PAPER *Special Section on Noise Cancellation and Reduction Techniques*

A Low-Distortion Noise Canceller and Its Learning Algorithm in Presence of Crosstalk

Akihiro HIRANO†, **Kenji NAKAYAMA**†, ***Regular Members***, **Shinya ARAI**††, ***and*** **Masaki DEGUCHI**†, ***Nonmembers***

SUMMARY This paper proposes a low-distortion noise canceller and its learning algorithm which is robust against crosstalk and is applicable for continuous sounds. The proposed canceller consists of two stages: cancellation of the crosstalk and cancellation of the noise. A recursive filter reduces the number of computations for noise cancellation stage. Separate filters for the adaptation and the filtering are introduced for crosstalk cancellation. Computer simulations show 10 dB improvement of the error power.

key words: *adaptive noise canceller, crosstalk, low distortion*

1. Introduction

Extracting desired signals from noise-corrupted signals is important in communication systems and sound recording. Adaptive noise cancellers (ANC's) [1]–[8] are widely used to reduce such noise. An ANC uses two microphones: the primary microphone for obtaining the noise-corrupted signals, and the reference microphone for capturing the noise. The noise from the reference microphone is filtered to generate the replica of the noise in the primary microphone output.

In ANC's, the desired signal components captured by the reference microphone, known as a "crosstalk [3]," is an important problem [2]–[7]. When the crosstalk is present, the desired signal is linearly distorted by the desired signal components mixed in the reference signals.

Crosstalk-resistant noise cancellers [2]–[7] have been proposed to improve the performance of ANC's in presence of the crosstalk. There are three classes of crosstalk-resistant ANC's: cross-coupled ANC [2], [3], ANC with pre-processing [4], and ANC with post-processing [5]–[7]. The cross-coupled ANC's uses two adaptive filters for crosstalk cancellation and noise cancellation. The ANC with pre-processing also has two stages for crosstalk cancellation and noise cancellation. ANC's with post-processing have an linear equalizer which reduces the distortion.

Most of crosstalk-resistant ANC's requires absence of the desired signal for correct adaptation. Though such absence is natural for speech signals, desired signals will be "continuous" for music recording. Even for speech signals, speech detectors for adaptation control is necessary and its performance affects the sound quality. Therefore, ANC's without such speech detection are preferable.

This paper proposes a low-distortion noise canceller and its learning algorithm which is robust against crosstalk and is applicable for continuous sounds. The cross-talk problem and some crosstalk-resistant noise cancellers are reviewed in Sect. 2. Section 3 derives a new low-distortion algorithm followed by the computer simulations.

Manuscript received May 12, 2000.

Manuscript revised August 14, 2000.

†The authors are with the Faculty of Engineering, Kanazawa University, Kanazawa-shi, 920-8667 Japan.

††The author is with Fujitsu Hokuriku Communication Systems Limited, Kanazawa-shi, 920-0031 Japan.

2. Crosstalk in Noise Cancellation

2.1 Classical Noise Canceller

Figure 1 depicts a classical noise canceller [1] in the presence of a crosstalk. A desired signal $S(z)$ propagates the acoustic path $H_4(z)$ and reaches the primary microphone "Mic1." A noise $V(z)$ via $H_1(z)$ is mixed at Mic1 and generates the primary signal $D(z)$. The reference microphone "Mic2" captures the noise $V(z)$ via $H_2(z)$ to output the reference signal $X(z)$. The adaptive filter $W(z)$ generates the replica of the noise. Subtracting the replica from the primary signal cancels the noise.

This noise canceller works fine if no desired signals are mingled with the noise in the reference signal. However, $S(z)$ reaches Mic2 via $H_2(z)$ and causes the distortion of the desired signal [2]–[7]. In the presence of the crosstalk, the calnceller output $E(z)$ is given by

$$E(z) = \{H_4(z) - H_3(z)W(z)\}S(z)$$

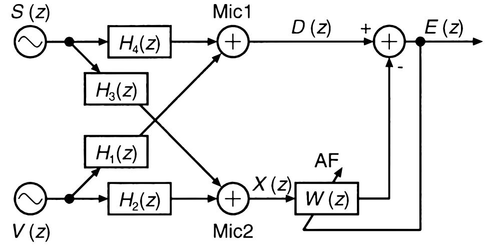


Fig. 1 Classical noise canceller in presence of crosstalk.

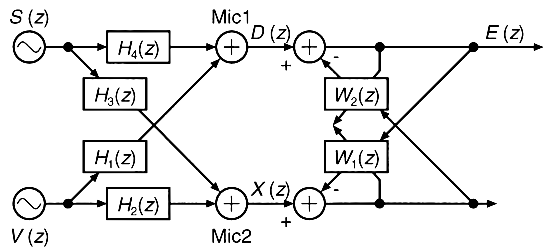


Fig. 2 Cross-coupled noise canceller.

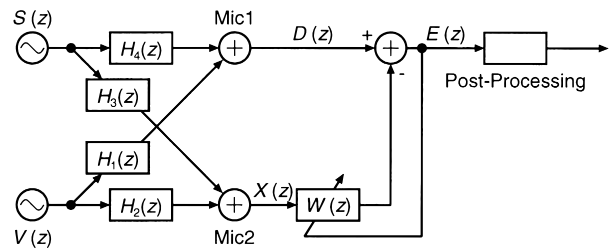


Fig. 4 Noise canceller with post-processing.

