AN ADAPTIVE MICROPHONE ARRAY WITH GOOD SOUND QUALITY USING AUXILIARY FIXED BEAMFORMERS AND ITS DSP IMPLEMENTATION

Osamu HOSHUYAMA and Akihiko SUGIYAMA

C&C Media Research Laboratories
NEC Corporation
E-mail: houshu@ccm.CL.nec.co.jp

ABSTRACT

This paper presents an adaptive microphone array using two auxiliary fixed beamfomers for good sound quality. One auxiliary fixed beamfomer is introduced in the target signal path to avoid suppression of high-frequency components in the total output. The other auxiliary fixed beamfomer is used for adaptation-mode control to eliminate the hysteresis in the relationship between signal direction and sensitivity. Both auxiliary fixed beamfomers bring about good sound quality, which improve intelligibility in speech communications and speech recognition rate. The proposed microphone array is implemented on a DSP system, which demonstrates flat frequency response and less hysteresis in its directivity pattern.

1. INTRODUCTION

Microphone arrays based on beamforming techniques [1]–[7] have been studied to reduce or eliminate directional interferences. Specifically, adaptive microphone arrays [1]–[6] are promising because they can attain high interference-reduction performance with a small number of microphones arranged in a small space. In the actual environment, they require robustness against array imperfections such as position error and target-direction error. For this robustness, an adaptive microphone array with an adaptive blocking matrix has been proposed [5]. It is a generalized sidelobe canceller comprising a fixed beamfomer (FBF), an adaptive blocking matrix (ABM), a multiple-input canceller (MC), and an adaptation-mode controller (AMC). The ABM provides robustness sufficient for actual applications, and high interference reduction performance. However, its performance may be limited by sharing the same component for multiple purposes.

First of all, the single FBF is shared by target-signal blocking in the ABM and target-signal enhancement in the target-signal path to the MC. The FBF can not be optimum for both purposes because the degrees of freedom are limited. An FBF optimized for target signal blocking has a narrow beam in its directivity pattern, which could cause attenuation of high frequency components in the target-signal path. This high frequency attenuation degrades sound quality in speech communications and can be a serious problem for speech recognition.

The ABM is also shared by target-signal blocking and interference extraction for the AMC. The AMC using the ABM output has a problem of hysteresis in the total directivity pattern because the immature convergence of the ABM causes imperfect target detection. Degradation of sound quality by the hysteresis is annoying for users.

This paper proposes a new robust adaptive microphone array with an adaptive blocking matrix (RAMA-ABM) with good sound quality. It introduces two auxiliary FBFs to preserve high frequency components and to eliminate the hysteresis. The proposed RAMA-ABM is implemented on a DSP (Digital Signal Processor) system and evaluated in a real environment.

2. CONVENTIONAL RAMA-ABM

Structure of the conventional RAMA-ABM is shown in Fig. 1. It consists of an FBF, a BM, an MC and an AMC. The FBF enhances the target signal. The ABM adaptively blocks the target signal and passes only interference signals using the output signal of the FBF. It may perform rough target tracking, which leads to robustness by absorbing the influence of array imperfections [3]. The MC adaptively extracts the target signal by subtracting the ABM output signals from the FBF output. In the MC, the norm of coefficients is constrained by an inequality, which also leads to robustness when target blocking in the ABM is insufficient [4]. The AMC controls the adaptation of the ABM and the MC based on target detection using an estimate of SIR (Signal-to-Interference Ratio) [5].

This RAMA-ABM has high interference reduction capability. However, its frequency response has directional dependency because of the FBF's directivity. If the target direction is off the assumed direction, high frequency components are attenuated. Let us assume that the target tracking by the ABM is perfect and the target signal is completely blocked at the outputs of the ABM. Frequency dependency of the directivity pattern in the allowable target region is dominated by that of the FBF in the signal path. If this FBF in Fig. 1 is a delay-and-sum beamformer, sensitivity at the total output has a frequency-dependent directional pattern. Even though the FBF is a constant-width beamformer [7], the target signal level at the total output is still attenuated, which may be significant for some speech recognition systems.

