Directivity characteristics of adaptive microphone-array for noise reduction (AMNOR)

Yutaka Kaneda

*NTT Human Interface Laboratories,
3-9-11 Midori-cho, Musashino, Tokyo, 180 Japan*

(Received 27 February 1991)

[English translation of the same article in *J. Acoust. Soc. Jpn. (J)* **44, 23-30 (1988)]**

AMNOR is a new type of noise-reduction microphone system which uses a microphone-array and digital signal processing. In this paper, the theory of AMNOR is described in the frequency domain, and its directivity characteristics are studied. Then, it is shown that, under some simple noise conditions, AMNOR forms the same directivity patterns as those formed by conventional directional microphones. In contrast to the directivity characteristics of conventional directional microphones which are assumed to be optimal for some restricted noise environments, AMNOR can realize the optimal directivity characteristics adaptively for a wide variety of existing noise environments. This is accomplished in AMNOR by the following three functions; (1) AMNOR forms $M-1$ (where M is the number of microphone elements) dead angles with respect to the noise arrival directions, (2) AMNOR optimizes the directivity lobe figure (3) AMNOR performs directivity control of (1) and (2) for each frequency band independently.

Keywords: Microphone array, Noise reduction, AMNOR, Directivity, Directional microphone

PACS Number: 43.88.Hz

1. INTRODUCTION

A technique for selectively receiving a specific sound in a noisy environment is strongly desired. Such a technique is required for hands-free telephones, teleconference systems and speech-understanding systems. These systems are becoming popular, but environmental noise seriously degrades their quality (This paper refers to all kinds of undesired sound—such as machine noise, TV sound and undesired speech—as “noise”).

Various types of directional microphones have been used for selective sound receiving. Super-directional microphone arrays have also been investigated recently.^{1,2)} Conventional directional microphones, however, do not sufficiently improve the signal-to-noise ratio (SNR), and super-directional microphone arrays require impractically

large microphone arrays for many applications.

An alternative technique called AMNOR (Adaptive Microphone Array for Noise Reduction) has recently been proposed.^{3,4)} In contrast to conventional techniques, AMNOR achieves a high SNR with a small microphone array. AMNOR comprises a microphone array, variable digital filters, and a filter controller. The characteristics of these filters are automatically optimized for noise conditions like arrival directions and power levels. Since the directivity of AMNOR is determined by these filter characteristics, optimal directivity is achieved by the optimal control of these filters.

The previous studies have presented methods for deriving the optimal filter characteristics, but their physical meaning remains unclear. That is, the relationship between the directivity characteristics and noise conditions has not been treated in depth.

This paper first presents the AMNOR principle for the frequency domain, in which the relationship between the AMNOR directivity and noise conditions is easily perceived. Based on this frequency domain representation, AMNOR directivity in some typical noise environments is compared with that of conventional directional microphones. Finally, the features of AMNOR's directivity characteristics are discussed.

2. AMNOR PRINCIPLE IN THE FREQUENCY DOMAIN

2.1 Problem Formulation for Noise Reduction

Figure 1 is a block diagram of the AMNOR sound receiving system in a sound field where a desired sound source coexists with noise sources. Here $S(\omega)$ represents a desired signal such as speech, and $N_1(\omega), \dots, N_N(\omega)$ represent N undesired signals that are hereafter called "noise signals" or simply "noise." Furthermore, $S(\omega), N_1(\omega), \dots, N_N(\omega)$ are assumed to be mutually independent stochastic signals. The term $G_i(\omega)$ ($i=1, \dots, M$) represents the transfer function between the desired sound source and the i -th microphone, and $G_{ji}(\omega)$ ($j=1, \dots, N; i=1, \dots, M$) represents the transfer function between the j -th noise source and the i -th microphone.

In the AMNOR system, sound is first received by the microphone array consisting of M microphone elements. Then the received signals $U_1(\omega), \dots, U_M(\omega)$ are applied to filters that have characteristics of $H_1(\omega), \dots, H_M(\omega)$. Finally the filter outputs are summed to produced the AMNOR output signal $Y(\omega)$.