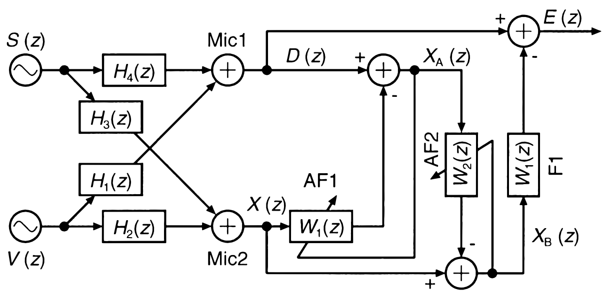


Fig. 3 Noise canceller with pre-processing.

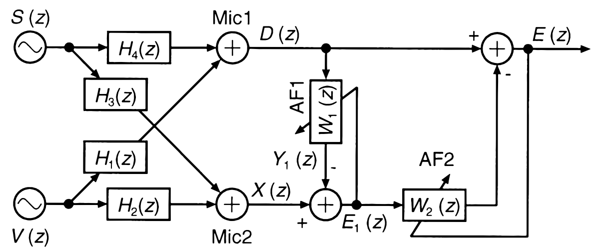


Fig. 5 Basic block diagram of proposed noise canceller.

$$+\{H_1(z) - H_2(z)W(z)\}V(z). \tag{1}$$

It is obvious that there are no solution for $W(z)$ which generates the optimum output $H_4(z)S(z)$. To reduce the signal distortion caused by the crosstalk, several noise cancellers has been proposed [2]–[7].

2.2 Cross-Coupled Noise Canceller

A cross-coupled structure [2], [3] shown in Fig. 2 consists of two adaptive filters. The adaptive filter $W_1(z)$ cancels $S(z)$ in $X(z)$ while $W_2(z)$ reduces the noise $V(z)$ in $D(z)$. The drawback of this structure is the adaptation control [2]. Simultaneous adaptation of $W_1(z)$ and $W_2(z)$ might causes performance degradation. A practical solution for intermittent signals such as a speech is updating $W_2(z)$ only when $S(z)$ is absent.

2.3 Noise Canceller with Pre-Processing

In the noise canceller with pre-processing [4] shown in Fig. 3, the adaptive filters $W_1(z)$ and $W_2(z)$ identify the paths $H_1(z)$ and $H_3(z)$, respectively. $W_2(z)$ also reduces $S(z)$ in $X(z)$. Finally, the filter F1, which uses the same filter coefficients as $W_1(z)$, cancels the noise. This canceller also requires adaptation control based on the absence of $S(z)$; $W_1(z)$ is updated when $S(z)$ is absent.

2.4 Noise Canceller with Post-Processing

Another class of noise cancellers use a post-processing to reduce the distortion [5]–[7]. Figure 4 depicts simplified block diagram. Since the distortion process is linear and is a combination of $H_i(z)$'s ($i = 1, \cdots, 4$), a linear equalizer can reduce the distortion. Some of these noise cancellers still requires adaptation control based on the absence of $S(z)$ [5].

3. Proposed Noise Canceller

3.1 Assumptions

Before deriving the algorithm, let us assume the followings.

1. both $S(z)$ and $V(z)$ are continuous sounds.
2. $S(z)$ and $V(z)$ are independent each other.
3. $H_1(z)$ has a longer delay than that of $H_2(z)$.
4. $H_3(z)$ has a longer delay than that of $H_4(z)$.
5. Correlation between $H_1(z)V(z)$ and $H_2(z)V(z)$ is negligible.

Because of the first assumption, many conventional algorithms are difficult to be applied though it is natural for music recording in a noisy environment. The third and fourth assumptions will be valid if Mic1 and Mic2 are located near the sound source and the noise source, respectively. The fifth assumption is also valid if Mic2 is located near the noise source. The influence of assumptions 4 and 5 will be discussed by using simulation results.

3.2 Basic Structure

Figure 5 depicts the basic block diagram of the proposed noise cacneller. It consists of two stages: cancellation of the desired signal $S(z)$ mixed in $X(z)$ and cancellation of the noise $V(z)$ in $D(z)$. The first stage consists of the adaptive filter $W_1(z)$, which cancels $S(z)$ in $X(z)$. Thus the reference signal for the second stage $E_1(z)$ is generated. In the second stage, $V(z)$ in $D(z)$

is canceled by $W_2(z)$ using $E_1(z)$.

The reference signal for the second stage $E_1(z)$ and the canceller output $E(z)$ are calculated by

$$\begin{aligned} E_1(z) &= \{H_3(z) - H_4(z)W_1(z)\}S(z) \\ &\quad + \{H_2(z) - H_1(z)W_1(z)\}V(z) \end{aligned} \tag{2}$$

and

$$\begin{aligned} E(z) &= \{H_4(z) - (H_3(z) \\ &\quad - H_4(z)W_1(z))W_2(z)\}S(z) \\ &\quad + \{H_1(z) - (H_2(z) \\ &\quad - H_1(z)W_1(z))W_2(z)\}V(z). \end{aligned} \tag{3}$$

From the optimum output $E(z) = H_4(z)S(z)$, the optimum transfer functions for the adaptive filters $W_1^{opt}(z)$ and $W_2^{opt}(z)$ are derived as

$$W_1^{opt}(z) = \frac{H_3(z)}{H_4(z)} \tag{4}$$

and

$$\begin{aligned} W_2^{opt}(z) &= \frac{H_1(z)}{H_2(z) - H_1(z)W_1(z)} \\ &= \frac{\dfrac{H_1(z)}{H_2(z)}}{1 - W_1(z)\dfrac{H_1(z)}{H_2(z)}}, \end{aligned} \tag{5}$$

respectively. Note that $W_1^{opt}(z)$ will cancel $S(z)$ component from $X(z)$.

