ESTIMATION OF THE SIGNAL-TO-INTERFERENCE RATIO BASED ON NORMALIZED CROSS-CORRELATION WITH SYMMETRIC LEAKY BLOCKING MATRICES IN ADAPTIVE MICROPHONE ARRAYS

[1]Thanh Phong HUA,[1]Akihiko SUGIYAMA,[2]Régine LE BOUQUIN JEANNES and [2]Gérard FAUCON
[1]aks@ak.jp.nec.com
[1]NEC Media and Information Research Laboratories Kawasaki 211-8666, Japan
[2]LTSI, Inserm U642, Université de Rennes 1, 35042 RENNES CEDEX, FRANCE

ABSTRACT

This paper proposes estimation of the signal-to-interference ratio (SIR) based on normalized cross-correlation (NCC) with symmetric leaky blocking matrices for adaptive microphone arrays. An analysis of the NCC shows that it is a function of the actual SIR and the phase difference between its inputs. This function provides its optimum SIR estimation when the phase difference is maximized for the interference and minimized for the target. Symmetric leaky blocking matrices are designed and optimized to approximately satisfy this condition. The actual and the theoretical NCCs show good agreement with each other at different SIRs using white signals. The proposed SIR estimation is insensitive to the threshold for target detection for a wide range of input signals with different directions of arrival. Simulation results show that the NCC is close to 1 for the target and -1 for the interference, representing a nearly optimum behavior.

1. INTRODUCTION

Adaptive microphone arrays are useful to capture a signal from a specific direction of arrival (DOA) and to suppress directional interferences [1]-[5]. To work properly, they require an adaptation mode controller (AMC) to adapt their filter coefficients [4, 5]. Based on an estimation of signal-to-interference ratio (SIR), target and interference sections are detected to apply different adaptations. The accuracy of the SIR estimate is crucial for correct adaptations, and therefore, for the quality of the output signal. When a misadaptation occurs, the interference may not be sufficiently suppressed and the target may be partially cancelled.

Greenberg and Zureck proposed SIR estimation by normalized cross-correlation (NCC) between adjacent-microphone signals [5]. The NCC reflects the similarities between the amplitudes and phases of the input signals. The amplitudes of the microphone signals are the same assuming that each microphone signal is a delayed version of another. Thus, the NCC reflects only the phase difference. Knowing the target signal direction, the phase can be aligned to obtain the maximum correlation in case of target signal only. When the interference alone arrives from a DOA that is sufficiently different from the target direction, the NCC is smaller. Consequently, the NCC can be used as an SIR estimate. However, the SIR estimate becomes inaccurate by weak distinction of the target from the interference in low frequencies, where the phase difference between microphone signals is almost independent of the DOA.

This paper proposes an SIR estimation based on an NCC between the outputs of symmetric leaky blocking matrices. Due to the symmetric structure, a good distinction of the target from the interference is achieved. In the next section, the relation between the NCC and the actual SIR for white signals is studied. The phase difference between the input signals of NCC is analyzed in Section 3 to optimize discrimination performance between the target and the interference. In Section 4, evaluation results confirm the analysis and show good target-detection performance of the proposed SIR estimation compared to the conventional one.

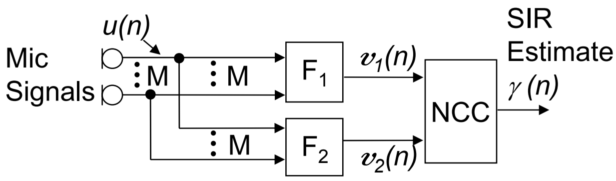

Figure 1: SIR estimation by normalized cross-correlation of two filtered signals $v_1(n)$ and $v_2(n)$.

