A NEW ROBUST ADAPTATION MODE CONTROLLER BASED ON A NORMALIZED CROSS-CORRELATION WITH MULTIPLE SYMMETRIC LEAKY BLOCKING MATRICES FOR ADAPTIVE MICROPHONE ARRAYS

Thanh Phong HUA†‡, *Akihiko Sugiyama*†, *Régine LE BOUQUIN JEANNES*‡, *Gérard FAUCON*‡

†NEC Common Platform Software Research Laboratories
Kawasaki 211-8666, JAPAN
‡LTSI, Inserm U642, Université de Rennes 1,
Campus de Beaulieu, 35042 RENNES CEDEX, FRANCE

ABSTRACT

This paper proposes an adaptation mode controller (AMC) based on a normalized cross-correlation with multiple symmetric leaky blocking matrices (SLBMs) for adaptive microphone arrays. Robust and wideband estimation of a signal-to-interference ratio (SIR) is performed by the use of multiple pairs of SLBMs. The SIR estimate is compared to theoretical thresholds to evaluate target-interference proportion for controlling filter-coefficient adaptation in the adaptive path of the beamformer. A design method for SLBMs is developed with an example. Evaluations with recorded signals show that the proposed AMC achieves better discrimination between speech-dominant and interference-dominant sections, resulting in desirable control of coefficient adaptation in the microphone array. The better output signal of the microphone array, due to its good control, provides a speech recognition rate by as much as 11% higher than the conventional AMC.

1. INTRODUCTION

A robust adaptive microphone array with an adaptive blocking matrix, RAMA-ABM [1], is a generalized sidelobe canceller (GSC) [2] that has tolerance in the target-direction error. It is composed of a fixed beamformer (FBF), an adaptive blocking matrix (ABM), and a multiple input canceller (MC). Adaptation of filter coefficients in ABM should be performed only during target-periods, and alternately in MC to optimally suppress the interference and enhance the target. This is controlled by an adaptation mode controller (AMC) based on an estimated signal-to-interference ratio (SIR) [3, 4].

Greenberg and Zureck proposed an AMC based on the normalized cross-correlation (NCC) between microphone signals [3]. Assuming plane waves, the phases of the microphone signals are aligned with each other so that the mutual correlation of microphone signals is maximized for the target DOA. It means that the NCC is small for other DOAs. Consequently, the NCC can be regarded as an SIR estimate. NCC is compared with a threshold to control the adaptations. However, misdetections occur and cause breathing noise [4].

An AMC based on nested and symmetric leaky blocking matrices performs two-stage SIR estimation and makes few misdetections. A first SIR estimate is given by the power ratio of a modified FBF output to a nested blocking matrix (NBM) output. A second SIR estimate by NCC between the outputs of symmetric leaky blocking matrices (SLBMs) controls a gain which refines the first SIR estimate in low frequencies, where it is less accurate. The power ratio is not sufficiently robust to the interference DOA. It contains an FBF gain in the numerator, and an NBM gain in the denominator. These gains have different dependency on the interference DOA. In addition, the SIR estimate by NCC is bandlimited. The target may have its dominant power outside this frequency range, resulting in its misdetection and cancellation.

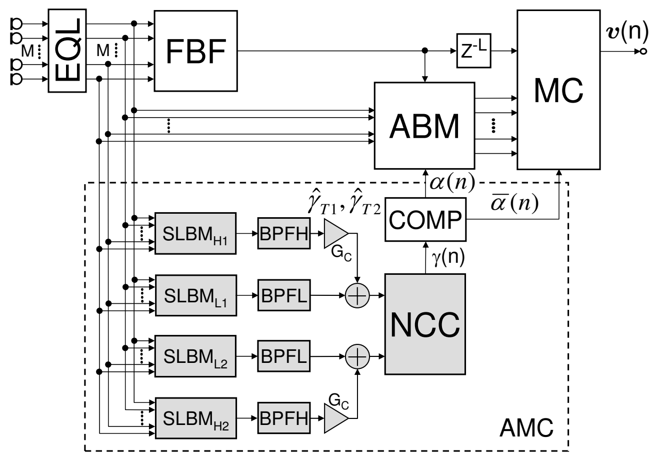

Figure 1: Structure of the proposed AMC.

This paper proposes an AMC based on a normalized cross-correlation between multiple pairs of SLBMs, which provides robustness and wideband coverage. Each pair of SLBMs operates in a distinct frequency range. In the next section, the proposed AMC is described. Section 3 compares the proposed AMC to the conventional AMC in a real acoustic environment by signal analysis and speech recognition with TV noise as interference.