3.3 Realization of $W_2(z)$ by Recursive Filter

The optimum transfer function for $W_2(z)$ shown in (5) suggests some difficulties on the adaptation and the implementation. Since $W_2^{opt}(z)$ consists of $W_1(z)$, $W_2(z)$ should track the change of $W_1(z)$ as well as learn the room acoustics. The recursive form of (5) might results in the huge number of taps for $W_2(z)$.

To reduce such difficulties, a recursive filter shown in Fig. 6 is introduced. The transfer function of the recursive filter is given by

$$W_2(z) = \frac{W_2'(z)}{1 - W_1(z)W_2'(z)}. \tag{6}$$

Comparing (5) and (6) leads us to the optimum transfer function of $W_2'(z)$ as

$$W_2'^{opt}(z) = \frac{H_1(z)}{H_2(z)}. \tag{7}$$

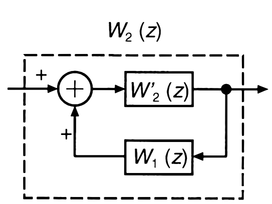


Fig. 6 Implementation of $W_2^{opt}(z)$ by recursive filter.

Note that $W_2'^{opt}(z)$ is independent of $W_1(z)$ and is the same as that for the classical noise canceller shown in Fig. 1.

The recursive form of (6) involves stability problem. A proof for the stability and adaptation algorithms which ensures the stability should be future study.

3.4 Adaptation of $W_1(z)$

The convergence value of $W_1(z)$ will be analyzed in the time domain to clarify the problem for adapting AF1. The primary signal $d(n)$ and the reference signal $x(n)$ at the time index n are given by

$$d(n) = \mathbf{h}_1^T\mathbf{v}(n) + \mathbf{h}_4^T\mathbf{s}(n) \tag{8}$$

and

$$x(n) = \mathbf{h}_2^T\mathbf{v}(n) + \mathbf{h}_3^T\mathbf{s}(n), \tag{9}$$

respectively. The vectors $\mathbf{h}_i$ $(i = 1, \cdots, 4)$ are the impulse responses corresponding to $H_i(z)$ and are defined by

$$\mathbf{h}_i = \Big[\overbrace{\underbrace{0 \ \cdots \ 0}_{N_{Di}} \ h_{i,0} \ \cdots \ h_{i,N_{Hi}-1}}^{N_H}\Big]^T \quad (i = 1, 3) \tag{10}$$

$$\mathbf{h}_i = \Big[\overbrace{h_{i,0} \ \cdots \ h_{i,N_{Hi}-1} \ \underbrace{0 \ \cdots \ 0}_{N_{Di}}}^{N_H}\Big]^T \quad (i = 2, 4). \tag{11}$$

N_{Hi} is the length of a non-zero section in $\mathbf{h}_i$, while N_{Di} is the that of a flat-delay section. N_H is the dimension of $\mathbf{h}_i$ and is equal to $N_{Hi} + N_{Di}$. $\mathbf{s}(n)$ and $\mathbf{v}(n)$ are the desired signal vector and the noise vector which are defined by

$$\mathbf{s}(n) = [s(n) \ \cdots \ s(n - N_H + 1)]^T \tag{12}$$

and

$$\mathbf{v}(n) = [v(n) \ \cdots \ v(n - N_H + 1)]^T, \tag{13}$$

respectively. $[\cdot]^T$ denotes the transpose of a matrix $[\cdot]$.

The AF1 output $y_1(n)$ is calculated by

$$y_1(n) = \mathbf{w}_1^T(n)\mathbf{d}(n) \tag{14}$$

where $\mathbf{w}_1(n)$ is the filter coefficients of AF1 defined by

$$\mathbf{w}_1(n) = [w_{1,0}(n) \ \cdots \ w_{1,N_{W1}-1}(n)]^T, \tag{15}$$

$\mathbf{d}(n)$ is the primary signal vector defined by

$$\mathbf{d}(n) = [d(n) \ \cdots \ d(n - N_{W1} + 1)]^T. \tag{16}$$

N_{W1} is the number of taps for AF1. Using $s(n)$ and $v(n)$, $y_1(n)$ is re-written as

$$y_1(n) = \mathbf{w}_1^T(n)\{\mathbf{H}_1^T\mathbf{v}'(n) + \mathbf{H}_4^T\mathbf{s}'(n)\} \tag{17}$$

where $\mathbf{s}'(n)$ and $\mathbf{v}'(n)$ are the desired signal vector and the noise vector defined by

$$\mathbf{s}'(n) = [s(n) \ \cdots \ s(n - N_{W1} - N_H + 2)]^T \tag{18}$$

and

$$\mathbf{v}'(n) = [v(n) \ \cdots \ v(n - N_{W1} - N_H + 2)]^T, \tag{19}$$

respectively. The difference between $\mathbf{s}(n)$ and $\mathbf{s}'(n)$ is the vector length. Matrices $\mathbf{H}_i$ defined by

$$\mathbf{H}_i = \begin{bmatrix} \mathbf{h}_i & \mathbf{0}_1 & \cdots & \mathbf{0}_{N_{W1}-2} & \mathbf{0}_{N_{W1}-1} \\ & \mathbf{h}_i & \cdots & \mathbf{h}_i & \\ \mathbf{0}_{N_{W1}-1} & \mathbf{0}_{N_{W1}-2} & \cdots & \mathbf{0}_1 & \mathbf{h}_i \end{bmatrix} \tag{20}$$

contains a vector $\mathbf{h}_i$ and corresponds to the convolution of $\mathbf{h}_i$ and its input. $\mathbf{0}_i$ is an i-th order zero vector. Multiplying each row vector of $\mathbf{H}_i$ by the input vector results in an output sample from a filter $\mathbf{h}_i$.