2. SIR ESTIMATE BY AN NCC

Figure 1 presents the general structure of SIR estimation by an NCC. The target direction is assumed to be perpendicular to the microphone array surface. The NCC is calculated between the outputs $v_1(n)$ and $v_2(n)$ of the filters F_1 and F_2 with the same gain but different phase except for the target DOA. The value $\gamma(n)$ of the NCC at sample n is given by

$$\gamma(n) = \frac{\sum_{k=0}^{N-1} v_1(n-k) \cdot v_2(n-k)}{\sqrt{\sum_{k=0}^{N-1} v_1^2(n-k)}\sqrt{\sum_{k=0}^{N-1} v_2^2(n-k)}}, \qquad (1)$$

with N past samples of $v_1(n)$ and $v_2(n)$.

Assuming the target and the interference are short-time stationary and white signals, it is shown in [6] that $\gamma(n)$ is approximated by $\hat{\gamma}(\rho, \theta)$ as

$$\hat{\gamma}(\rho,\theta) = \frac{\sum_{i=0}^{N-1} \left\{G^2(i,\theta)cos\left[\varphi(i,\theta)\right] + \rho G^2(i,0)\right\}}{\sum_{i=0}^{N-1} \left\{G^2(i,\theta) + \rho G^2(i,0)\right\}}, \qquad (2)$$

where ρ is the actual SIR, defined by $\rho = T^2(i)/I^2(i)$. $T^2(i)$ and $I^2(i)$ are the magnitude of the target and the interference at the i-th frequency bin. $G(i,\theta)$ is the gain of F_2 and F_1 at a DOA θ and $\varphi(i,\theta)$ is the phase difference between F_1 and F_2 as

$$\varphi(i,\theta) = q_1(i,\theta) - q_2(i,\theta), \qquad (3)$$

where $q_1(i,\theta)$ and $q_2(i,\theta)$ are the phases introduced by F_1 and F_2, respectively. It should be noted that (2) is the general expression of the SIR estimate based on an NCC. In the conventional method [5], F_1 and F_2 pass one microphone signal and the other.

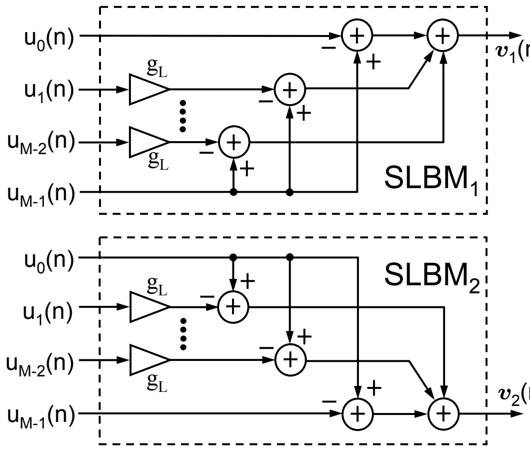

Figure 2: Structure of symmetric leaky blocking matrices.

Thus, their gains are unity. The phase difference between these signals corresponds to the phase shift of the signals arriving at the microphones. This phase shift is small in low frequencies where the wavelength is too large for the array size. It causes insufficient distinction between the target and the interference.

3. PROPOSED SIR ESTIMATION

The proposed SIR estimation introduces symmetric leaky blocking matrices (SLBMs), as depicted in Fig. 2, as F_1 and F_2. SLBMs have a similar structure to that of the nested blocking matrix [7], but characterized by a leaky factor g_L.

3.1. Definition of the symmetric leaky blocking matrices

The outputs $v_1(n)$ and $v_2(n)$ of SLBM$_1$ and SLBM$_2$, respectively, at sample n are defined by

$$v_1(n) = (M-1)\,u_{M-1}(n) - g_L \sum_{m=1}^{M-2} u_m(n) - u_0(n), \quad (4)$$

$$v_2(n) = (M-1)\,u_0(n) - g_L \sum_{m=1}^{M-2} u_m(n) - u_{M-1}(n), \quad (5)$$

where $u_m(n)$ is the m-th microphone signal and g_L is a leaky factor satisfying $g_L \neq 1$. The symmetric structures of SLBM$_1$ and SLBM$_2$ create signals with large phase difference in low frequencies for the interference. If $g_L = 1$, they become pure blocking matrices for the target, resulting in $v_1(n) = v_2(n) = 0$. The essential information about the target is therefore missing in the output. The constraint $g_L \neq 1$ allows leakage of the in-phase target in the correlated outputs.