2. PROPOSED ADAPTATION MODE CONTROLER

The proposed AMC contains multiple pairs of SLBMs which are optimized in multiple passbands for robust and wideband SIR estimation. An example of the proposed AMC with two SLBM pairs is depicted in Fig. 1. SLBM_{L1} and SLBM_{L2} form a pair and so do SLBM_{H1} and SLBM_{H2}. The former pair covers a low frequency range and the latter, a high frequency range. The structure of each SLBM is identical to that of the SLBMs in [5].

The input of SLBMs are automatically caliburated by an equalizer (EQL) [6] for gain imbalance among microphone signals. The outputs of SLBM_{L1} and SLBM_{H1} are combined to form one of the input signals with the sample index n,

$v_1(n)$, for NCC calculation. The other input signal, $v_2(n)$, is obtained as a combination of the output signals of SLBM$_{L2}$ and SLBM$_{H2}$. To represent the designated frequency range, the output of each SLBM is filtered by a corresponding band-pass filter (BPF) before integration. The inputs $v_1(n)$ and $v_2(n)$ of NCC are defined as a linear combination of filtered SLBM outputs in each frequency range as follows.

$$v_1(n) = z_{SLBML1}(n) + G_c \cdot z_{SLBMH1}(n), \tag{1}$$

$$v_2(n) = z_{SLBML2}(n) + G_c \cdot z_{SLBMH2}(n), \tag{2}$$

where G_c is a gain which compensates for the small power in high frequencies. $z_{SLBMx1}(n)$ and $z_{SLBMx2}(n)$ denote the filtered outputs of $SLBM_{x1}$ and $SLBM_{x2}$, respectively, with x being a suffix L or H. The filtered outputs are expressed as[5]

$$z_{SLBML1}(n) = (M-1)\,u_{M-1}(n) - g_L\sum_{i=1}^{M-2} u_i(n) - u_0(n), \tag{3}$$

$$z_{SLBML2}(n) = (M-1)\,u_0(n) - g_L\sum_{i=1}^{M-2} u_i(n) - u_{M-1}(n), \tag{4}$$

$$z_{SLBMH1}(n) = (M-1)\,u_{M-1}(n) - g_H\sum_{i=1}^{M-2} u_i(n) - u_0(n), \tag{5}$$

$$z_{SLBMH2}(n) = (M-1)\,u_0(n) - g_H\sum_{i=1}^{M-2} u_i(n) - u_{M-1}(n), \tag{6}$$

where M is the number of microphones, $u_i(n)$ the output signal of the i-th microphone. g_L and g_H are the leaky factors for $SLBM_L$ and $SLBM_H$, respectively and $g_L \neq 1$ and $g_H \neq 1$.

NCC $\gamma(n)$ for the n-th sample is calculated from $v_1(n)$ and $v_2(n)$ as follows [5]:

$$\gamma(n) = \frac{\sum_{p=0}^{N-1} v_1(n-p) \cdot v_2(n-p)}{\sqrt{\sum_{p=0}^{N-1} v_1^2(n-p) \cdot \sum_{p=0}^{N-1} v_2^2(n-p)}}. \tag{7}$$

N is the number of past samples for calculating $\gamma(n)$. $\gamma(n)$ can be considered as an estimate of the actual SIR ρ. Thus, it is compared by COMP with detection thresholds, $\hat{\gamma}_{T1}$ and $\hat{\gamma}_{T2}$, to obtain the control signal $\alpha(n)$ for filter coefficient adaptations as

$$\alpha(n) = \begin{cases} 0, & \gamma(n) \leq \hat{\gamma}_{T1}, \\ (\gamma(n) - \hat{\gamma}_{T1})/(\hat{\gamma}_{T2} - \hat{\gamma}_{T1}), & \hat{\gamma}_{T1} < \gamma(n) < \hat{\gamma}_{T2}, \\ 1, & \gamma(n) \geq \hat{\gamma}_{T2}. \end{cases} \tag{8}$$

To achieve alternate adaptations, the step sizes of coefficient adaptation in ABM are multiplied by $\alpha(n)$, whereas those in MC are multiplied by $\bar{\alpha}(n) = 1 - \alpha(n)$.