The error $e_1(n)$ for AF1 is generated by

$$e_1(n) = x(n) - y_1(n). \tag{21}$$

By taking the ensemble average, $E[e_1^2(n)]$ becomes

$$E[e_1^2(n)] = E[x^2(n)] - 2E[x(n)y_1(n)] + E[y_1^2(n)]. \tag{22}$$

The second term of (22) is further calculated as

$$\begin{aligned} E[x(n)y_1(n)] &= \mathbf{h}_3^T E[\mathbf{s}(n)\mathbf{s}'(n)]\mathbf{H}_4\mathbf{w}_1(n) \\ &\quad + \mathbf{h}_3^T E[\mathbf{s}(n)\mathbf{v}'(n)]\mathbf{H}_1\mathbf{w}_1(n) \\ &\quad + \mathbf{h}_2^T E[\mathbf{v}(n)\mathbf{s}'(n)]\mathbf{H}_4\mathbf{w}_1(n) \\ &\quad + \mathbf{h}_2^T E[\mathbf{v}(n)\mathbf{v}'(n)]\mathbf{H}_1\mathbf{w}_1(n) \\ &= \mathbf{h}_3^T \mathbf{R}_{ss'}\mathbf{H}_4\mathbf{w}_1(n) \\ &\quad + \mathbf{h}_2^T \mathbf{R}_{vv'}\mathbf{H}_1\mathbf{w}_1(n). \end{aligned} \tag{23}$$

Cross correlations $E[\mathbf{s}(n)\mathbf{v}'(n)]$ and $E[\mathbf{v}(n)\mathbf{s}'(n)]$ are equal to zero because of the assumption 2. From the assumption 5, $\mathbf{h}_2^T\mathbf{R}_{vv'}\mathbf{H}_1$, which corresponds to the correlation between $H_1(z)V(z)$ and $H_2(z)V(z)$, is also zero. Thus $E[x(n)y_1(n)]$ becomes

$$E[x(n)y_1(n)] = \mathbf{h}_3^T\mathbf{R}_{ss'}\mathbf{H}_4\mathbf{w}_1(n). \tag{24}$$

Similarly,

$$\begin{aligned} E[y_1^2(n)] &= \mathbf{w}_1(n)\mathbf{H}_1^T\mathbf{R}_{v'v'}\mathbf{H}_1\mathbf{w}_1(n) \\ &\quad + \mathbf{w}_1(n)\mathbf{H}_4^T\mathbf{R}_{s's'}\mathbf{H}_4\mathbf{w}_1(n) \end{aligned} \tag{25}$$

is derived for $E[y_1^2(n)]$.

The convergence value of AF1 $\mathbf{w}_1(\infty)$ is derived from

$$\begin{aligned} \frac{\partial E[e_1^2(n)]}{\partial \mathbf{w}_1} &= 2\left(\mathbf{H}_1^T\mathbf{R}_{v'v'}\mathbf{H}_1 + \mathbf{H}_4^T\mathbf{R}_{s's'}\mathbf{H}_4\right)\mathbf{w}_1(n) \\ &\quad -2\mathbf{h}_3^T\mathbf{R}_{ss'}\mathbf{H}_4 = 0. \end{aligned} \tag{26}$$

The convergence value is calculated as

$$\mathbf{w}_1(\infty) = \left(\mathbf{H}_1^T\mathbf{R}_{v'v'}\mathbf{H}_1 + \mathbf{H}_4^T\mathbf{R}_{s's'}\mathbf{H}_4\right)^{-1}\mathbf{h}_3^T\mathbf{R}_{s's'}\mathbf{H}_4. \tag{27}$$

The optimum value $\mathbf{w}_1^{opt}$ is derived by equating $v(n)$ to zero and is given by

$$\mathbf{w}_1^{opt}(\infty) = \left(\mathbf{H}_4^T\mathbf{R}_{s's'}\mathbf{H}_4\right)^{-1}\mathbf{h}_3^T\mathbf{R}_{s's'}\mathbf{H}_4, \tag{28}$$

which corresponds to $H_3(z)/H_4(z)$.

Obviously, the convergence value (27) is different from the optimum value (28). The difference is $\mathbf{H}_1^T\mathbf{R}_{v'v'}\mathbf{H}_1$ in the inverse matrix, which corresponds to the auto-correlation of the noise component $v(n)$ in the primary input $d(n)$. Thus the noise component $v(n)$ in $d(n)$ disturbs the adaptation of AF1.

The influence of $v(n)$ in $d(n)$ will be significant if the noise $v(n)$ is wide-band, or if $v(n)$ is large in the frequency range where $s(n)$ is dominant. On the other hand, the influence of $v(n)$ is negligible if $v(n)$ is small in the frequency range where $s(n)$ is dominant. A narrow-band noise corresponds to a latter case.

3.5 Proposed Structure

For correct adaptation, AF1 should be updated by using a reference input which contains less $v(n)$ component than $d(n)$. Thus AF1 is divided into two filters: an adaptive filter to update $\mathbf{w}_1(n)$ and a digital filter to generate $y_1(n)$ from $d(n)$ using $\mathbf{w}_1(n)$. To update $\mathbf{w}_1(n)$, $e(n)$ is used as the reference input rather than $d(n)$.