3.2. Gain analysis for SLBM$_1$ and SLBM$_2$

With the plain-wave assumption, the time delay $t_0(\theta)$ between adjacent-microphone signals depends on the source DOA θ as

$$t_0(\theta) = \frac{D\,sin\theta}{c}, \quad (6)$$

where D and c are the microphone spacing and the sound velocity. Therefore, one reference signal can be viewed as one and only input of the blocking matrix. Let us choose $u_0(n)$ as the reference signal. Other signals are its shifted versions by an integer multiple of $t_0(\theta)$. The transfer functions $\mathrm{H}_{\mathrm{SLBM1}}(j\omega_i, \theta)$ and $\mathrm{H}_{\mathrm{SLBM2}}(j\omega_i, \theta)$ of the blocking matrices SLBM$_1$ and SLBM$_2$, respectively, can be expressed with respect to the input $u_0(n)$ as

$$H_{\mathrm{SLBM1}}(j\omega_i, \theta) = (M-1)e^{-(M-1)j\omega_i t_0(\theta)} - g_L \sum_{m=1}^{M-2} e^{-jm\omega_i t_0(\theta)} - 1, \quad (7)$$

$$H_{\mathrm{SLBM2}}(j\omega_i, \theta) = (M-1) - g_L \sum_{m=1}^{M-2} e^{-jm\omega_i t_0(\theta)} - e^{-(M-1)j\omega_i t_0(\theta)}, \quad (8)$$

where f_s is the sampling frequency, $\omega_i = 2\pi i f_s/N$, and $j = \sqrt{-1}$. Knowing that the gain is the norm of the transfer function, $G_{SLBM1}(i,\theta)$ and $G_{SLBM2}(i,\theta)$ are expressed as

$$G_{SLBM1}(i,\theta) = \left\{\left[(M-1)e^{-(M-1)j\omega_i t_0(\theta)} - g_L \sum_{m=1}^{M-2} e^{-jm\omega_i t_0(\theta)} - 1\right] \times \left[(M-1)e^{(M-1)j\omega_i t_0(\theta)} - g_L \sum_{m=1}^{M-2} e^{+jm\omega_i t_0(\theta)} - 1\right]\right\}^{1/2}. \quad (9)$$

$$G_{SLBM2}(i,\theta) = \left\{\left[(M-1) - g_L \sum_{m=1}^{M-2} e^{-jm\omega_i t_0(\theta)} - e^{-(M-1)j\omega_i t_0(\theta)}\right] \times \left[(M-1) - g_L \sum_{m=1}^{M-2} e^{+jm\omega_i t_0(\theta)} - e^{+(M-1)j\omega_i t_0(\theta)}\right]\right\}^{1/2}. \quad (10)$$

They both lead to the same gain $G(i,\theta)$ expressed by (11) in the following page. As a consequence, the gains of SLBM$_1$ and SLBM$_2$ are equal.

3.3. Phase analysis for SLBM$_1$ and SLBM$_2$

The phases $q_1(i,\theta)$ and $q_2(i,\theta)$ of the transfer functions $\mathrm{H}_{\mathrm{SLBM1}}(j\omega_i, \theta)$ and $\mathrm{H}_{\mathrm{SLBM2}}(j\omega_i, \theta)$, respectively, are expressed by

$$q_p(i,\theta) = arctan\frac{Im\left[H_{\mathrm{SLBMp}}(j\omega_i,\theta)\right]}{Re\left[H_{\mathrm{SLBMp}}(j\omega_i,\theta)\right]} + \Delta_p(i,\theta), \quad (12)$$

$$\Delta_p(i,\theta) = \frac{1 - sign\left\{Re\left[H_{\mathrm{SLBMp}}(j\omega_i,\theta)\right]\right\}}{2}\pi \times sign\left\{Im\left[H_{\mathrm{SLBMp}}(j\omega_i,\theta)\right]\right\}, \quad (13)$$

where $Re[\cdot]$ denotes the real-part operator, $Im[\cdot]$ the imaginary-part operator, $sign[\cdot]$ the sign operator, and $p = 1$ or 2. An example of the phase difference $\varphi(i,\theta)$ defined in (3) is shown in Fig. 3 for $M = 4$ and $g_L = 0.92$. A sharp transition in directivity of $\varphi(i,\theta)$ in the low frequencies is achieved.