2.1 Design of SLBMs and BPFs

Parameters g_L and g_H for SLBM$_{L1}$, SLBM$_{L2}$, SLBM$_{H1}$, and SLBM$_{H2}$ are determined such that the phase difference between the paired SLBM outputs is large for the interference in the designated frequency range and equal to zero for the target. Assuming a passband of 500-1500Hz, design of SLBM$_{L1}$ and SLBM$_{L2}$ can be performed referring to [5] to obtain $g_L = 0.92$. The corresponding BPFs are desinged so that the signals outside 500-1500 Hz are sufficiently attenuated. Similarly, SLBM$_{H1}$ and SLBM$_{H2}$ are designed from viewpoints of the leaky factor g_H and the BPF passband.

The following conditions are assumed for determining the optimum values of g_H and specifications for the corresponding BPF.

Table 1: Parameter Range for SLBM$_H$ Design.

Parameter	Range	Search step
g_H	$[-2,\ 2]$	0.01
f_{ctr} [Hz]	[4000, 7400]	1
f_{Wmin} [Hz]	$[f_{ctr} - 449,\ f_{ctr} - 550]$	1
f_{Wmax} [Hz]	$[f_{ctr} + 449,\ f_{ctr} + 550]$	1
f_W [Hz]	[900, 1100]	1

- The target DOA may not be larger than the minimum interference DOA θ_{min}.
- The bandwidth f_W of the BPF passband is set to the range of $[f_{Wmin},\ f_{Wmax}]$. The center frequency of the passband is higher than f_{ctr} where speech onsets still have considerable power.

For designing g_H, an approximated NCC, $\hat{\gamma}_H(\rho,\theta)$, for the outputs of SLBM$_{H1}$ and SLBM$_{H2}$ is utilized. Assuming uncorrelated white signals with the target at 0° and the interference at an angle θ, $\hat{\gamma}_H(\rho,\theta)$ as a function of the actual SIR ρ and an interference DOA θ is shown [7] to be

$$\hat{\gamma}_H(\rho,\theta) = \frac{\sum_{i=0}^{N-1} G_H^2(i,\theta) cos(\varphi_H(i,\theta)) + \rho G_H^2(i,0)}{\sum_{i=0}^{N-1} G_H^2(i,\theta) + \rho G_H^2(i,0)}, \tag{9}$$

where $G_H(i,\theta)$ is the gain of SLBM$_{H1}$ and SLBM$_{H2}$ and $\varphi_H(i,\theta)$ is the phase difference between their output signals at the normalized frequency i. It should be noted that the gains of the paired SLBMs are identical to each other [7].

Then, similar design criteria to those in [5] can be applied.

I No overlap between $\hat{\gamma}_H(\rho = 0\,dB,\theta)$ and $\hat{\gamma}_H(\rho \to -\infty\,dB,\theta)$
This criterion ensures discrimination between interference-only situation and $0\,dB$-SIR situation in $\theta_{min} \leq \theta$.

II Centered NCC $\hat{\gamma}_H(\rho = 0\,dB,\theta)$ around 0
This criterion is derived from an assumption that $\hat{\gamma}_H(\rho \to +\infty\,dB,\theta)$ and $\hat{\gamma}_H(\rho \to -\infty\,dB,\theta)$ approaches $+1$ and -1, respectively. In such a case, it is natural that the range of $\hat{\gamma}_H(\rho = 0\,dB,\theta)$ should be centered around zero. This is because of the three typical cases in the value of $\hat{\gamma}_H(\rho,\theta)$, namely, target-only, interference-only, and mixed-signal situations. To minimize distinction errors among the three cases, $\hat{\gamma}_H(\rho,\theta)$ of each case should be equally distant from each other.

III Minimum variance of $\hat{\gamma}_H(\rho = 0\,dB,\theta)$
When the centered NCC condition is not satisfied, the variation of $\hat{\gamma}_H(\rho,\theta)$ along the DOA should be minimized instead. This criterion makes the $0\,dB$-SIR estimate robust to the DOA. A value of g_H, which gives the minimum variance of $\hat{\gamma}_H(\rho = 0\,dB,\theta)$ along the DOA axis, is selected.