Figure 7 shows the block diagram of the proposed noise canceller. It consists of two adaptive filters, AF1 and AF2, and two filters, F1 and F2. Two filters F1 and F2 use the same coefficients as AF1. To ensure correct convergence of $W_1(z)$, AF1 uses $E(z)$ as its reference

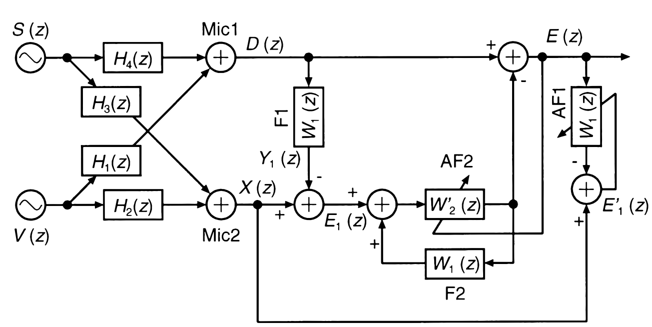


Fig. 7 Proposed noise canceller.

input rather than $D(z)$. Since $E(z)$ contains less noise components, the influence of $V(z)$ on the convergence of $W_1(z)$ is reduced. The recursive structure shown in Fig. 6 is used as AF2.

The canceller output $E(z)$ is calculated as

$$E(z) = \left\{ H_4(z) - \frac{(H_3(z) - H_4(z)W_1(z))W_2'(z)}{1 - W_1(z)W_2'(z)} \right\} S(z) + \frac{H_1(z) - H_2(z)W_2'(z)}{1 - W_1(z)W_2'(z)} V(z). \tag{29}$$

From (29), it is easy to show that the optimum transfer functions are (4) and (5). The noise $V(z)$ can be canceled regardless of $W_1(z)$.

The error $E'(z)$ for AF1 is given by

$$\begin{aligned} E'(z) &= X(z) - W_1(z)E(z) \\ &= \frac{H_3(z) - H_4(z)W_1(z)}{1 - W_1(z)W_2'(z)} S(z) \\ &\quad + \frac{H_2(z) - H_1(z)W_1(z)}{1 - W_1(z)W_2'(z)} V(z). \end{aligned} \tag{30}$$

After convergence of $W_2'(z)$ to its optimum value, $E'(z)$ becomes

$$\begin{aligned} E'(z) &= X(z) - W_1(z)E(z) \\ &= \frac{H_3(z) - H_4(z)W_1(z)}{1 - W_1(z)W_2'(z)} S(z) + H_2(z)V(z). \end{aligned} \tag{31}$$

Only $S(z)$ component in $E'(z)$ can be manipulated by $W_1(z)$. Therefore, $W_1(z)$ can be updated free from the influence of $V(z)$ in $D(z)$.

3.6 Number of Computations

Number of computations for some algorithms are compared. Table 1 shows the number of multiplications for the LMS algorithm [1]. The proposed algorithm requires $2N_{W1}$ multiplications for AF1, $2N_{W2}$ for AF2, N_{W1} for F1, and N_{W1} for F2. N_{W2} is the number of taps for AF2. Though removing the recursive filter F2 (without IIR) seems to reduce the number of computations, it will cause the increase of N_{W2} as suggested in Sect. 3.3. The effects of removing F2 on the performance and on N_{W2} will further be examined by computer simulations.

The proposed ANC seems to require more number of computations than that of conventional crosstalk-resigtant ANC's. However, speech detection required for conventional ANC's is not included in Table 1.

Table 1 Number of computations.

Algorithm	Number of Multiplications
Conventional	$2N_W + 1$
Cross-Coupled	$4N_W + 2$
Pre-Processing	$3N_{W1} + 2N_{W2} + 2$
Proposed	$4N_{W1} + 2N_{W2} + 2$
without IIR	$3N_{W1} + 2N_{W2} + 2$

4. Computer Simulations

Simulations have been carried out to show the performance of the proposed noise canceller. In the simulations, the error power $EP(n)$ defined by

$$EP(n) = \frac{\sum_{i=0}^{499} |d_s(n-i) - e(n-i)|^2}{\sum_{i=0}^{499} |d_s(n-i)|^2}, \tag{32}$$

the residual noise $RN(n)$ defined by

$$RN(n) = \frac{\sum_{i=0}^{499} |e_v(n-i)|^2}{\sum_{i=0}^{499} |d_s(n-i)|^2}, \tag{33}$$

the signal distortion $SD(n)$ defined by

$$SD(n) = \frac{\sum_{i=0}^{499} |d_s(n-i) - e_s(n-i)|^2}{\sum_{i=0}^{499} |d_s(n-i)|^2}, \tag{34}$$

have been compared. $d_s(n)$ is the signal component in $d(n)$ and is given by

$$d_s(n) = \mathbf{h}_4^T \mathbf{s}(n). \tag{35}$$

Note that $d_s(n)$ is the optimal output of the noise canceller. $e_s(n)$ and $e_v(n)$ are the signal component and the noise component in the output $e(n)$, respectively. $e_s(n)$ corresponds to the first term of (29), while $e_v(n)$ corresponds to the second term. The convergence value of $W_1(z)$ has also been compared to examine the convergence of AF1.

Performance for both white-noise and colored noise have been evaluated. The colored signals have been generated by second-order autoregressive (AR) models. The frequency responses of the AR models are depicted by Fig. 8. For AR signals, $V(z)$ is small in the frequency range where $S(z)$ is dominant. On the contrary, $V(z)$ is large in the frequency range where $S(z)$ is dominant for the white-noise case.