3.4. Optimum leaky factor for phase-difference control

g_L controls the shape of $\hat{\gamma}(\rho,\theta)$ through the gains and the phases of SLBM$_1$ and SLBM$_2$. There are two criteria for the choice of g_L, namely, minimum variance of $\hat{\gamma}(\rho,\theta)$ along the interference-DOA axis and no intersection between $\hat{\gamma}(\rho = 0dB,\theta)$ and $\hat{\gamma}(\rho =$

$$G(i,\theta) = \left\{ \left[(M-1)^2 + 1 + (M-2)g_L^2\right] + \sum_{m=1}^{M-2} \left(2g_L\left[(M-2-m)g_L + (2-M)\right]\cos\left[\frac{2\pi\, m\, i\, f_s}{N} t_0(\theta)\right]\right) - 2(M-1)\cos\left[\frac{2\pi\,(M-1)\, i\, f_s}{N}\, t_0(\theta)\right]\right\}^{1/2}. \quad (11)$$

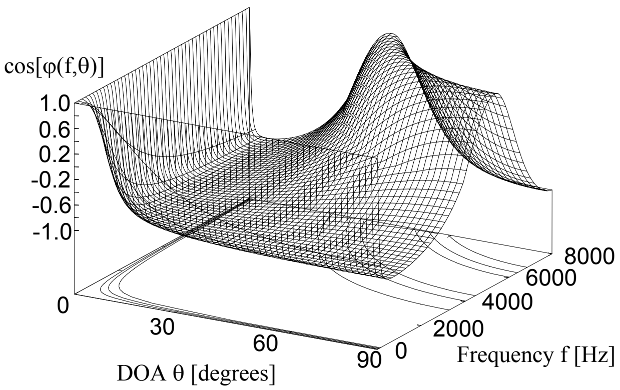

Figure 3: Cosine of the phase difference between output signals of $SLBM_1$ and $SLBM_2$ as a function of the frequency and the DOA for $g_L = 0.92$ and using 4 microphones for a sampling frequency of 16 kHz.

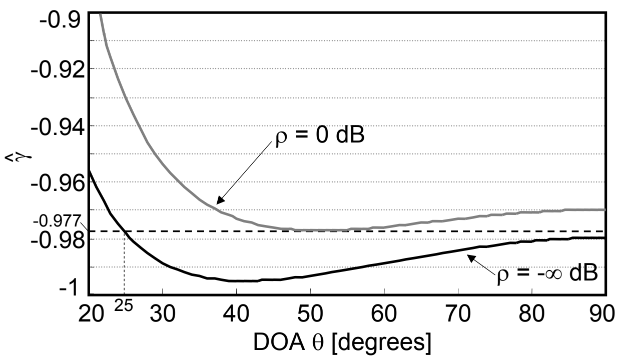

Figure 4: NCC vs. the interference DOA with $g_L = 0.92$ and $M = 4$ for the passband 500-1500 Hz assuming white signals.

$-\infty dB, \theta)$. These conditions are essential for accurate and consistent SIR estimation independent of the interference DOA greater than its minimum, θ_{min}. An optimum value of $g_L = 0.92$ was found for $500-1500$ Hz, where speech has significant power, by exhaustive search in the range $[-1,\ 3]$ with a resolution of 0.01. The corresponding NCC values versus the DOA at $\rho = 0$ and $-\infty$ dB are depicted in Fig. 4, where the criteria are satisfied for interference DOAs above 25°.