The received signals $U_1(\omega), \dots, U_M(\omega)$ can be represented using a matrix formula as follows.

$$U = GV \quad (1)$$

Where U and V are column vectors of respective orders M and $N+1$, and G is an $M \times (N+1)$ matrix:

$$U \triangleq [U_1(\omega), U_2(\omega), \dots, U_M(\omega)]^T \quad (2)$$

$$V \triangleq [S(\omega), N_1(\omega), N_2(\omega), \dots, N_N(\omega)]^T \quad (3)$$

$$G \triangleq \begin{pmatrix} G_1(\omega) & G_{11}(\omega) & \dots & G_{N1}(\omega) \\ G_2(\omega) & G_{12}(\omega) & \dots & G_{N2}(\omega) \\ \vdots & \vdots & \ddots & \vdots \\ G_M(\omega) & G_{1M}(\omega) & \dots & G_{NM}(\omega) \end{pmatrix} \quad (4)$$

T : transpose

An M -th order row vector H whose elements are $H_1(\omega), \dots, H_M(\omega)$ is then defined as

$$H \triangleq [H_1(\omega), H_2(\omega), \dots, H_M(\omega)] \quad (5)$$

And the output $Y(\omega)$ is represent by the product of the row vector H and the column vector U .

$$Y(\omega) = HU \quad (6)$$

Because noise components in the received signals $U_1(\omega), \dots, U_M(\omega)$ are reduced by applying the filters $H_1(\omega), \dots, H_M(\omega)$, the vector H is called the noise reduction filter.

The noise reduction filter H completely determines the directivity characteristics of AMNOR for a given number and arrangement of microphones. Optimal determination of the noise reduction filter H is thus the main problem in AMNOR.

2.2 Optimal Noise Reduction Filter

Considering the frequency response of AMNOR in the desired direction, $F(\omega)$, the deviation D_1 of $F(\omega)$ from a flat response is defined by the following equation.

$$D_1 \triangleq \frac{1}{2\omega_0} \int_{-\omega_0}^{\omega_0} |1 - F(\omega)|^2 d\omega, \quad (7)$$

Fig. 1 Block diagram of AMNOR (Advantage Microphone-array for Noise Reduction).

where ω_0 denotes the highest frequency of the objective band. In Eq. (7), D_1 is derived by square integration of the deviation of $F(\omega)$ from the flat response, which is represented by the value 1. Hereafter D_1 is called degradation. The noise power observed in the AMNOR output $Y(\omega)$ is now denoted by D_2 . Then, there is a trade-off relationship between the degradation D_1 and the output noise power D_2 . That is, reducing the output noise power D_2 increases the degradation D_1 , and vice versa.

Since humans cannot perceive, or do not care about, slight degradation of the desired signal, it is effective to reduce the output noise power D_2 by allowing some degradation. Large degradation, however, cannot be accepted even if the output noise is greatly reduced. Considering this, AMNOR determines the optimal noise reduction filter based on the criterion of minimizing the output noise power D_2 while keeping the degradation D_1 below a threshold value $\hat{D}_1$. The efficiency of this criterion has been demonstrated in a previous paper,³⁾ where it was also shown that the tradeoff relationship between D_1 and D_2 makes the degradation D_1 of the optimal filter equal to the threshold value $\hat{D}_1$.

AMNOR assumes two conditions:

- 1) the direction of the desired sound source is known.
- 2) signals $U_{N1}(\omega), \dots, U_{NM}(\omega)$ that contain only noise can be received (Practically, this noise-only reception can be accomplished during the silent period of the desired speech).

When the desired sound wave can be regarded as a plane wave, the transfer functions $G_1(\omega), \dots, G_M(\omega)$ —except the transfer time delay—are derived from the desired sound direction and the arrangement of microphones in the array. Condition 1 thus gives $G_1(\omega), \dots, G_M(\omega)$.

Consider now the block diagram shown in Fig. 2. In this figure, $\sqrt{P_s} S'(\omega)$ denotes a fictitious desired signal (FD signal), where P_s is a constant called the FD signal level, and the expectation $E[|S'(\omega)|^2] = 1$. The FD signal is first applied to the filters $G_1(\omega), \dots, G_M(\omega)$. Then the outputs are added to the received signals that contain only noise. Thus, the signals $U_i'(\omega)$ ($i=1, \dots, M$) are produced as

$$U_i'(\omega) = U_{Ni}(\omega) + G_i(\omega) \sqrt{P_s} S'(\omega). \quad (8)$$

Considering $U_1'(\omega), \dots, U_M'(\omega)$ as inputs for the

Fig. 2 Block diagram indicating AMNOR noise reduction filter calculation method which is performed in the filter controller in Fig. 1.