These evaluations will also confirm the influence of the noise component $V(z)$ in the primary input $D(z)$ on the adaptation of $W_1(z)$, which is analyzed in Sect. 3.4. For the white-noise case, the analysis shows that the

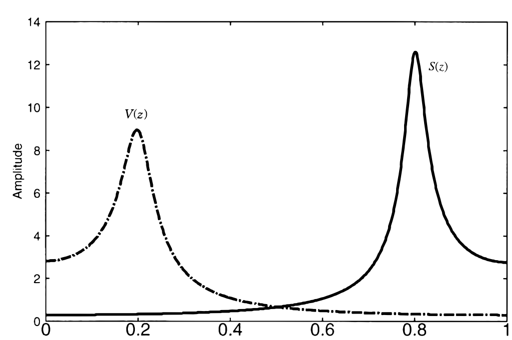


Fig. 8 AR model for colored signals.

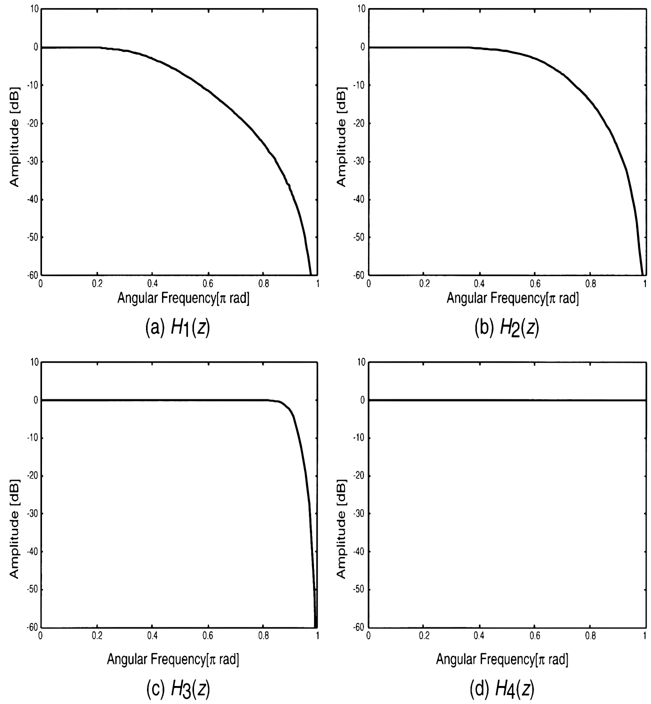


Fig. 9 Frequency responses of unknown systems.

adaptation of $W_1(z)$ using $D(z)$ would be affected by $V(z)$. On the contrary, the performance would be independent of the adaptation of $W_1(z)$ for the AR signals shown in Fig. 8. These facts suggest that a noise canceller which updates $W_1(z)$ using $D(z)$ would work fine for these AR signals, but not for white signals.

Figure 9 shows the frequency responses of the unknown systems. $H_1(z)$ and $H_2(z)$ are second-order Butterworth low-pass filters (LPF's) with the cut-off frequencies of 0.4 and 0.6, respectively. $H_3(z)$ is a third-order Butterworth LPF with the cut-off frequency of 0.9. $H_4(z) = 1$. $H_1(z)$ and $H_3(z)$ also contain flat delay sections. The delay length N_{D1} and N_{D3} are 30, which corresponds to 1.28 m distance for 8 kHz sampling.

The performance of the classical noise canceller shown in Fig. 1 (Conventional) and the proposed noise canceller (Proposed) are compared. Low-distortion algorithms which depend on the absent of $S(z)$ are not examined because they are not applicable for continuous sounds. In order to evaluate the effect of recursive structure, the proposed noise canceller using an FIR adaptive filter as AF2 (without IIR) and a longer-tap version (without IIR, $N_{W2} = 128$) are also compared. The performance of the proposed noise canceller which updates AF1 using $D(z)$ as in Fig. 5 (AF1 using $D(z)$) shows the influence of $V(z)$ in $D(z)$ on the adaptation of AF1.

As the adaptive filters, FIR adaptive filters based on the normalizes LMS algorithm (also known as learning identification method) [9] are used. The number of taps for AF1 and AF2, N_{W1} and N_{W2}, are chosen as 64 except for specially mentioned. The adaptation stepsize for both AF1 and AF2 is 0.01.

Figure 10(a) compares the the error power $EP(n)$ for a white-noise case. The proposed algorithm improves the error power by almost 10 dB compared with the conventional algorithm. The performance is degraded by removing the recursive part, or by changing the adaptation of AF1.

Figure 10(b) shows the residual noise component $RN(n)$ in $EP(n)$. From $RN(n)$, the effect of the recursive structure is clear. Without the recursive part, $RN(n)$ becomes almost 10 dB larger. Though increasing N_{W2} improves the residual noise level, the convergence speed becomes slow. Furthermore, the number of computations for $N_{W2} = 128$ without recursive part is larger than that for proposed method as shown in Table 1. $RN(n)$ is not affected by the adaptation of AF1.

As demonstrated by Fig. 10(c), the adaptation of AF1 has a great role on the signal distortion $SD(n)$. By changing the adaptation of AF1, $SD(n)$ is improved by 8 dB. The structure of AF2 has no influence on the signal distortion. Though increase of N_{W2} seems to improve $SD(n)$, this is because of the slow convergence; the signal distortion is small if $W_2(z)$ is almost zero.

Figure 10(d) compares the convergence value of $W_1(z)$, which affects the signal distortion. By the proposed method, $W_1(z)$ converges closer to the optimum than that using $D(z)$.

Figure 11 compares the simulation results for colored signals. In this simulation, the frequency of the signal $S(z)$ and the noise $V(z)$ is different. Therefore, the adaptation of AF1 has almost no influence on the performance. For both algorithms, $W_1(z)$ converges to the optimum value where the signal $S(z)$ is dominant.