The conventional RAMA-ABM has another problem of hysteresis in its AMC. It detects target signal based on the power ratio of the output of the FBF to an output signal of the ABM. However, the power ratio is influenced by the convergence status. For simplicity, let us assume an environment where the interference is stationary and the target source has burst characteristics like speech. When the target signal stops, adaptation of the ABM, i.e. target tracking, also stops. As shown in Fig. 2, the ABM has a dip in the directivity pattern. The direction and the depth of the dip vary with convergence status of the ABM. If the target signal

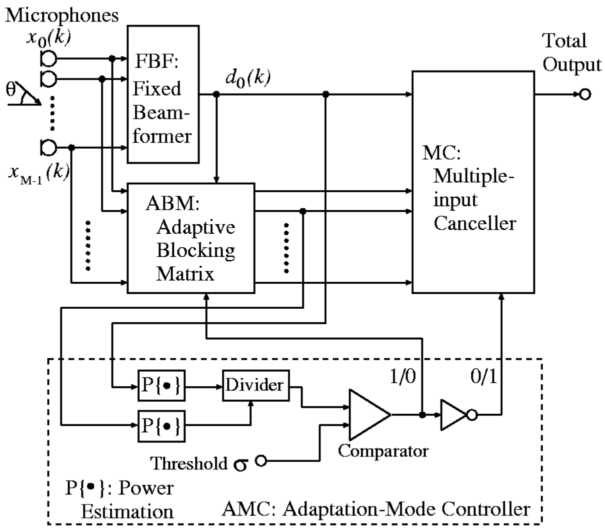

Figure 1: Structure of Conventional RAMA-ABM.

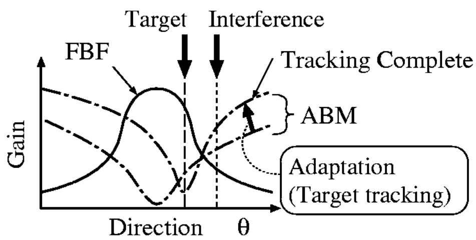

Figure 2: Directivity Patterns of FBF and ABM. (Not quantitative)

stops when the dip of the ABM is not sufficiently deep, the gain of the ABM to the interference is larger than its desirable value. Therefore, the output power of the ABM is larger, and a larger target signal power is required for starting next adaptation of the ABM.

If signal source is single, this effect is observed as a hysteresis in its directivity pattern. When the dip is off the assumed target direction, it is smaller than when it is at the target direction. This is illustrated in Fig. 2. For the users, the beginning of the target speech gives low intelligibility or sometimes the speech itself may be missing because of the failure in target tracking.

3. PROPOSED RAMA-ABM

Figure 3 shows the structure of the proposed RAMA-ABM, which includes two auxiliary FBFs. One auxiliary FBF is introduced in the target signal path and the other auxiliary FBF is used in adaptation-mode control.

3.1. Auxiliary FBF in Signal Path

Auxiliary FBF is first introduced to avoid suppression of high-frequency components in the total output. A new FBF (FBF1 in Fig. 3) is inserted in the target signal path instead of the original FBF (FBF in Fig. 1 or FBF0 in Fig. 3). FBF1 has an almost flat frequency response in the allowable target-direction region as shown in Fig. 4(b). It is flatter than that of the delay-and-sum beamformer used for FBF0. Because the response in the allowable target region is dominated by the FBF in the signal path, directional dependency is eliminated by using FBF1 for the original FBF. Therefore, the sound quality of the total output signal is improved. As FBF1 can just be an arbitrary one of the microphones, this structure requires almost no additional computation for sufficient performance.

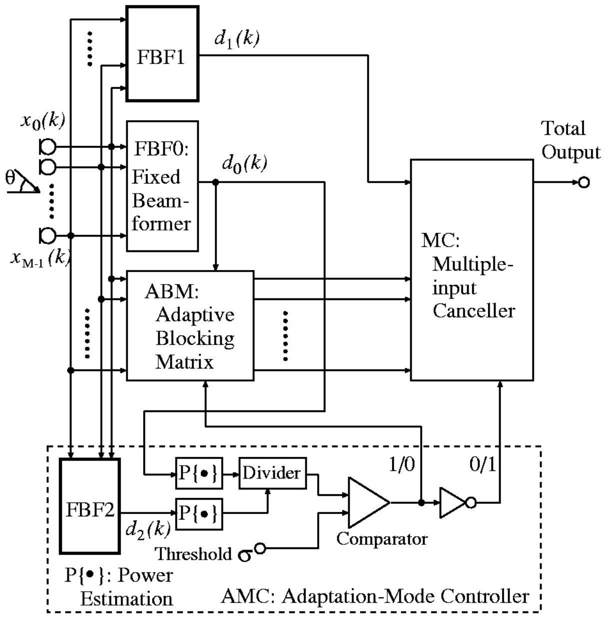

Figure 3: Structure of Proposed Adaptive Microphone Array.

