also set a photo detector 130 cm apart from the PC as it could receive the reflected light. The

Fig. 5. The photograph of the microphone setup.

sound source from 0 degree (Sound source 1) was set 50 cm apart from the PC. The sound source from 90 degree (Sound source 2) was set 80 cm apart from the PC. Figs.4 and 5 show the photograph of the experimental setup and the microphone setup, respectively. As shown in Fig.5, a microphone (RP-VC200, Panasonic) was set onto the PC. The recording room was not a particular room such as an anechoic room but an ordinary studio because we aim to apply the proposed method to the systems such as machines, vehicles and robots in the real environment. The reverberation time is 72 ms. The transfer functions from the motor to the laser listening device and to the microphones were measured by cross-spectrum method. As the sound sources, we selected ten male voices and ten female voices from "Japanese Newspaper Article Sentences" edited by the Acoustical Society of Japan. We evaluated Signal to Noise Ratio (SNR) as follows:

$$SNR = 10\log_{10}\frac{\sum_{t=0}^{T} s^2(t)}{\sum_{t=0}^{T} n^2(t)} \quad (26)$$

where T represents the duration of observation. Noise Reduction Ratio (NRR) is defined as follows:

$$NRR = SNR_{after} - SNR_{before} \quad (27)$$

where SNR_{before} and SNR_{after} represent the SNR before and after noise reduction, respectively. Tables.I and II show the results of SNR_{before} and NRR when the sound was generated from 0 degree and 90 degree, respectively. As shown in Tables.I and II, the proposed method could reduce the noise about 5dB throughout the experiments although the SNR_{before} was changed and the noise was sometimes extremely big.

VI. Conclusion

In this paper, we proposed a method for noise reduction combining microphones and laser listening devices. The proposed method can reduce the internal noise even if the number of the noise is larger than the microphones and laser

Gender	Sound	SNR_{before}(dB)	NRR(dB)
Male	Sound1	4.6756	5.4056
	Sound2	7.1097	5.682
	Sound3	6.5189	5.633
	Sound4	10.5039	5.6273
	Sound5	5.9519	5.4342
	Sound6	7.7476	5.5302
	Sound7	4.6702	5.5195
	Sound8	3.4505	5.4542
	Sound9	5.88	5.5499
	Sound10	1.6207	5.4721
Female	Sound11	12.4034	5.603
	Sound12	12.2813	5.6377
	Sound13	13.9176	5.5741
	Sound14	13.5222	5.6044
	Sound15	14.3532	5.5563
	Sound16	9.0463	5.5791
	Sound17	8.4928	5.5395
	Sound18	10.4704	5.6186
	Sound19	7.9323	5.5039
	Sound20	9.3797	5.6509

TABLE I

NRR RESULTS REGARDING MICROPHONE 1 WHEN THE SOUND WAS GENERATED FROM 0 DEGREE.

Gender	Sound	SNR_{before}(dB)	NRR(dB)
Male	Sound1	-3.7229	5.1815
	Sound2	-1.9621	5.254
	Sound3	-2.6817	5.1595
	Sound4	-0.1354	4.9431
	Sound5	-2.7837	4.7964
	Sound6	-1.6411	5.0186
	Sound7	-3.5784	5.2232
	Sound8	-3.6856	4.6572
	Sound9	-2.8947	5.1458
	Sound10	-4.1261	5.1912
Female	Sound11	1.0794	4.9512
	Sound12	1.7097	5.0225
	Sound13	4.1714	4.9956
	Sound14	2.2844	5.0263
	Sound15	2.819	4.978
	Sound16	-1.6268	4.924
	Sound17	-1.8483	4.8164
	Sound18	-1.8483	4.8164
	Sound19	-1.4002	5.1145
	Sound20	-0.7583	5.1122

TABLE II

NRR RESULTS REGARDING MICROPHONES 1 WHEN THE SOUND WAS GENERATED FROM 90 DEGREE.

listening devices without the contact on the objects. The proposed method is not conflicting but compatible with the conventional techniques for noise reduction. We can employ the proposed method as the preprocessing of the conventional microphone array. For future works, we aim to develop the method combining the proposed method and the conventional algorithm of the microphone array such as beam-forming, adaptive microphone array, ICA and sparseness approaches. The optimal settings of the laser listening devices should also be considered.

ACKNOWLEDGEMENT

This research was supported by the research grant of Support Center for Advanced Telecommunications Technology Research (SCAT), by the research grant of Tateisi Science and Technology Foundation, and by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Young Scientists (B), 20700168, 2008. This research was also supported by Waseda University Grant for Special Research Projects(B), No.2007B-142, No.2007B-143 and No.2007B-168, by "Establishment of Consolidated Research Institute for Advanced Science and Medical Care", Encouraging Development Strategic Research Centers Program, the Special Coordination Funds for Promoting Science and Technology, Ministry of Education, Culture, Sports, Science and Technology, Japan, by the CREST project "Foundation of technology supporting the creation of digital media contents" of JST, by the Grant-in-Aid for the WABOT-HOUSE Project by Gifu Prefecture, the 21st Century Center of Excellence Program, "The innovative research on symbiosis technologies for human and robots in the elderly dominated society", Waseda University and by Project for Strategic Development of Advanced Robotics Elemental Technologies (NEDO: 06002090).