In order to evaluate the performance when assumptions 4 and 5 are not valid, the relation between the delay N_{D3} of $H_3(z)$ and the error power $NP(z)$ has been examined. The relation between N_{D3} and $NP(z)$ is shown in Fig. 12. If AF1 is adapted using $D(z)$, N_{D3} should be longer than 20 samples for $NP(n)$ less than -20 dB. By updating AF1 using $E(z)$, $NP(n)$ is almost -25 dB when N_{D3} is longer than 2 samples. Thus requirements on N_{D3}, i.e. limitations on the microphone arrangement, can be alleviated.

In all simulations shown above, the recursive filter in the proposed noise canceller is stable. Simulations for other signals, noises and transfer functions also show no stability problem.

5. Conclusion

A low-distortion noise canceller and its learning algorithm has been proposed which is robust against crosstalk and is applicable for continuous sounds. The proposed canceller consists of two stages: cancellation of the desired signal from the reference signal, and can-

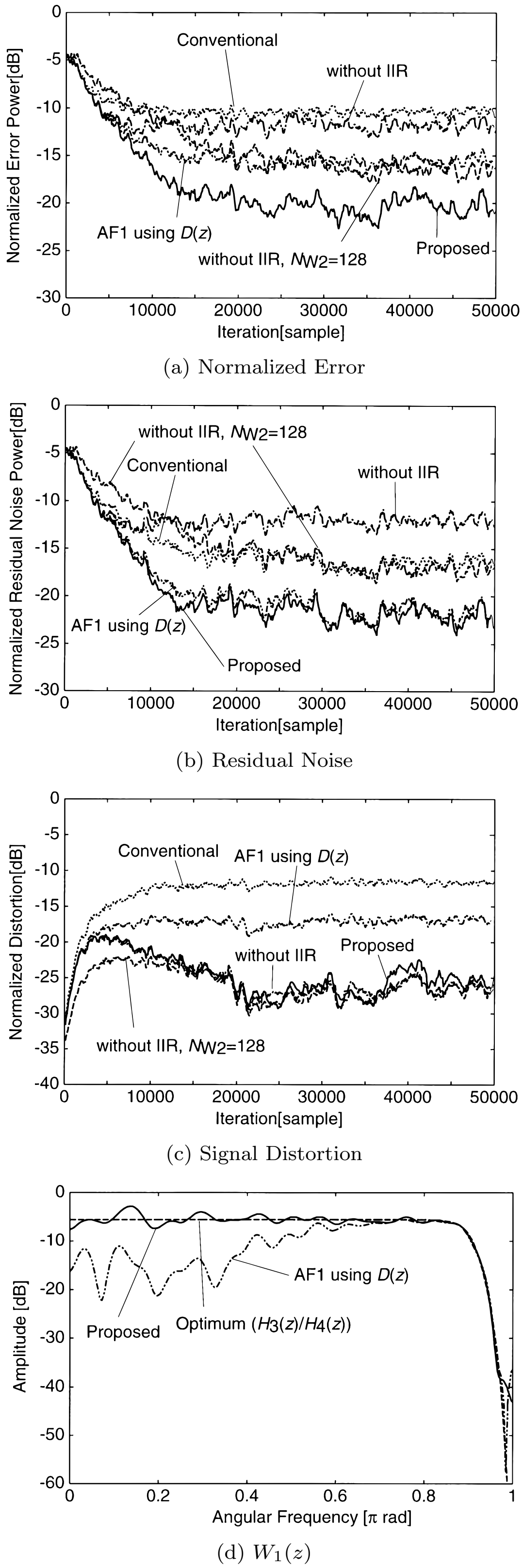


Fig. 10 Simulations for white-noise input.

