Acoust. Sci. & Tech. **24**, 1 (2003)

ACOUSTICAL LETTER

An adaptive microphone array for howling cancellation

Kazunori Kobayashi*, Ken'ichi Furuya and Akitoshi Kataoka

NTT Cyber Space Laboratories, NTT Corporation,
3–9–11 Midori-cho Musashino, 180–8585 Japan

(*Received 2 September 2002, Accepted for publication 25 September 2002*)

Keywords: Microphone array, Howling cancellation, Adaptive array, Null beamformer
PACS number: 43.38.Hz, 43.60.Lq

1. Introduction

Howling is a problem that occurs in public-address systems because the speech signals picked up by microphones in an auditorium or conference room are output by the loudspeakers, those signals are in turn picked up by the microphones, forming a feedback loop. In this paper, we propose an adaptive microphone array that suppresses howling by setting the directional response of the microphones such that the loudspeakers are at null points. A feature of this method is that the directions of the loudspeakers are determined while the person talking and the loudspeakers are simultaneously producing sound. Conventional adaptive microphone arrays lack this feature. Simulation results show that the proposed method is effective in suppressing howling.

2. Conventional howling cancellation techniques

Figure 1 shows an example of the public-address system. A loudspeaker outputs a speech signal that's been picked up by a microphone in the same room. The frequency response of the path from the microphone to the loudspeaker, including the acoustic feedback, is given by

$$P(\omega) = \frac{W(\omega)}{1 - W(\omega)F(\omega)}. \tag{1}$$

Here, $W(\omega)$ is the frequency response of the equipment between the microphone and the loudspeaker. $F(\omega)$ is the frequency response of the acoustic path from the speaker to the microphone, including the microphone's response. Howling occurs when the denominator of Eq. (1) is 0. The sufficient condition for howling to not occur is $|W(\omega)F(\omega)| < 1$.

The conventional technique for suppressing howling is to insert a filter that makes $|W(\omega)F(\omega)| < 1$ across the whole frequency range [1,2]. However, the volume from the loudspeaker is limited by the amplitude of the acoustic response, $F(\omega)$.

To increase this limit, the amplitude of $F(\omega)$ must be reduced by using a unidirectional microphone, a delay-and-sum microphone array, or an adaptive microphone array [3]. However, unidirectional microphones and delay-and-sum microphone arrays are not very effective in suppressing howling. With the conventional adaptive microphone array, the direction of the loudspeaker must be determined before anyone talks into the microphones. If the microphones or loudspeaker are moved, the array becomes ineffective as a howling-suppressor.

3. The proposed adaptive microphone array

Figure 2 shows the proposed adaptive microphone array for howling cancellation. With the conventional adaptive microphone array, it is difficult to determine the direction of the loudspeaker [4]. This is because the signal from the loudspeaker is strongly correlated with the speech signal from the talker. This problem is solved in the proposed method by the insertion of null beam-formers (NBFs) before the inputs to the adaptive filter to reduce the speech signal from the talker. An artificial source is placed at the talker's position. This is used to constrain the array's sensitivity in the talker's direction to unity. Suitable coefficients for the adaptive filter are then found by applying the normalized least-mean-square (NLMS) algorithm, which is given by

$$\boldsymbol{h}(n+1) = \boldsymbol{h}(n) - \alpha \frac{\boldsymbol{u}(n)e(n)}{\boldsymbol{u}(n)^{\mathrm{T}}\boldsymbol{u}(n)}, \tag{2}$$

where α is the step size, $e(n)$ is the error signal, and $\boldsymbol{h}(n)$ and $\boldsymbol{u}(n)$ are the vectors of adaptive filter coefficients $h_{m,j}(n)$ and input signals $u_m(n)$, respectively, and are given by

$$\boldsymbol{h}(n) = \left[h_{1,0}(n), \cdots, h_{1,L-1}(n), h_{2,0}(n), \cdots, h_{M-1,L-1}(n)\right]^{\mathrm{T}} \quad \text{and} \tag{3}$$

$$\boldsymbol{u}(n) = [u_1(n), \cdots, u_1(n-L+1), u_2(n), \cdots, u_{M-1}(n-L+1)]^{\mathrm{T}}. \tag{4}$$

Here, $h_{m,j}(n)$ is the coefficient of the j-th tap and the m-th channel of the adaptive filter at time n.