This substitution causes some degradation in interference reduction performance. However, when the array size is small, which is a common situation with applications of adaptive microphone arrays, the interference reduction by FBF0 is small. It will be less significant than ripples in the frequency and directional responses to the target signal. FBF1 has another advantage that FBF0 for the ABM could only be optimized for the target tracking in the ABM.

3.2. Auxiliary FBF for Adaptation-Mode Controller

The proposed RAMA-ABM utilizes another auxiliary FBF (FBF2 in Fig. 3) in the AMC instead of the ABM. FBF2 operates as a fixed blocking matrix, and eliminates the hysteresis in the relationship between signal direction and sensitivity. FBF2 has a similar beam pattern to that of the ABM. However, it is non-adaptive. The simplest realization of FBF2 is a delay-and-subtraction beamformer, which is used in simple Griffiths-Jim beamformer [6].

FBF2 is independent of the convergence or tracking of the ABM. As a result, the RAMA-ABM with the proposed AMC is free from hysteresis in its directivity pattern. Therefore, users can enjoy good sound quality.

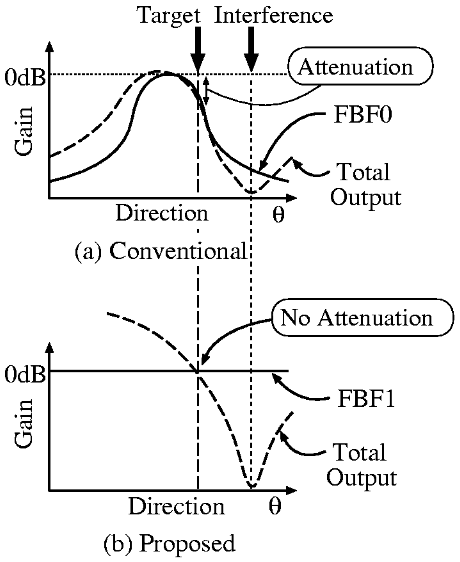

Figure 4: Comparison of Two FBFs in Signal Path. (Not quantitative)

4. IMPLEMENTATION AND EVALUATION ON A REALTIME PROCESSOR

The proposed RAMA-ABM was implemented on a DSP system and its directivity was evaluated. A single board DSP system [8] accommodating a floating point DSP, ADSP-21060 by Analog Devices [9] was employed for implementation. This realtime system is shown in Fig. 5. The assembly language programming was performed on a personal computer, which displays all the filter coefficients every second. The sampling frequency was 8 kHz. The system has processing power sufficient for the proposed RAMA-ABM with 8 channel signals. However, only 4 microphones were used. An equi-spaced linear array using omni-directional microphones with a spacing of 4.1 cm was used.

On the realtime processing system, the proposed RAMA-ABM was compared to the conventional RAMA-ABM. FBF0 and FBF1 were also included in comparison as references. FBF in Fig. 1 and FBF0 were delay-and-sum beamformers expressed by:

$$d_0(k) = \frac{1}{M}\sum_{i=0}^{M-1} x_i(k), \tag{1}$$

where $d_0(k)$ is the output signal of FBF0 at the k-th sample, M is the number of microphones ($M = 4$), and $x_i(k)$ is the i-th microphone signal. FBF1 was implemented by the center microphone (the second for $M = 4$), and FBF2 was constructed to provide the difference of the $\frac{M}{2}$-th and $(\frac{M}{2} - 1)$-th (the second and the third for $M = 4$) microphones as follows:

$$d_1(k) = x_{\frac{M}{2}}(k), \tag{2}$$

$$d_2(k) = x_{\frac{M}{2}}(k) - x_{(\frac{M}{2}-1)}(k), \tag{3}$$

where $d_1(k)$ and $d_2(k)$ are the output signals of FBF1 and FBF2, respectively. The step sizes selected were 0.02 for the ABM and 0.005 for the MC. The constraints of the ABM were set so that the allowable target-direction range was approximately $\pm 15°$, and the norm constraint [4] of the MC was 10. In both AMCs, time constant for averaging the estimates of powers was about 10 mS. The threshold σ for the AMC [5] was 1.7 for the allowable direction range of $\pm 15°$.