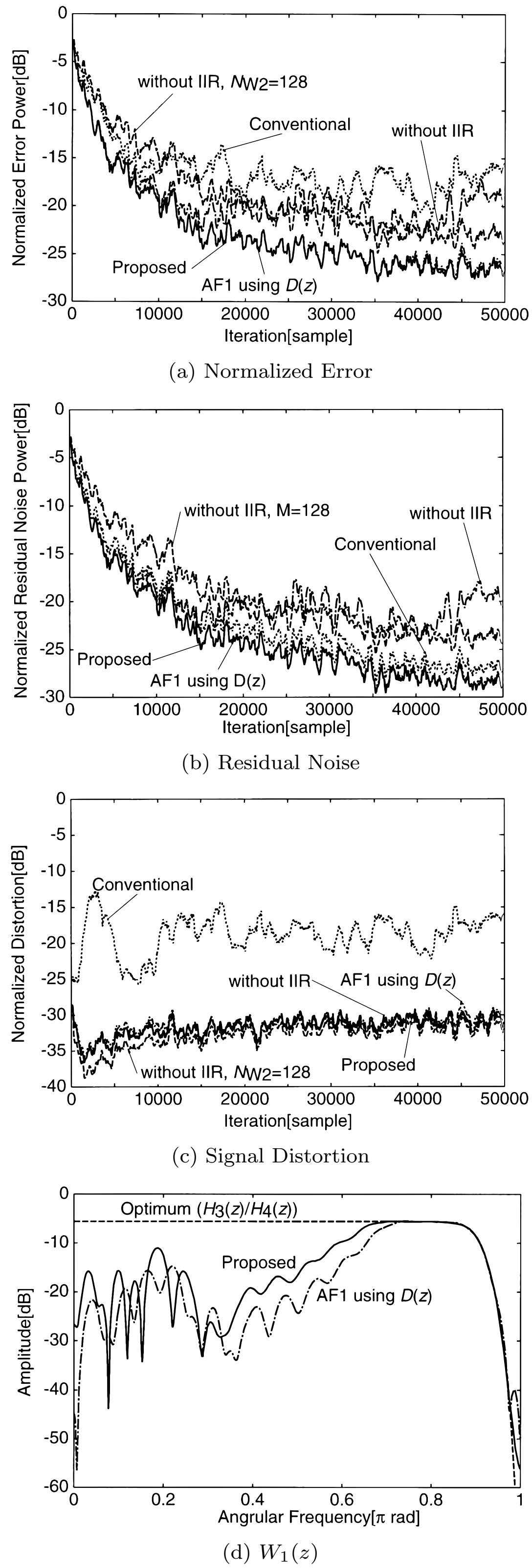


Fig. 11 Simulations for colored input.

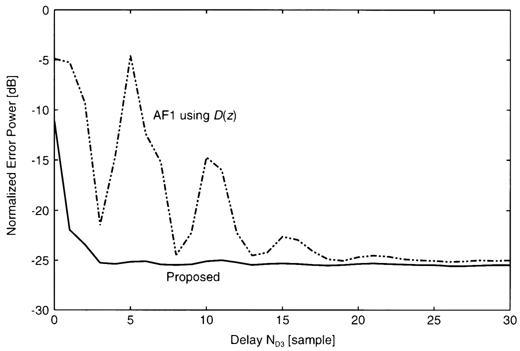


Fig. 12 Relation between N_{D3} and error power.

cellation of the noise. A recursive filter reduces the number of computations for noise cancellation stage. Separate filters for the adaptation and the filtering are introduced for desired signal cancellation. Computer simulations show that the proposed noise canceller improves the error power by almost 10 dB. The stability of the recursive filter, the uniqueness of the adaptive filters, evaluation in real environments should be future study.

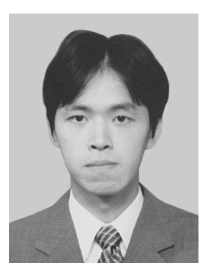
References

[1] B. Widrow and S.D. Stearns, "Adaptive noise canceling: Principles and applications," Proc. IEEE, vol.63, no.12, pp.1692–1716, Dec. 1975.
[2] M.J. Al-Kindi and J. Dunlop, "A low distortion adaptive noise cancellation structure for real time applications," Proc. ICASSP '87, pp.2153–2156, 1987.
[3] G. Mirchandani, R.L. Zinser, Jr., and J.B. Evans, "A new adaptive noise cancellation scheme in the presence of crosstalk," IEEE Trans. CAS–II, vol.39, no.10, pp.681–694, Oct. 1992.
[4] H. Kubota, T. Furukawa, and H. Itakura, "Pre-processed noise canceller design and its performance," IEICE Trans. vol.J69-A, no.5, pp.584–591, May 1986.
[5] J. Dunlop and M.J. Al-Kindi, "Application of adaptive noise cancelling to diver voice communications," Proc. ICASSP '87, pp.1708–1711, 1987.
[6] T. Taniguchi, Y. Tsukahara, T. Obara, and S. Minami, "A study on reducing distortion for acoustic noise canceller," Proc. IEICE Fall Conf., vol.A, p.126, 1994.
[7] V. Parsa, P.A. Parker, and R.N. Scott, "Performance analysis of a crosstalk resistant adaptive noise canceller," IEEE Trans. CAS–II, vol.43, no.7, pp.473–482, 1996.
[8] S. Ikeda and A. Sugiyama, "An adaptive noise canceller with low signal distortion for speech codecs," IEEE Trans. SP, vol.43, no.3, pp.665–674, March 1999.
[9] J. Nagumo and A. Noda, "A learning method for system identification," IEEE Trans. AC, vol.12, no.3, pp.282–287, March 1967.

Akihiro Hirano received the B.Eng., M.Eng. and Dr.Eng. degrees from Kanazawa University, Kanazawa, Japan in 1987, 1989 and 2000, respectively. He joined NEC Corporation, Kawasaki, Japan in 1989, where he had been a Research Engineer in Research and Development Group. Since 1998, he is a Research Associate at Faculty of Engineering, Kanazawa University. He has been engaged in researches on adaptive signal processing and neural networks. He was awarded the 1995 Academic Encouragement Award by IEICE. Dr. Hirano is a member of IEEE.

Kenji Nakayama received the B.Eng. and Dr.Eng. degrees in Electronics Engineering from Tokyo Institute of Technology (TIT), Tokyo, Japan, in 1971 and 1983, respectively. From 1971 to 1972 he was engaged in the research on classical network theory in TIT. He was involved in NEC Corporation from 1972 to 1988, where his research subjects were filter design methodology and signal processing algorithms. He joined the Department of Electrical and Computer Engineering at Kanazawa University, in August 1988, where he is currently a Professor. His current research interests include neural networks, adaptive signal processing, and signal theory. Dr. Nakayama is a senior member of IEEE and a member of INNS.

Shinya Arai received the B.Eng. degree from Kanazawa University, Kanazawa, Japan in 2000. He is currently a student of Graduate School of Natural Science & Technology, Kanazawa University. Mr. Arai has been engaged in researches on adaptive interference cancellation.



Masaki Deguchi received the B.Eng. degree from Kanazawa University, Kanazawa, Japan in 1999. He joined Fujitsu Hokuriku Communication Systems Limited, Kanazawa, Japan in April 1999, where he is a software engineer. He has been engaged in developments of software for telecommunication systems.