The result of this process is a design for the adaptive filter such that the response in the direction of the talker is constrained to unity and the direction of the loudspeaker is set as a null point in the directional response. These filter coefficients are copied to the output filter that is connected from the first microphone to the $(M-1)$-th microphone. The filter then produces output signals in which the signals from the loudspeaker have been suppressed. For the copied filter coefficients to be effective, the microphones must be in a

*e-mail: kobayashi.kazunori@lab.ntt.co.jp

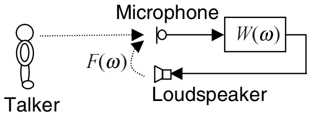

Fig. 1 An example of the public-address system.

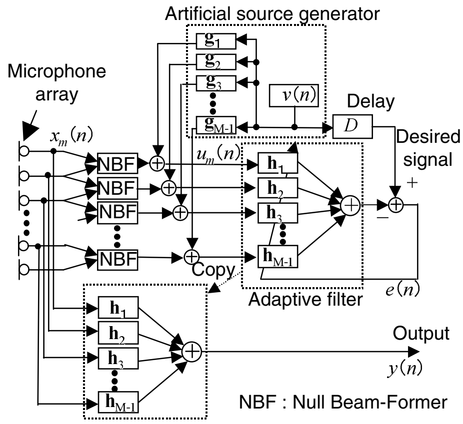

Fig. 2 Block diagram of the proposed method.

linear and equally spaced arrangement with much greater distances between the sound sources and microphones than the length of the array. This is because the time delays between the signals in the respective pairs of channels must be the same and this must still be so when the NBFs have been inserted.

4. Theoretical analysis

Firstly, an optimum solution is derived in the following way on the assumption that only direct-path sound is present. The conditions for an optimum filter are that the response from the talker to the output is unity but with the filter's delay D, and the response from the loudspeaker to the output is 0. Therefore, the equation for the optimum filter is

$$\begin{bmatrix} \boldsymbol{a}_S(\omega) \\ b \cdot \boldsymbol{a}_Y(\omega) \end{bmatrix} \boldsymbol{H}(\omega) = \begin{bmatrix} e^{-j\omega D} \\ 0 \end{bmatrix}, \quad (5)$$

where b is the gain between the speaker and the microphones, and $\boldsymbol{a}_S(\omega)$ and $\boldsymbol{a}_Y(\omega)$ are the vectors of the time delays τ_S and τ_Y between the respective pairs of microphones, on the signals from the talker and loudspeaker, respectively. These are given by

$$\boldsymbol{a}_S(\omega) = \left[1, e^{-j\omega\tau_S}, \cdots, e^{-j\omega(M-2)\tau_S}\right]^{\mathrm{T}} \quad \text{and} \quad (6)$$

$$\boldsymbol{a}_Y(\omega) = \left[1, e^{-j\omega\tau_Y}, \cdots, e^{-j\omega(M-2)\tau_Y}\right]^{\mathrm{T}}. \quad (7)$$

$\boldsymbol{H}(\omega)$ is the matrix of the filter's coefficient and is given by

$$\boldsymbol{H}(\omega) = [H_1(\omega), \cdots, H_{M-1}(\omega)]^{\mathrm{T}}. \quad (8)$$

The optimum filter $\boldsymbol{H}(\omega)$ is obtained as the minimum norm solution of Eq. (5), which is given by

$$\boldsymbol{H}(\omega) = \left(\boldsymbol{a}_S(\omega)\boldsymbol{a}_S(\omega)^{*\mathrm{T}} + b^2 \cdot \boldsymbol{a}_Y(\omega)\boldsymbol{a}_Y(\omega)^{*\mathrm{T}}\right)^{-1} \boldsymbol{a}_S(\omega)e^{-j\omega D} \quad (9)$$

where * denotes the complex conjugate.

Next, the equation for the filter after convergence of the proposed method is

$$\boldsymbol{H}(\omega) = \left\{\boldsymbol{a}_S(\omega)\boldsymbol{a}_S(\omega)^{*\mathrm{T}} + 2b^2 \cdot \{1 - \cos(\omega(\tau_S - \tau_Y))\} \frac{P_Y(\omega)}{P_V(\omega)} \boldsymbol{a}_Y(\omega)\boldsymbol{a}_Y(\omega)^{*\mathrm{T}}\right\}^{-1} \boldsymbol{a}_S(\omega)e^{-j\omega D} \quad (10)$$

where $P_Y(\omega)$ and $P_V(\omega)$ are the power spectra of the sound from the loudspeaker and artificial source, respectively. The derivation of this equation is not given in this paper.