REFERENCES

[1] W. Li, C. Miyajima, T. Nishino, K. Itou, K. Takeda and F. Itakura, "Adaptive Nonlinear Regression Using Multiple Distributed Microphones for In-car Speech Recognition," *IEICE trans on Fundamentals.*, Vol.E88-A, No.7, pp.1716–1723, 2005.

[2] J. Cho, A. Krishnamurthy, "Speech Enhancement Using Microphone Array in Moving Vehicle Environment," *Prof. of IEEE Intelligent Vehicles Symposium*, pp.366–371, 2003.

[3] J.-M. Valin, J. Rouat, and F. Michaud, "Enhanced Robot Audition Based on Microphone Array Source Separation with Post-filter", *Proc. of IEEE/RSJ Int'l Conf. on Robots and Intelligent Systems.*, pp.2123–2128, 2004.

[4] K. Kiyohara, Y. Kaneda, S. Takahashi, H. Nomura and J. Kojima, "A Microphone Array System for Speech Recognition," *Proc. of IEEE Int'l Conf. on Acoustics, Speech, and Signal Processing.*, pp.215–218, 1997.

[5] Y. Kaneda, and J. Ohga, "Adaptive Microphone Array System for Noise Reduction," *IEEE Trans on Acoust. Speech Source Process.*, Vol.ASSP-34, No.6, pp.1391–1400, 1986.

[6] K. Takao, M. Fujita, and T. Nishi, "An Adaptive Antenna Array under Directional Constraint," *IEEE Trans on Antennas Propagat.*, Vol.24, pp.662–669, 1976.

[7] M. Matsumoto, and S. Hashimoto, "A Miniaturized Adaptive Microphone Array under Directional Constraint Utilizing Aggregated Microphones," *J. Acoustic Society America.*, Vol.119, No.1, pp.352–359, 2006.

[8] J. Capon, "High Resolution Frequency-wavenumber Spectrum Analysis," *Proc. of IEEE.*, Vol.57, pp.2408–2418, 1969.

[9] R. O. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," *IEEE Trans on Antennas Propagat.*, Vol.AP-34, pp.276–280, 1986.

[10] M. Matsumoto, and S. Hashimoto, "Multiple Signal Classification by Aggregated Microphones," *IEICE trans on Fundamentals.*, Vol.E88-A, No.7, pp.1701–1707, 2005.

[11] A. Belouchrani, K. Abed Meraim, J. F. Cardoso and E. Moulines, "A Blind Source Separation Technique Based on Second Order Statistics," *IEEE Trans on Signal Processing.*, Vol.45 No.2, pp.434–444, 1997.

[12] D. Nuzillard, J. M. Nuzillard, "Second Order Blind Source Separation on the Fourier Space of Data," *Signal Processing.*, Vol.83, No.3, pp.627–631, 2003.

[13] K. I. Diamantaras, A. P. Petropulu and B. Chen, "Blind Two-Input-Two-Output FIR Channel Identification Based On Frequency Domain Second-Order Statistics," *IEEE Trans on Signal Processing.*, Vol.48, No.2, pp.534–542, 2000.

[14] B. Chen and A. P. Petropulu, "Frequency Domain Blind MIMO System Identification Based on Second- And Higher-Order Statistics," *IEEE Trans on Signal Processing.*, Vol. 49, No. 8, pp.1677–1688, 2001.

[15] T. Ihara, M. Handa, T. Nagai and A. Kurematsu, "Multi-channel Speech Separation and Localization by Frequency Assignment," *IEICE trans on Fundamentals.*, Vol.J86-A, No.10, pp.998–1009, 2003.

[16] S. Rickard and O. Yilmaz, "On the Approximate W-disjoint Orthogonality of Speech," *Proc. of IEEE Int'l Conf. on Acoustics, Speech, and Signal Processing 2002.*, pp.529–532, 2002.

[17] O. Yilmaz and S. Rickard, "Blind Separation of Speech Mixtures via Time-frequency Masking," *IEEE Trans on Signal Processing.*, Vol.52, No.7, pp.1830–1847, 2004.

[18] M. Matsumoto, S. Hashimoto, "Noise Reduction Combining Microphones and Piezoelectric Devices," *Proc. of Int'l Symp. on Communications, Control and Signal Process.*, pp.368–371, 2008.

[19] M. Matsumoto, S. Hashimoto, "An Acoustical Array Combining Microphones and Piezoelectric Devices," *J. Acoustic Society America.*, Vol.123, No.4, pp.2697–2705, 2008.

[20] P. D. O'Grady, B. A. Pearlmutter, and S. T. Rickard. "Survey of Sparse and Non-Sparse Methods in Source Separation," *Int'l Journal of Imaging Systems and Technology.*, Vol.15, pp.18–33, 2005.