noise reduction filter and $Y'(\omega)$ as the filter output at that time, the optimal filter that minimizes the mean square error $E[|\sqrt{P_s} S'(\omega) - Y'(\omega)|^2]$ is derived by the following equation with parameter P_s .

$$H(P_s) = P_s G_0^H R'^{-1} \\ = P_s G_0^H (R_N + P_s G_0 G_0^H)^{-1}, \quad (9)$$

where H represents conjugate transpose, and G_0 denotes a M -th order column vector whose elements are the transfer functions between the desired sound source and each microphone; i.e.,

$$G_0 \triangleq [G_1(\omega), G_2(\omega), \dots, G_M(\omega)]^T. \quad (10)$$

The term R' is an $M \times M$ cross spectrum matrix of $U_1'(\omega), \dots, U_M'(\omega)$, whose ij -th element is $E[U_i'(\omega) U_j'(\omega)^*]$ (where $*$ denotes complex conjugate). And R_N is a cross spectrum matrix of the received noise signals $U_{N1}(\omega), \dots, U_{NM}(\omega)$. Note that $H(P_s)$ is also a function of frequency and should strictly be represented by $H(P_s, \omega)$; parameter ω will be neglected, though, to simplify the notation.

Let $D_1(P_s)$ denote the degradation of the filter $H(P_s)$ of Eq. (9). Then $D_1(P_s)$ is a monotonically decreasing function of parameter P_s .^{3,4)} The value of P_s that satisfies the condition $D_1(P_s) = \hat{D}_1$ can therefore be found with arbitrary precision by adjusting P_s and repeatedly calculating $D_1(P_s)$. Substituting this P_s value into Eq. (9) finally gives the optimal filter based on the AMNOR criterion. All these calculations are performed at the filter controller shown in Fig. 1, and the optimal directivity characteristic is achieved by setting the derived filter characteristics into variable filters $H_1(\omega), \dots, H_M(\omega)$.

The filter $H(P_s)$ of Eq. (9) usually has a complicated characteristic and does not satisfy stability

condition. The optimal FIR filter design method described in the discrete time domain^{3,4)} is much more useful for practical application. This is because the FIR filter is appropriate for achieving a complex characteristic and is intrinsically stable. The discussion in the frequency domain is nonetheless continued in this paper because the better insight for directivity analysis is derived from consideration of the frequency domain.

2.3 Dead Angles in the Directivity Pattern

When the number of microphones M is greater than the number of noise sources N ; i.e., $M > N$; it is appropriate to rewrite Eq. (9) in the following more general formulation because the matrix $\mathbf{R}'$ in this equation is not necessarily a nonsingular matrix. Thus,

$$\mathbf{H}(\mathbf{P}_s) = \mathbf{P}_s \mathbf{G}_0^H \mathbf{R}'^+, \quad (11)$$

where $^+$ denotes a Moore-Penrose generalized inverse matrix.⁵⁾ Then, independent of the FD signal level $\mathbf{P}_s$, Eq. (11) is reduced to

$$\mathbf{P}_s \mathbf{G}_0^H \mathbf{R}'^+ = [1, 0, 0, \dots, 0] \mathbf{G}^+ \quad (12)$$

(see proof in the Appendix). The AMNOR output $Y(\omega)$ is now derived from Eqs. (1)–(4), (6), (11) and (12):

$$\begin{aligned} Y(\omega) &= \mathbf{H}(\mathbf{P}_s) \mathbf{U} \\ &= [1, 0, 0, \dots, 0] \mathbf{G}^+ \mathbf{U} \\ &= [1, 0, 0, \dots, 0] \mathbf{V} \\ &= S(\omega). \end{aligned} \quad (13)$$

This equation indicates that AMNOR with the optimal filter $\mathbf{H}(\mathbf{P}_s)$ retrieves the desired signal perfectly when $M > N$.

AMNOR thus completely reduces noise when there are fewer noise sources than microphones. When the noise sound waves are plane waves, canceling the noise arriving from a certain direction means creating a dead angle (a direction in which the microphone system has no sensitivity). It is clear that AMNOR can create at most $M-1$ dead angles in its directivity pattern.