2.2 Design example

Exhaustive search is performed for the parameters shown in Table 1. The sampling frequency was set to $16\,kHz$. If there is any set of parameters that satisfies the above-mentioned Criterion I in the exhaustive search, the set is stored. After the search, all stored parameter sets are evaluated with Criterion II. If there is only one set that meets Criterion II, that is the design result. When there are multiple sets, then, the set with the minimum value of $|\hat{\gamma}_{min} + \hat{\gamma}_{max}|/2$ is selected. If no set satisfies Criterion II, Criterion III is considered for them to select one set of parameters. Finally, in this example, $g_H = -0.5$, $f_{Wmin} = 3786\,Hz$, and $f_{Wmax} = 4769\,Hz$ were found to satisfy Criterion II.

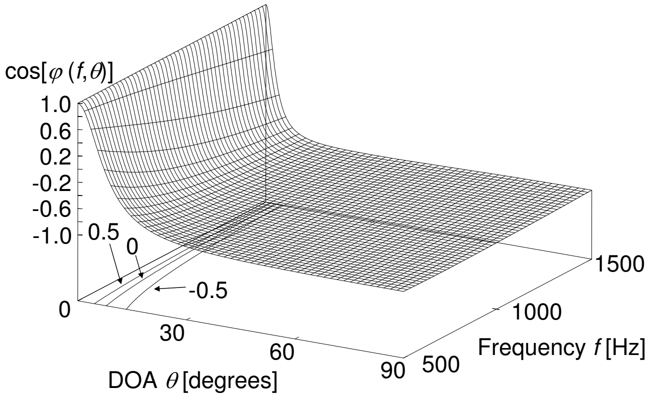

Figure 2: Cosine of the phase difference between output signals of $SLBM_1$ and $SLBM_2$ as a function of frequency and DOA with $g_L = 0.92$, $M = 4$, and $f_s = 16\,kHz$.

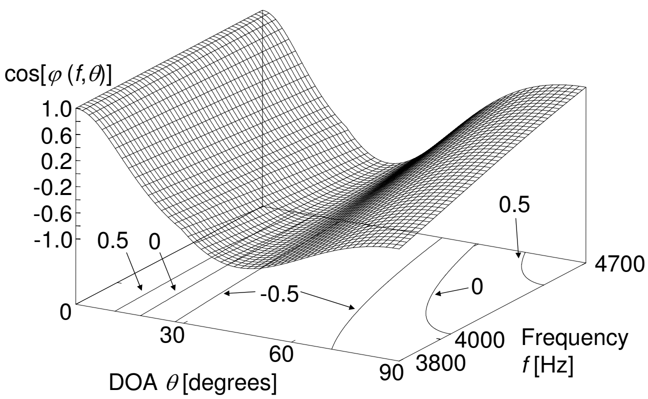

Figure 3: Cosine of the phase difference between output signals of $SLBM_{H1}$ and $SLBM_{H2}$ as a function of frequency and DOA with $g_L = -0.5$, $M = 4$, and $f_s = 16\,kHz$.

2.3 Phase and gain analysis for $\mathbf{SLBM_L}$ and $\mathbf{SLBM_H}$

Referring to [7], the transfer functions $H_{SLBMx1}(j\omega_i, \theta)$ and $H_{SLBMx2}(j\omega_i, \theta)$ of, $SLBM_{x1}$ and $SLBM_{x2}$, respectively, at a normalized frequency i and an interference DOA θ, are given by

$$\begin{aligned} &H_{SLBMx1}(j\omega_i, \theta) \\ &= (M-1)e^{-(M-1)j\omega_i t_0(\theta)} - g_x \sum_{m=1}^{M-2} e^{-jm\omega_i t_0(\theta)} - 1, \quad (10) \\ &H_{SLBMx2}(j\omega_i, \theta) \\ &= (M-1) - g_x \sum_{m=1}^{M-2} e^{-jm\omega_i t_0(\theta)} - e^{-(M-1)j\omega_i t_0(\theta)}, \quad (11) \end{aligned}$$

where $\omega_i = 2\pi i f_s/N$ is the radian frequency, f_s the sampling frequency, $j = \sqrt{-1}$, $t_0(\theta) = D\,sin\theta/c$ the relative delay between microphone signals, D the microphone spacing, and c the sound velocity. The phases $q_{x1}(i, \theta)$ and $q_{x2}(i, \theta)$ of $SLBM_{x1}$ and $SLBM_{x2}$, respectively, defined in [-π; π], are