The experimental set-up is shown in Fig. 6. Reverberation time of the room was 0.3 second. A white-noise source was scanned in two ways from 0° to 50° at a distance of 2.0 m. Output powers of the system with 5-degree intervals and its power spectra in two directions (0° and 15°) were measured.

Figure 7 shows power spectra of the total output in the two directions ($\theta = 0°$ and 15° in the right scan). Because of flatter echo of the room, these spectra have common large ripples. Therefore, only the envelopes were compared. When a signal source was located in front of the array ($\theta = 0°$), both the conventional and the proposed RAMA-ABMs caused equal suppression in high frequency components (from 2.5k to 3.5kHz) as shown in Fig. 7 (a) and (b). However, at $\theta = 15°$, a little off the center in the allowable target-direction range, the proposed RAMA-ABM had higher power than the conventional RAMA-ABM by as much as 6dB at 3 kHz. This is because FBF1 has higher sensitivity to high frequency components than FBF0 as shown in Fig. 7 (c) and (d). This response of the proposed RAMA-ABM means that the extracted target signal at the MC output has good sound quality.

Output powers normalized by that at the center are plotted in Fig. 8. As is clear from Fig. 8 (a), the proposed RAMA-ABM exhibited almost no hysteresis thanks to FBF2. On the other hand, the conventional RAMA-ABM had over 10dB difference at $\theta = 15°$ according to the scanning direction of the signal source.

In Fig. 8 (a), the sensitivity of the proposed RAMA-ABM at $\theta = 15°$ was 1 or 2dB higher than the conventional RAMA-ABM. This is because FBF1 has a flatter response than FBF0 as illustrated in Fig. 8 (b). However, the interference reduction performance of the proposed RAMA-ABM was not so high as that of the conventional RAMA-ABM. In Fig. 8 (a), output power of the proposed RAMA-ABM at $\theta \geq 30°$ is 6dB higher. Therefore, there is a trade off between the interference reduction performance and the frequency response to the target signal. An FBF with medium response between FBF0 and FBF1 will provide a compromised performance.

5. CONCLUSION

This paper has presented an adaptive microphone array using two auxiliary fixed beamfomers for good sound quality. One auxiliary fixed beamfomer is introduced in the target signal path to avoid suppression of high-frequency components in the total output. The other auxiliary fixed beamfomer is used for adaptation-mode control to eliminate the hysteresis in the relationship between signal direction and sensitivity. Both auxiliary fixed beamfomers bring about good sound quality, which improve intelligibility in speech communications and speech recognition rate. The adaptive microphone array with the proposed techniques has been implemented on a DSP system, which has demonstrated flat frequency response to the target signal and less hysteresis in its directivity pattern.

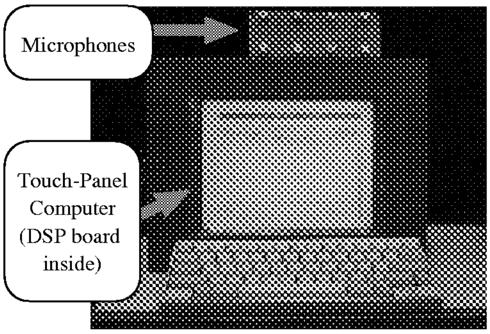

Figure 5: Realtime AMA System.

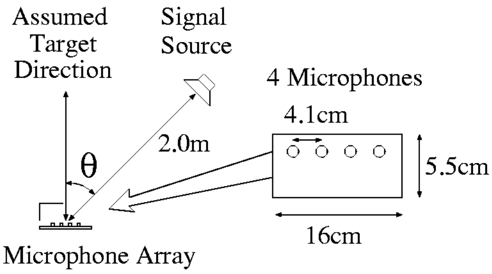

Figure 6: Experimental Set-up.