3. DIRECTIVITY CHARACTERISTICS

The section studies the directivity characteristics of AMNOR comparing with those of conventional first-order gradient microphones in two representative noise environments: one in which noise comes from one direction, and one in which noise is uniformly distributed (Because most conventional direc-

tional microphones are first-order gradient microphones, the first-order gradient microphone is also referred to as the “directional microphone” in this paper).

3.1 First-Order Gradient Microphone

Conventional first-order gradient microphones can be functionally modeled by a processing diagram with two omnidirectional microphones separated by a small distance as shown in Fig. 3. In this figure, the first microphone's output represents the signal caused by the sound pressure at the front of the gradient microphone's membrane, and the second microphone's output represents the signal caused by the sound pressure at the back of this membrane. The delay unit represents the time difference between the sound pressures at the front and back of the membrane, and the subtraction operation represents the opposing ways in which each sound pressure contributes to the output: positively and negatively.

Suppose the desired signal comes from the direction $\theta = 180^\circ$ in Fig. 3. It first arrives at the first microphone and then it arrives at the second microphone after a delay time τ_d .

$$\tau_d = d/c, \quad (14)$$

where d denotes the distance between microphones and c is the sound velocity. The desired signal received by the second microphone is then delayed by the delay unit and subtracted from the output of the first microphone. Denoting the desired signal by $s(t)$, the subtraction output $w(t)$ is expressed as

$$w(t) = s(t) - s(t - \tau_d - \tau). \quad (15)$$

The time of origin for the desired signal $s(t)$ is adjusted to that observed at the first microphone. Taking the Fourier transform of Eq. (15),

Fig. 3 Model of conventional first-order gradient microphone.

Table 1 Directional patterns of first-order gradient microphone.

bidirection	unidirection	hypercardioid
$\tau = 0$	$\tau = \tau_d$	$\tau = \tau_d / 3$

$$W(\omega) = S(\omega) - S(\omega) e^{-j\omega(\tau_d + \tau)} = S(\omega)K(\omega), \quad (16)$$

where

$$K(\omega) = 1 - e^{-j\omega(\tau_d + \tau)}. \quad (17)$$

Equation (16) indicates the desired signal is distorted with the frequency characteristic $K(\omega)$ by the subtraction. The compensation filter in Fig. 3 is the inverse filter, with the characteristic $1/K(\omega)$, to compensate this distortion. Conventional directional microphones are designed to have the relationship, $\omega(\tau_d + \tau) \ll 1$. Due to this relationship, $K(\omega)$ can be represented as

$$K(\omega) \doteq j\omega(\tau_d + \tau). \quad (18)$$

Based on Eq. (18), the conventional microphones achieve a flat frequency response by choosing the proper method for controlling the vibration system in the microphone; *e.g.*, the mass control method is used in moving coil microphones.

In Fig. 3, unidirectional directivity is attained by designing the microphone structure to satisfy $\tau = \tau_d$, and bidirectional directivity attained by satisfying $\tau = 0$. Unidirectionality cancels the noise arriving as a plane wave at $\theta = 0^\circ$, and bidirectionality cancels the noise arriving at $\theta = 90^\circ$. Setting $\tau = \tau_d/3$ creates hypercardioid directivity, which achieves the greatest reduction of diffuse noise among the gradient-type directivities. The relationships between the delay time τ and the directivity pattern are listed in Table 1.

3.2 Noise Reduction Filter for Two-Element AMNOR

A block diagram of the noise reduction part of AMNOR using two omnidirectional microphone elements is shown in Fig. 4. Here only noise is assumed to be received by the microphones, whose outputs are denoted by $U_{N1}(\omega)$ and $U_{N2}(\omega)$. The FD signal level is again denoted by P_s and the transfer functions from the desired sound source to

Fig. 4 Noise reduction of AMNOR using 2 microphone elements.