The only difference between Eqs. (9) and (10) is in the multiplication of the $\boldsymbol{a}_Y(\omega)\boldsymbol{a}_Y(\omega)^{*\mathrm{T}}$ term by different weights. This reflects the difference at the inputs to the filter between the power spectra for sound from the loudspeaker and from the talker, which in turn is because of the NBFs and the artificial source. Therefore, while the proposed method will not get the optimum filter, it is applicable to the design of a filter such that the response in the direction of the talker is constrained to unity and the direction of the loudspeaker is set as a null point in the directional response.

5. Computer simulation

Figure 3 shows the positions of the microphone array, the talker and the loudspeaker in the simulation. The microphone array is a 5-element 2-cm-spaced linear array. The output of the loudspeaker is sound that has been picked up by the microphone array and processed by the proposed method. Reflections from the walls of the room are simulated by a mirror-image method. Each channel of the filter has a length of 128 taps, and the NLMS is used as the adaptive algorithm.

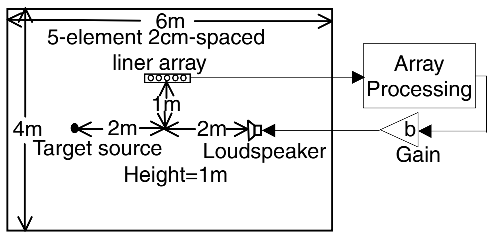

Fig. 3 Location of the microphone array, target source and loudspeaker.

5.1. Case I: non-reverberant room with a white-noise source

The effectiveness of the proposed method for a non-reverberant room and a white-noise source as the talker was demonstrated. The performance of the conventional delay-and-sum array is compared with that of the proposed method. Figure 4 shows the frequency responses of the paths from the loudspeaker and from the talker to the array's output with each of the methods. The frequency response from the loudspeaker with the proposed method is less than −20 dB in

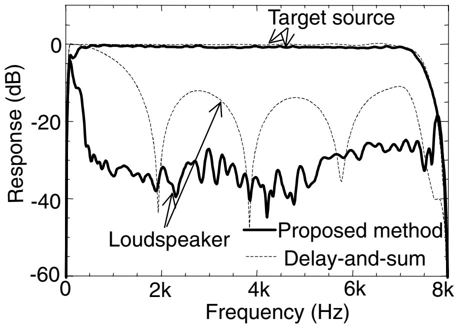

Fig. 4 Frequency response with the proposed method and delay-and-sum array.

the whole range above 300 Hz. On average, the response is 15 dB lower than that of the delay-and-sum array. Therefore, the proposed method allows us to get howling-free output from the loudspeaker at a volume 15 dB higher than with a delay-and-sum array. These results clarify the strong effectiveness of the proposed method in howling suppression.

5.2. Case II: reverberant room with a speech source

Table 1 shows the improvement in the howling margin for a speech source as the talker in a reverberant room with a reverberation time of 300 ms. Howling margin means the increase in howling onset volume relative to the case for a single microphone with no filtering. In case II, the proposed method provides a howling margin of 15 dB, while the margin for a delay-and-sum array is only 5 dB. This result clarifies the proposed method's effectiveness in suppressing howling under more realistic conditions.

Table 1 Howling margin with the proposed method and delay-and-sum array.

Proposed method	Delay-and-sum
14.5 dB	4.6 dB

6. Conclusions

We have proposed an adaptive microphone array that is operated to suppress howling by setting the directional response of the array such that the loudspeakers are at null points. The results of simulation showed that the proposed method provides a howling margin better by 10 dB than the conventional delay-and-sum array. Its strong effectiveness in the suppression of howling had thus been demonstrated.

References

[1] Y. Hirata, M. Nishimura and T. Itow, "Howlback suppressor in loudspeaking telephony using comb filters," *J. Acoust. Soc. Jpn. (J)* **32**, 247–252 (1976).

[2] S. Ushiyama, T. Hirai, M Tohyama and Y. Shimizu, "Howling suppression by smoothing the open-loop transfer function," *Tech. Rep. IEICE*, EA94-4, pp. 23–28 (1994).

[3] T. Ohga, Y. Yamazaki and Y. Kaneda, *Acoustic Systems and Digital Processing* (IEICE, Tokyo, 1995).

[4] B. Widrow, S. Strearns, *Adaptive Signal Processing* (Prentice-Hall, Englewood Cliffs, 1985).