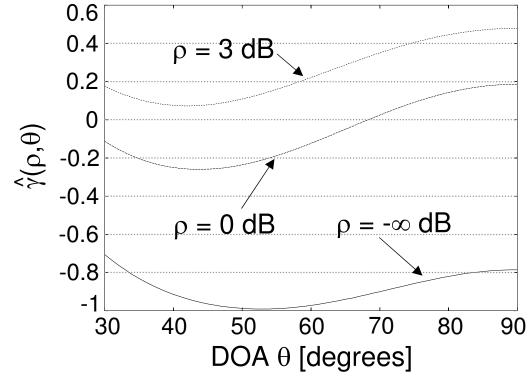

Figure 4: $\hat{\gamma}(\rho, \theta)$ versus the interference DOA θ for different SIRs.

expressed by [7]

$$q_{xp}(i, \theta) = tan^{-1}\frac{Im\left[H_{SLBMxp}(j\omega_i, \theta)\right]}{Re\left[H_{SLBMxp}(j\omega_i, \theta)\right]} + \Delta_{xp}(i, \theta), \quad (12)$$

$$\begin{aligned} \Delta_{xp}(i, \theta) \;=\; & \frac{1 - sign\left\{Re\left[H_{SLBMxp}(j\omega_i, \theta)\right]\right\}}{2}\pi \\ & \times sign\left\{Im\left[H_{SLBMxp}(j\omega_i, \theta)\right]\right\}, \quad (13) \end{aligned}$$

where $Re[\cdot]$ and $Im[\cdot]$ denote the real- and the imaginary-part operator, $sign[\cdot]$ the sign operator, and $p = 1$ or 2. The phase difference $\varphi_x(i, \theta)$ between $SLBM_{x1}$ and $SLBM_{x2}$ are defined as follows.

$$\varphi_x(i, \theta) = q_{x1}(i, \theta) - q_{x2}(i, \theta). \quad (14)$$

The cosine of the phase difference between $SLBM_{L1}$ and $SLBM_{L2}$ with $g_L = 0.92$ and the corresponding characteristics for $SLBM_{H1}$ and $SLBM_{H2}$ with $g_H = -0.5$ are depicted in Figs. 2 and 3, respectively, in the corresponding passband for $M = 4$. Both cosine functions exhibit constant responses in the designated frequency range and $\theta_{min} \leq \theta$. They provide a value close to unity in the vicinity of the look direction, *i.e.* $\theta = 0$, and a much smaller value otherwise. In the case of $SLBM_H$, the value of the cosine function increases for a large value of θ. However, it is still close to zero except high frequencies where speech does not have dominant power.

The gains $G_x(i, \theta)$ of $SLBM_{x1}$ and $SLBM_{x2}$ at the normalized frequency i and the interference DOA θ can be derived as in [7]. It is straightforward to show that they are expressed as

$$\begin{aligned} G_x(i, \theta) = \Bigg\{ & \left[(M-1)^2 + 1 + (M-2)\,g_x^2\right] \\ & + \sum_{m=1}^{M-2} \Bigg(x\left[2\,(M-2-m)\,g_x^2 + (4-2M)\,g_x\right] \\ & \qquad cos\left[\frac{2\pi\, m\, i\, f_s}{N}\, t_0(\theta)\right] \Bigg) \\ & -2(M-1)cos\left[\frac{2\pi\,(M-1)\,i\,f_s}{N}\, t_0(\theta)\right] \Bigg\}^{1/2} \quad (15) \end{aligned}$$

and identical to each other for both $SLBM_{x1}$ and $SLBM_{x2}$.

2.4 NCC between the combined SLBM outputs

From (1) and (2), the overall gain $G(i, \theta)$ between the output $v_p(n)$ and the input $u_0(n)$ is obtained as a linear combination of the SLBM gains $G_x(i,\theta)$ as

$$G(i,\theta) = G_L(i,\theta)G_{BPFL}(i) + G_c \cdot G_H(i,\theta)G_{BPFH}(i), \quad (16)$$

where $G_{BPFL}(i)$ and $G_{BPFH}(i)$ are the gains of BPFL and BPFH, respectively. Assuming that BPFL and BPFH are both ideal bandpass filters, the gains are given by

$$G_{BPFL}(i) = \begin{cases} 1, & i_{Lmin} \le i \le i_{Lmax} \\ 0, & otherwise \end{cases}, \quad (17)$$

$$G_{BPFH}(i) = \begin{cases} 1, & i_{Hmin} \le i \le i_{Hmax} \\ 0, & otherwise \end{cases}, \quad (18)$$

where i_{xmin} and i_{xmax} are cutoff frequencies of the BPFx. Therefore, the gain $G(i,\theta)$ in (16) becomes

$$G(i,\theta) = \begin{cases} G_L(i,\theta), & i_{Lmin} \le i \le i_{Lmax} \\ G_c \cdot G_H(i,\theta), & i_{Hmin} \le i \le i_{Hmax} \end{cases}. \quad (19)$$