each microphone are denoted by $G_1(\omega)$ and $G_2(\omega)$. Assuming equal noise power levels at both microphones; *i.e.*,

$$E[U_{Ni}(\omega)U_{Ni}(\omega)^*] = P_N \quad (i=1, 2), \quad (19)$$

the cross spectrum matrix of the received noise R_N is expressed by

$$R_N = P_N \begin{pmatrix} 1 & b^* \\ b & 1 \end{pmatrix}, \quad (20)$$

where

$$b \triangleq E[U_{N2}(\omega)U_{N1}(\omega)^*]/P_N. \quad (21)$$

Here $G_1(\omega)$ and $G_2(\omega)$ represent the transmission delays of a plane wave, and they are represented in vector form by the following equation:

$$G_0 = [G_1(\omega), G_2(\omega)]^T = [1, e^{-j\omega\tau_s}]^T, \quad (22)$$

where

$$\tau_s = -d \cos \theta_s / c, \quad (23)$$

and θ_s denotes the direction of the desired signal. The transmission delay is normalized to that of the first microphone. The vector product $G_0 G_0^H$ is expressed by

$$G_0 G_0^H = \begin{pmatrix} 1 & a^* \\ a & 1 \end{pmatrix}, \quad (24)$$

where

$$a \triangleq G_2(\omega)G_1(\omega)^* = e^{-j\omega\tau_s}. \quad (25)$$

Considering $aa^* = 1$, the noise reduction filter $H(P_s)$ is then derived from Eqs. (9), (20) and (24) as

$$H(P_s) = \frac{P_s}{P_s(2 - ab^* - a^*b) + P_N(1 - bb^*)} \cdot [1 - a^*b, a^* - b^*]. \quad (26)$$

Based on the definition of H ; *i.e.*, Eq. (5); the formulas for $H_1(\omega)$ and $H_2(\omega)$ are now derived as

$$H_1(\omega) = P_s(1 - a^*b)/K'(\omega) \quad (27)$$

$$H_2(\omega) = P_s(a^* - b^*)/K'(\omega) \quad (28)$$

where

$$K'(\omega) = P_s(2 - ab^* - a^*b) + P_N(1 - bb^*). \quad (29)$$

In the following discussion, the desired signal direction θ_s is assumed to be 180° , as with conventional directional microphones. This assumption and Eqs. (23) and (25) gives

$$a = e^{-j\omega\tau_d}. \quad (30)$$

3.3 Noise from a Single Direction

Consider noise incident from a single direction θ_N as a plane wave. This noise arrives at the second microphone τ_N before (or, if $\theta_N > 90^\circ$, after) it arrives at the first microphone.

$$\tau_N = \frac{d}{c} \cos \theta_N \quad (31)$$

Therefore,

$$U_{N2}(\omega) = U_{N1} e^{j\omega\tau_N}, \quad (32)$$

and from Eqs. (19) and (21)

$$b = e^{j\omega\tau_N}. \quad (33)$$

Substituting (30) and (33) into (27), (28) and (29) gives the following equations:

$$H_1(\omega) = 1/K'(\omega) \quad (34)$$

$$H_2(\omega) = -e^{-j\omega\tau_N}/K'(\omega) \quad (35)$$

$$K'(\omega) = 1 - e^{-j\omega(\tau_d + \tau_N)} \quad (36)$$

Consider the filters $H_1(\omega)$ and $H_2(\omega)$ in Fig. 4 to have characteristics expressed by Eqs. (34) and (35). Regarding the common term $1/K'(\omega)$ in $H_1(\omega)$ and $H_2(\omega)$ as the compensation filter characteristic, the input-output characteristic of the block diagram in Fig. 4 coincides with that of Fig. 3. For example, when $\theta_N = 0^\circ$, τ_N becomes τ_d using Eqs. (31) and (14). As a result, AMNOR forms a unidirectional pattern and suppresses the noise. Furthermore, τ_N becomes 0 when $\theta_N = 90^\circ$, and AMNOR forms a bi-directional pattern suppressing noise.

In these examples, the number of microphones M is 2 and the number of noise sources N is 1. This satisfies the condition $M > N$ described in the previous section, and AMNOR forms $M - 1 = 1$ dead angles in its directivity pattern (Bidirectional and hypercardioid patterns appear to have two dead angles because the directivity of a linear array is axis-symmetric, but the number of controllable dead angles is only one). AMNOR thus reduces noise by automatically pointing the dead angle in the direction of noise arrival.

The above discussion is also valid for multiple

noise sources when each noise has different frequency components. This is because only one arrival direction is observed at a specific frequency band. In this case, because the noise arrival direction θ_N is a function of frequency ω , so is the τ_N of Eq. (31). The $H_2(\omega)$ of Eq. (35) thus functions as a frequency-dependent delay unit. As a result, AMNOR forms different directivity patterns for different frequency bands. In other words, AMNOR forms the dead angle in the noise arrival direction of each frequency band; thus reducing multiple noise signals.