4. EVALUATIONS

A uniform linear array of 4 microphones was used with a 5-th order elliptic filters for bandlimiting the NCC input to the passband $500-1500$Hz. A target-detection threshold, $\hat{\gamma}_T$, was calculated as the theoretical maximum value of $\hat{\gamma}(\rho = 0dB, \theta)$ and set to -0.953. Whenever $\hat{\gamma}(0dB, \theta) > \hat{\gamma}_T$, the SIR is likely to be higher than 0dB. Parameters are summarized in Table 1.

Table 1: Parameter values for the evaluations

Parameter	Value	Parameter	Value
f_s	16 kHz	g_L	0.92
Passband	$500-1500$ Hz	γ_T	-0.953
N	256	c	340 m/s
θ_{min}	30°	D	2.1 cm

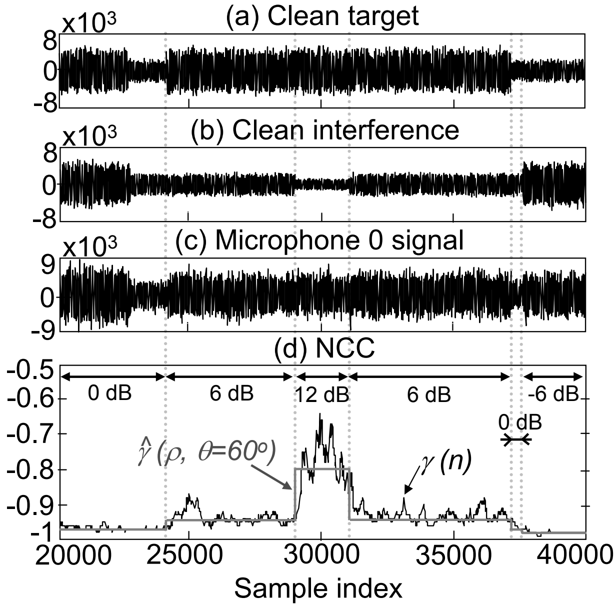

Figure 5: $\gamma(n)$ obtained with uncorrelated white signals at 0°, as the target signal, and at 60°, as the interference, for $g_L = 0.92$.

4.1. Validation of the theoretical NCC, $\hat{\gamma}(\rho, \theta)$

An approximated NCC $\hat{\gamma}(\rho, \theta)$ in (2) was evaluated with uncorrelated white target and interference signals at 0° and at 60°, respectively. Figure 5 (a) and (b) present these white signals with sudden changes in magnitude to generate SIRs of -6, 0, 6, and 12dB. Shown in Fig. 5 (c) is the signal at microphone 0. Figure 5 (d) compares the approximated NCC $\hat{\gamma}(\rho, \theta = 60^\circ)$ with the calculated NCC $\gamma(n)$. Validity of $\hat{\gamma}(\rho, \theta)$ is confirmed by their good agreement.

4.2. Comparison of the conventional and the proposed NCCs

$\gamma(n)$ by the conventional [5] and the proposed were compared using a female speech as the target and a TV noise as the interference with an average SIR of 0dB. Figures 6 and 7 depict, from the top to the bottom, the signal at microphone 0, clean target, pure interference, the conventional NCC [5], and the proposed NCC, representing an interference DOA of 30° and 60°, respectively. Sections highlighted in gray show target-dominant sections. Figures 6 (e) and 7 (e) confirm that the proposed NCC exhibits an optimum behavior, *i.e.* close to 1 during target-dominant sections and to -1 during interference-dominant sections. On the contrary, the conventional NCC is distributed in a narrow range close to 1. The narrow range needs high resolution for target-interference discrimination, leading to a sensitive threshold. This sensitivity is likely to cause target cancellation and/or insufficient interference suppression. Moreover, the conventional NCC cannot use a common and constant threshold as is clear from Fig. 6 (d) and Fig. 7 (d). In case of the NCC with SLBMs, when the actual SIR is under 0 dB, $\gamma(n)$ is under the theoretical threshold $\hat{\gamma}_T$ (dotted line) and vice-versa. As a result, the NCC by the proposed method is superior to that by the conventional one for target detection.