Finally, recalling

$$\hat{\gamma}_L(\rho,\theta) = \frac{\sum_{i=0}^{N-1} G_L^2(i,\theta)cos(\varphi_L(i,\theta)) + \rho G_L^2(i,0)}{\sum_{i=0}^{N-1} G_L^2(i,\theta) + \rho G_L^2(i,0)}, \quad (20)$$

in [7], (20), (9), and (19) lead to

$$\gamma(n) \approx \hat{\gamma}(\rho,\theta) = \frac{\sum_{i=i_{Lmin}}^{i_{Lmax}} N_L(i,\theta) + G_c^2 \sum_{i=i_{Hmin}}^{i_{Hmax}} N_H(i,\theta)}{\sum_{i=i_{Lmin}}^{i_{Lmax}} D_L(i,\theta) + G_c^2 \sum_{i=i_{Hmin}}^{i_{Hmax}} D_H(i,\theta)}, (21)$$

where

$$N_x(i,\theta) = G_x^2(i,\theta)cos\left[\varphi_x(i,\theta)\right] + \rho G_x^2(i,0), \quad (22)$$
$$D_x(i,\theta) = G_x^2(i,\theta) + \rho G_x^2(i,0). \quad (23)$$

Now, the relation between $\hat{\gamma}(\rho,\theta)$ and the actual SIR ρ has been developed. $\gamma(n)$ is an estimate of $\hat{\gamma}(\rho,\theta)$, and therefore provides an estimate of the actual SIR ρ depending on the interference DOA θ. $\hat{\gamma}(\rho,\theta)$ in (21) as a function of the interference DOA and different SIRs is depicted in Fig. 4.

3. EVALUATIONS

A uniform linear array of four microphones was placed in the middle of a reverberant room of $5 \times 5 \times 3$ m in width, depth, and height, to acquire the data. The target speech comprising child, female, or male voice and TV-commercial as the interference including advertisement, music, ambient noise was presented by loudspeakers. They were recorded separately so that they can be mixed for arbitrary SIRs.

3.1 Parameters

5-th order elliptic filters were used as BPFL and BPFH. The passband of BPF for AMC-NCC and BPFL for AMC-DNCC were set to $500 - 1500\,Hz$, whereas the BPFH passband was set to $3800 - 4700\,Hz$. $\hat{\gamma}_{T1}$ and $\hat{\gamma}_{T2}$ were set to the theoretical minimum and the maximum values of $\hat{\gamma}\,(\rho = 0dB, \theta)$ in (21), respectively (See Fig. 4). The gain G_C was selected such that $G_C^2 = V_{HFE}$, and $f_s = 16\,kHz$. The parameter values are summarized in Table 2.

3.2 Quality of the microphone array output

Improved output quality by AMC-DNCC is shown in Figs. 5 (e) and Fig. 6 (e), where the interference was located 1.0 m away at 30° with an average SIR of 0 dB, and at 90° with an average SIR of 10 dB , respectively. Whenever a

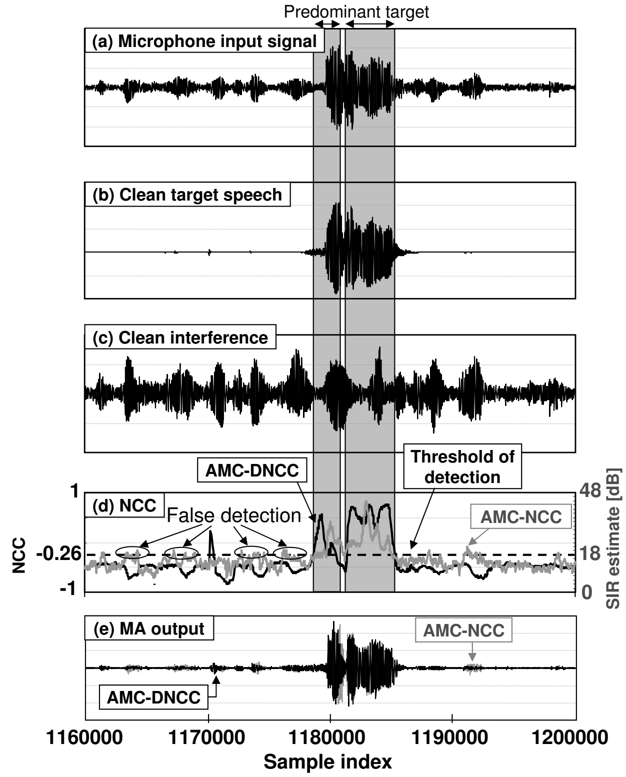

Figure 5: Output of the microphone array using a target speech at 0°, and an interference DOA of 30° and an average SIR of 0 dB.