3.4 Noise in a Diffuse Sound Field

The cross spectrum between the signals received by two microphones coincides with the spatial correlation coefficient in a diffuse sound field.⁶⁾ That is,

$$b = \sin x / x = \rho, \quad (37)$$

where

$$x = \omega d / c. \quad (38)$$

Substituting (30) and (37) into (27), (28) and (29) gives the following equations:

$$H_1(\omega) = (1 - a\rho)/K'(\omega) \quad (39)$$

$$H_2(\omega) = (a - \rho)/K'(\omega) \quad (40)$$

$$K'(\omega) = 2 - \rho(a + a^*) + (P_N/P_s)(1 - \rho^2) \quad (41)$$

For comparison with the conventional directional microphone, the distance d between two microphones is assumed to be much smaller than the wave length; i.e., $d \ll 2\pi c/\omega$. This makes $x \ll 1$. By using the following approximations,

$$\sin x \doteq x - x^3/6 \quad \text{and} \quad (42)$$

$$a = e^{-j\omega d/c}$$

$$= e^{-jx}$$

$$\doteq 1 - x^2/2 - j(x - x^3/6), \quad (43)$$

Eqs. (39), (40) and (41) are rewritten as

$$H_1(\omega) = 1/K'(\omega) \quad (44)$$

$$H_2(\omega) = -e^{-j\omega(1/3)\tau_d}/K'(\omega) \quad \text{and} \quad (45)$$

$$K'(\omega) = (1 + (P_N/4P_s))(1 - e^{-j\omega(4/3)\tau_d}). \quad (46)$$

In these equations, the FD signal level P_s appears only in $K'(\omega)$, which is a common factor of $H_1(\omega)$ and $H_2(\omega)$. The value of P_s can therefore be neglected when AMNOR's directivity pattern is treated as a relative quantity. The input-output characteristic of the block diagram in Fig. 4; with $H_1(\omega)$ and $H_2(\omega)$ represented by Eqs. (44), (45) and (46); therefore coincides with that of Fig. 3 with

$\tau = \tau_a/3$ (except the gain factor). This means that AMNOR forms hypercardioid directivity, which is known to be optimal for the diffuse noise field.

As mentioned before, AMNOR can control one dead angle with two microphone elements. The contribution of one dead angle to reducing the noise in a diffuse noise field, however, is very small. The noise reduction effect, in this case, is therefore considered to be achieved by the total shape of the directivity pattern.

It might appear that the total directivity pattern control results from the dead angle control. To clarify this, assume that the desired signal direction is 90° . Substituting $a=1$ (which is derived from Eqs. (23) and (25)) into Eqs. (39), (40) and (41) gives the optimal filter as follows:

$$H_1(\omega) = 1/K'(\omega) \quad (47)$$

$$H_2(\omega) = 1/K'(\omega) \quad (48)$$

$$K'(\omega) = 2 + (P_N/P_S)(1 + \rho) \quad (49)$$

Equations (47) and (48) indicate that two microphone outputs are simply added together in AMNOR. In this case, assuming $d < c\pi/\omega$, sound from any direction is not completely canceled. In other words, the AMNOR directivity pattern has no dead angle. AMNOR's noise reduction effect is therefore attained not by control of the dead angle but by control of the total directivity pattern.

In conclusion, AMNOR achieves its noise reduction performance by optimally controlling not only dead angles but also the total directivity pattern shape.

4. FEATURES OF AMNOR

The main feature of AMNOR is that it adaptively achieves the optimal directivity characteristic for a given noise environment, desired signal direction, and number and arrangement of microphones.

The previous section demonstrated that AMNOR forms the same directivity patterns as conventional directional microphones when the number of microphones is two, the desired signal direction is 180° , and noise arrives from one direction or from all directions. This showed the basic principle of AMNOR is not unlike that of conventional directional microphones.

But while the directivity pattern of a conventional directional microphone is fixed and optimal only for a specific noise environment, AMNOR forms the optimal directivity pattern for any noise environ-

ment. In other words, the functional principle of AMNOR includes those of conventional directional microphones.