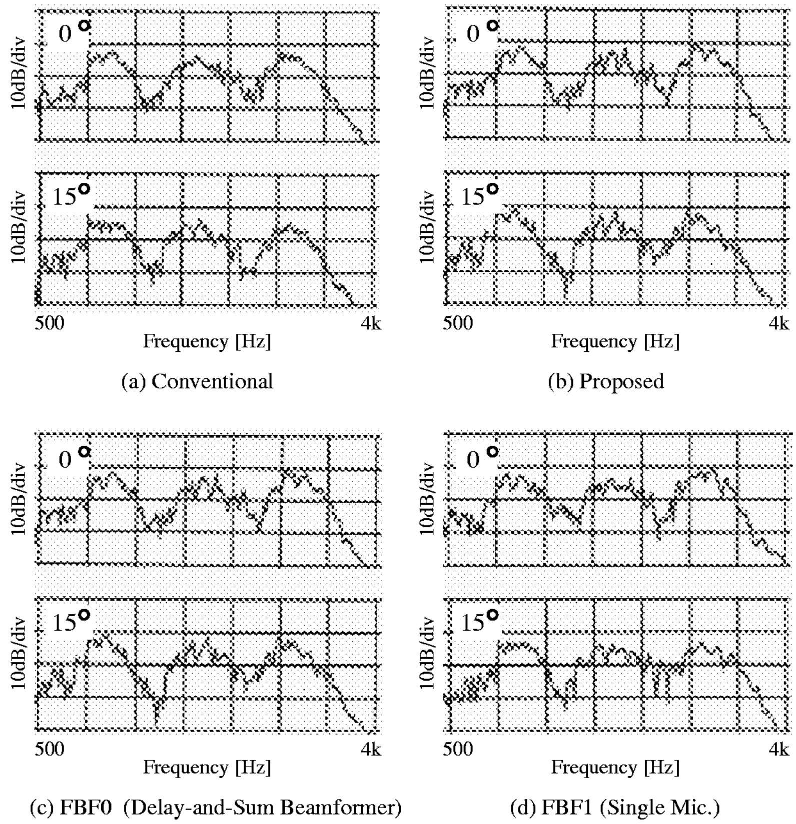

Figure 7: Spectra of Output Signals.

ACKNOWLEDGEMENT

The authors would like to thank Dr. Kazunori Ozawa, Senior Manager of C&C Media Research Laboratories, NEC Corporation for his continuous encouragement on this study.

6. REFERENCES

[1] Special Session on Microphone Array Signal Processing, in Proc. IEEE ICASSP'97, vol. I, pp. 211–254, Apr. 1997.

[2] G. L. Fudge, and D. A. Linebarger, "Steered Response Control of the Generalized Sidelobe Canceller," in Proc. IEEE ICASSP'95, pp.3623-3626, May. 1995.

[3] O. Hoshuyama, A. Sugiyama, and A. Hirano, "A Robust Adaptive Beamformer for Microphone Arrays with a Blocking Matrix Using Constrained Adaptive Filters," in Proc. IEEE ICASSP'96, pp.925–928, May 1996.

[4] O. Hoshuyama, A. Sugiyama, and A. Hirano, "A Robust Adaptive Microphone Array with Improved Spatial Selectivity and Its Evaluation in a Real Environment," in Proc. IEEE ICASSP'97, vol. I, pp.367–370, Apr. 1997.

[5] O. Hoshuyama, B. Begasse, A. Sugiyama, and A. Hirano, "A Realtime Robust Adaptive Microphone Array Controlled by an SNR Estimation," in Proc. IEEE ICASSP'98, pp.3605–3678, May 1998.

[6] L. J. Griffiths, and C. W. Jim, "An Alternative Approach to Linear Constrained Adaptive Beamforming," IEEE Trans. AP, vol. AP-30, no. 1, pp. 27–34, jan. 1982.

[7] T. Taniguchi, "Broadband Frequency Invariant Beamforming Method with Low Computational Cost," in Proc. IEEE ICASSP'98, pp. 2029–2032, May 1998.

[8] Bittware Research Systems, "User's Guide: Blacktip-PCI," Oct. 1997.

[9] Analog Devices, "ADSP-2106x SHARC User's Manual," Mar. 1995.

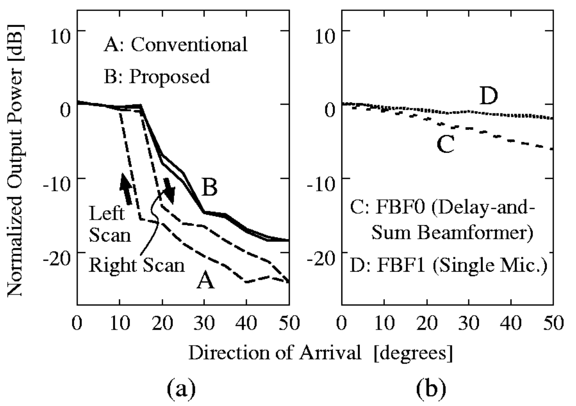

Figure 8: Outpur Powers in Real Environment.