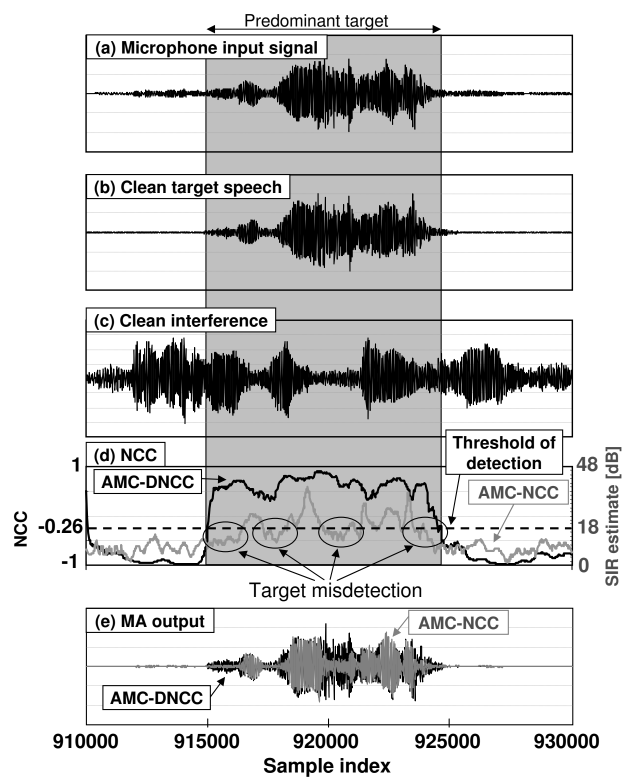

Figure 6: Output of the microphone array using a target speech at 0°, and an interference DOA of 90° and an average SIR of 10 dB.

Table 2: Parameter values for the evaluations

AMC-NCC		AMC-DNCC	
Parameter	Value	Parameter	Value
g_L	0.92	g_L, g_H	0.92, -0.50
σ_1, σ_2 [dB]	18, 21	$\hat{\gamma}_{T1}$, $\hat{\gamma}_{T2}$	-0.26, 0.20
N, K	256, 128	N	512
BPF [Hz]	500 - 1500	BPFL [Hz]	500 - 1500
δ	70	BPFH [Hz]	3800 - 4700
V_{HFE}	9	G_c	3
$\hat{\gamma}_T$	-0.95	C_p	14.5
T_p	2/3		

burst of interference occurs, a sudden decrease in NCC can be observed for AMC-DNCC. On the contrary, this is not the case for the SIR estimate by AMC-NCC. This difference leads to a higher interference suppression in the proposed method as in Fig. 5 (d). The word in the target in Fig. 6 is "suzushii (cool)" in Japanese that contains onsets in the beginning and the middle of the word, at sample indexes around 916000 and 920000. These onsets are well detected when the SIR is larger than $0\,dB$. Consequently, the onset is better preserved in the output signal of the microphone array with AMC-DNCC as in Fig. 6 (e).

3.3 Speech recognition rates

The AMCs were also evaluated in a speech recognition scenario. The target, located at $0.5\,m$ or $1.5\,m$ away, was composed of 30 Japanese speakers (10 males, 10 females and 10 children) with 50 utterances for each speaker. The interference was placed at $1.0\,m$ in 3 different DOAs: 30°, 60°, and 90°. Signals with average SIRs of 0, 5, 10, and $15\,dB$ as well as clean speech (CS) signals formed 780 signals processed by RAMA-ABM with AMC-NCC, and that with AMC-DNCC. Their outputs were evaluated by a Japanese speech recognition system, *Julius* [8]. The results are shown in Figs. 7 and 8.