As previously described, AMNOR adaptively controls the dead angles and the total directivity pattern to achieve optimal directivity for a given noise environment. These controls are also performed independently at each frequency. For example, noise environment information such as arriving directions and magnitude at a frequency ω_1 is contained in the cross spectrum matrix $R_N(\omega_1)$. The response of the optimal filter for frequency ω_1 , $H(P_S, \omega_1)$, is derived from Eq. (9) by using the $R_N(\omega_1)$. The response of $H(P_S, \omega_1)$ therefore decides the directivity characteristic for the frequency ω_1 and this directivity is optimal for noise at the frequency ω_1 . Similarly, calculating $H(P_S, \omega_2)$ achieves the optimal directivity for the noise environment at frequency ω_2 . Thus, AMNOR independently produces the optimal directivity for each frequency band.

In the practical application of FIR filters,^{3,4)} only a finite number of frequency bands can be controlled independently. When the order of an FIR filter is $(2L-1)$, the number of frequency bands whose amplitude and phase characteristics are independently controllable is $(L+1)$. Therefore, the number of frequency bands for which AMNOR can independently control the directivity by using FIR filters is $(L+1)$.

A simulation confirming the above discussion is now described. The noise reduction FIR filter coefficients are calculated according to Ref. 3) or 4).

Figure 5 shows the noise environment simulated. The number of microphone elements is two, and the desired signal direction θ_s is 180° . Two noise signals that have the same power but different frequency bands, of 300–600 and 1,250–3,100 Hz, re-

Fig. 5 Noise environment for simulation: two noise sound waves having no common frequency component incident from different directions.

Fig. 6 AMNOR directivity pattern, (a) for 500 Hz and (b) for 2,000 Hz.

Fig. 7 Frequency response of AMNOR for directions of 0°, 126° (noise directions), and 180° (desired direction). On the bottom of the figure, ① shows the frequency range of noise coming from the 0° direction and ② shows the frequency range of noise coming from the 126° direction.

spectively arrive at 0° and at 126°. The conventional directional microphones can only reduce one of the noise signals in this environment. However, AMNOR with two microphone elements can reduce the noise signals from different directions because noise in each frequency band comes from only one direction; thus, the condition $M > N$ holds.

The simulation results with the 32-nd order FIR filters are shown in Figs. 6 and 7. Figure 6 shows the AMNOR directivity patterns formed for 500 Hz (a) and 2,000 Hz (b). AMNOR creates a dead angle in the noise direction at each frequency. Figure 7 shows the AMNOR frequency sensitivity characteristic at 0°, 126° and 180°. Here, ① and ② denote the frequency bands of noise signals coming from 0° and 126°. This figure shows that the sensitivity in each noise direction is greatly reduced at the frequency band of that noise signal. Figure 6(b) seems to indicate the sensitivity at

180° is also decreased, but this is only due to its relative expression; the flatness of the sensitivity at 180° is shown in Fig. 7.

Finally, the signal-to-noise ratio improvement derived in this simulation is about 40 dB for AMNOR and 7 dB for a conventional unidirectional microphone, a difference of 33 dB.

5. CONCLUSION

This paper describes the theory of AMNOR in the frequency domain and explores its directional character.

When noise arrives from one direction or is diffuse, AMNOR with two microphone elements automatically forms the same directivity patterns as those formed by conventional directional microphones. But in contrast to the directivity characteristics of conventional directional microphones, which are optimal for specific noise environments, AMNOR adaptively adjusts for optimal directivity in a wide variety of noise environments. This is accomplished by the following three functions:

- (1) With M microphone elements, AMNOR forms at most $M-1$ dead angles in its directivity pattern.
- (2) When effective noise reduction is not attained only by controlling the dead angle(s), AMNOR optimizes the directivity pattern figure.
- (3) AMNOR independently controls directivity for multiple frequency bands.

In a simulation experiment where two noise sound waves with no common frequency component came from different directions. AMNOR reduced noise about 33 dB more than a conventional directional microphone.

ACKNOWLEDGMENTS

The authors is grateful to Dr. Sadaoki Furui, Chief of NTT Speech and Acoustic Research Laboratory, and the colleagues of this laboratory for their discussion and comments. And Dr. Juro Ohga at Fujitsu Labs. and former head of the NTT Acoustic group, who proposed the AMNOR system with the author, is acknowledged for his continuing suggestions.