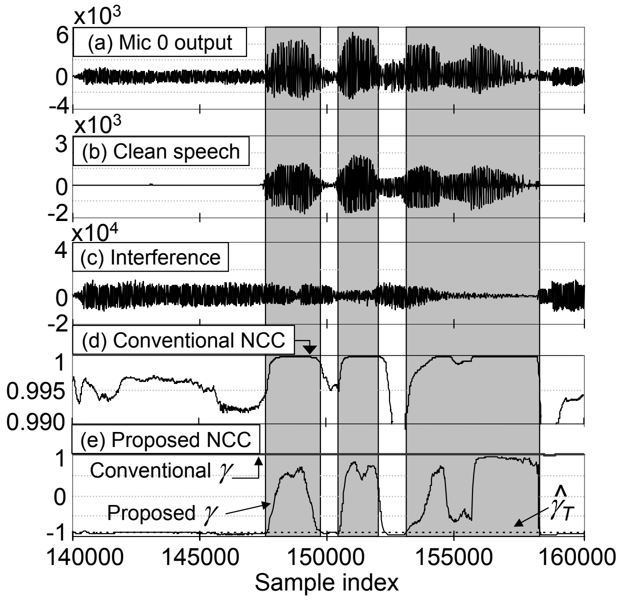

Figure 6: NCC method for a simulated input signal using a target speech at 0°, and an interference DOA of 30°.

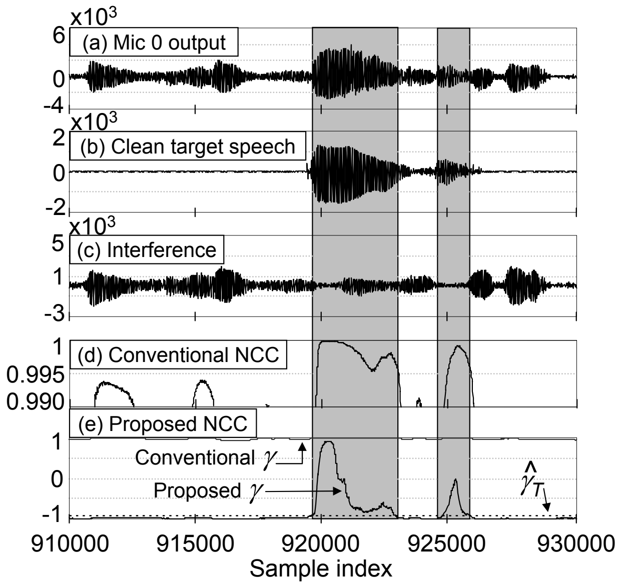

Figure 7: NCC method for a simulated input signal using a target speech at 0°, and an interference DOA of 60°.

5. CONCLUSION

Estimation of the signal-to-interference ratio (SIR) based on normalized cross-correlation (NCC) with symmetric leaky blocking matrices has been proposed for adaptive microphone arrays. NCC has been analyzed to show that it best approximates the SIR when the input phases are opposite for the interference and identical for the target. To satisfy this condition, symmetric leaky blocking matrices have been developed and optimized. The actual and the theoretical NCCs have shown good agreement with each other at different SIRs for white signals. Simulation results have shown that the estimated SIR exhibits a nearly optimum behavior, *i.e.* the NCC close to 1 for the target and -1 for the interference.

6. REFERENCES

[1] 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.

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

[3] O. Hoshuyama, B. Begasse, A. Sugiyama, A. Hirano "A realtime robust adaptive microphone array controlled by an SNR estimate," Proc. of ICASSP'98, pp. 3605–3608, 1998.

[4] O. Hoshuyama and A. Sugiyama, "An adaptive microphone array with good sound quality using auxiliary fixed beamformers and its single-chip implementation," Proc. of ICASSP'99, pp. 949–952, Mar. 1999.

[5] 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.

[6] 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.

[7] 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," to appear in Proc. of EUSIPCO2006, Sep. 2006.