Compared to AMC-NCC, AMC-DNCC always achieves a better recognition rate with an increase of up to 11% and an average increase of 4%. The improvement by the proposed method is more noticeable in low SIRs such as $0\,dB$, where the average increase in the recognition rate is 9.5%. This was expected because the robustness to the interference DOA has been improved in low SIRs. Compared to the single-microphone case, the recognition rate was increased by up to 52%, except for the clean speech sources with approximately 1% degradation. This degradation was caused by target cancellation at high SIRs due to reverberations [3].

4. CONCLUSION

An adaptation mode controller (AMC) based on a normalized cross-correlation with multiple symmetric leaky blocking matrices (SLBMs) for adaptive microphone arrays has been proposed. Multiple pairs of symmetric leaky blocking matrices have been introduced for robust and wideband estimation of a signal-to-interference ratio (SIR). A design method for SLBMs has been developed and a design example has been demonstrated. Evaluations with recorded signals have shown that the proposed AMC achieves better discrimination between speech-dominant and interference-dominant sections, resulting in desirable control of coefficient adaptation in the microphone array. The better output signal of the microphone array, due to its good control, provides a higher speech recognition rate by as much as 11% than the conventional AMC.

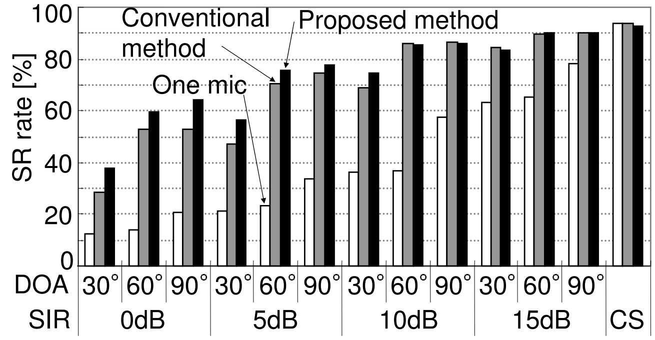

Figure 7: Speech recognition rates with the target at 0.5 m away.

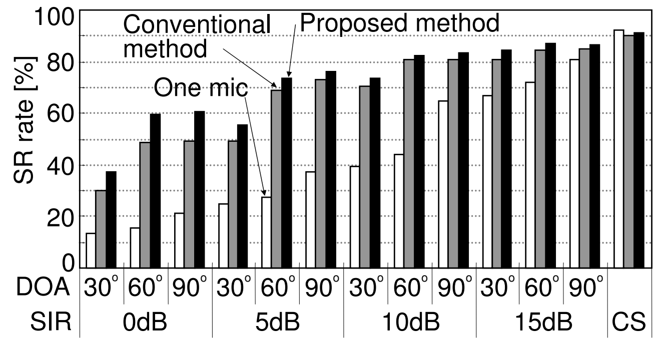

Figure 8: Speech recognition rates with the target at 1.5 m away.

REFERENCES

[1] O. Hoshuyama and A. Sugiyama, "Robust adaptive beamforming," *Microphone Arrays*, Chap. 5, Brandstein and Ward, ed. Springer, 2001.

[2] 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.

[3] J. Greenberg and P. Zureck, "Evaluation of an adaptation beamforming method for hearing aids," *J.A.S.A*, vol.91, no.3, pp.1662–1676, Mar. 1992.

[4] O. Hoshuyama, B. Begasse, A. Sugiyama, A. Hirano "A realtime robust adaptive microphone array controlled by an SNR estimate," Proc. of ICASSP98, pp. 3605–3608, 1998.

[5] T. P. Hua, A. Sugiyama, R. Le Bouquin Jeannes, G. Faucon, "A new adaptation mode controller for adaptive microphone arrays based on nested and symmetric leaky blocking matrices," Proc. of EUSIPCO2006, Sep. 2006.

[6] T. P. Hua, A. Sugiyama, G. Faucon, "A new self-calibration technique for adaptive microphone arrays," Proc. of IWAENC2005, pp. 237–240, Sep. 2005.

[7] T. P. Hua, A. Sugiyama, R. Le Bouquin Jeannes, G. Faucon, "Estimation of the signal-to-interference ratio based on normalized cross-correlation with symmetric leaky blocking matrices in adaptive microphone arrays," *IEICE Tech. Rep.*, vol. 106., No. 21, pp. 61–66, Apr. 2006.

[8] A. Lee, T. Kawahara and K. Shikano, "Julius – An open source real-time large vocabulary recognition engine," Proc. EUROSPEECH2001, pp. 1691–1694, Sep. 2001.