REFERENCES

- 1) J. L. Flanagan, J. D. Johnston, R. Zahn, and G. W. Elko, "Computer-steered microphone arrays for sound transduction in large room," *J. Acoust. Soc.*

- Am. **78**, 1508–1518 (1985).
- 2) F. Pirz, "Design of a wideband, constant beam-width, array microphone for use in the near field," *Bell Syst. Tech. J.* **58**, 1839–1850 (1979).
- 3) Y. Kaneda and J. Ohga, "Adaptive microphone-array system for noise reduction," *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-34**, 1391–1400 (1986).
- 4) Y. Kaneda, "AMNOR (Adaptive microphone array for noise reduction)," *Tech. Rep. Eng. Acoust. Group IECE*, EA 85-36 (1985) (in Japanese).
- 5) C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications* (John Wiley & Sons, New York, 1971).
- 6) H. Cox, "Spatial correlation in arbitrary noise fields with application to ambient sea noise," *J. Acoust. Soc. Am.* **54**, 1289–1301 (1973).

APPENDIX: THE PROOF OF EQ. (12)

The $(N+1)$ -th and M -th order column vectors V' and U' are first defined.

$$V' \triangleq [\sqrt{P_s} S'(\omega), N_1(\omega), N_2(\omega), \dots, N_N(\omega)]^T \quad (\text{A.1})$$

$$U' \triangleq [U_1'(\omega), U_2'(\omega), \dots, U_M'(\omega)]^T \quad (\text{A.2})$$

Using Eq. (8); the definition of $U_1'(\omega), \dots, U_M'(\omega)$; and Eq. (4); the definition of the matrix G ; the following relationship is then derived.

$$U' = GV' \quad (\text{A.3})$$

The cross spectrum matrix R' of $U_1'(\omega), \dots, U_M'(\omega)$ is expressed as

$$\begin{aligned} R' &= E[U' U'^H] \\ &= E[GV' V'^H G^H] \\ &= G E[[\sqrt{P_s} S'(\omega), N_1(\omega), \dots, N_N(\omega)]^T \\ &\quad [\sqrt{P_s} S'(\omega)^*, N_1(\omega)^*, \dots, N_N(\omega)^*]^H] G^H. \end{aligned} \quad (\text{A.4})$$

Since the FD signal $\sqrt{P_s} S'(\omega)$ and noise signals $N_1(\omega), \dots, N_N(\omega)$ are mutually independent, R' is derived as

$$R' = GQQG^H, \quad (\text{A.5})$$

where, Q is an $(N+1) \times (N+1)$ diagonal matrix expressed by

$$Q = \text{diag}[\sqrt{P_s}, \sqrt{P_{N1}}, \sqrt{P_{N2}}, \dots, \sqrt{P_{NN}}] \quad (\text{A.6})$$

and

$$P_{Nj} = E[N_j(\omega) N_j(\omega)^*] \quad (j=1, \dots, N). \quad (\text{A.7})$$

Equation (12) is therefore proved as follows:

$$\begin{aligned} P_s G_0^H R'^+ &= P_s G_0^H (GQQG^H)^+ \\ &= [1, 0, \dots, 0] P_s G^H (GQQG^H)^+ \\ &= [1, 0, \dots, 0] QQG^H (GQQG^H)^+ \\ &= [1, 0, \dots, 0] Q(GQ)^H ((GQ)(GQ)^H)^+ \\ &= [1, 0, \dots, 0] Q(GQ)^+ \\ &= [1, 0, \dots, 0] G^+ \end{aligned} \quad (\text{A.8})$$

The 5-th and 6-th equalities in Eq. (A.8) are shown in Ref. 5), and the 6-th equality holds only when $\text{rank}(G) \leq (N+1)$; i.e., when $M > N$.

Yutaka Kaneda was born in Osaka, Japan, on February 20, 1951. He received his Masters and Doctoral degrees in Engineering from Nagoya University, Japan, in 1977 and 1990. Since joining the Electrical Communication Laboratories, Nippon Telegraph and Telephone Corporation, Musashino, Tokyo, Japan, in 1977, he has been engaged in research on acoustical processing. He is now a Senior Research Engineer at the Speech and Acoustic Laboratory, NTT Human Interface Laboratories.