

# An Adaptive Noise Canceller with Low Signal Distortion for Speech Codecs

Shigeji Ikeda and Akihiko Sugiyama, *Member, IEEE*

**Abstract**— This paper proposes an adaptive noise canceller (ANC) with low signal distortion for speech codecs. The proposed ANC has two adaptive filters: a main filter (MF) and a subfilter (SF). The signal-to-noise ratio (SNR) of input signals is estimated using the SF. To reduce signal distortion in the output signal of the ANC, a step size for coefficient update in the MF is controlled according to the estimated SNR. Computer simulation results using speech and diesel engine noise recorded in a special-purpose vehicle show that the proposed ANC reduces signal distortion in the output signal by up to 15 dB compared with a conventional ANC. Results of subjective listening tests show that the mean opinion scores (MOS's) for the proposed ANC with and without a speech codec are one point higher than the scores for the conventional ANC.

## I. INTRODUCTION

EXTRACTING a desired speech signal from noisy speech corrupted by additive noise is an important problem in digital voice communication systems. The background noise encountered in such environments as fighter jets, helicopters, tanks, and automobiles affects the performance of narrowband speech codecs based on the linear predictive coding (LPC) technique [1], [2]. For code excited linear predictive (CELP) coders with vector quantization (VQ), the intelligibility of coded speech is most significantly reduced by the presence of the additive noise. This is caused by the VQ codebooks, which are designed for noise-free speech.

There have been a number of single-microphone approaches for noise cancellation based on speech-enhancement techniques such as Kalman filtering [3], [4]. These methods are applicable even when only a noise-corrupted signal is available. However, the performance of noise reduction is degraded for a low signal-to-noise ratio (SNR) below 0 dB.

Adaptive noise cancellation [5]–[15] is another approach for powerful noise reduction. In an adaptive noise canceller (ANC), there are two microphones: the primary microphone to obtain the noise-corrupted speech and the reference microphone to obtain only a correlated component of the noise present in the primary microphone. The noise in the reference microphone is processed by an adaptive filter to generate a replica of the noise component in the primary input.

Manuscript received October 15, 1996; revised December 23, 1997. The associate editor coordinating the review of this paper and approving it for publication was Dr. Ronald D. DeGroat.

S. Ikeda was with the Communication System Engineering Department, Radio Application Division, NEC Corporation, Tokyo, Japan. He is now with the General Audio Division, Personal A&V Products Company, Sony Corporation, Tokyo, Japan.

A. Sugiyama is with the Signal Processing Research Laboratory, C&C Media Research Laboratories, NEC Corporation, Kawasaki, Japan.

Publisher Item Identifier S 1053-587X(99)01332-X.

Although the ANC is an effective technique for lower SNR signals, the quality of the processed speech may be degraded by uncorrelated additive noise components, signal distortion, and reverberation. The uncorrelated additive noise components appear, in an equivalent model, at the output of the primary and the reference microphone. The uncorrelated noise at the output of the primary microphone cannot be canceled by the ANC and remains at the output terminal. It serves as an interference for adaptation of the ANC. The other uncorrelated noise has the same effect on ANC adaptation. As a result, performance of the ANC is limited [5]. In the application that is dealt with this paper, the ratio of the speech to the above uncorrelated noise components is generally over 40 dB, which is true for most of ANC's applied to speech communications. Therefore, performance limitation comes from the other two factors: distortion and reverberation.

The signal distortion is caused when the reference signal includes a correlated component of the speech in the primary microphone, which is called "crosstalk." The crosstalk makes the adaptive filter partially cancel the speech in the primary microphone, resulting in distortion. The reverberation is caused by misadjustment errors in the adaptive filter. A large number of taps may cause large misadjustment errors, which lead to audible reverberation in the processed speech due to the feedback nature of the speech through adaptation of the adaptive filter [14]. Although there have been a numerous studies [7]–[13] on the crosstalk problem, they cannot resolve the reverberation problem.

Another important key to ANC's is reduction of the convergence time. For this purpose, a fast convergence algorithm for ANC's has been reported [15]. This algorithm estimates the speech power and subtracts it from the error signal to obtain a better estimate of misadjustment power. The step size for coefficient update is controlled according to the estimate of the misadjustment power. Thanks to the estimation of the speech power, this algorithm can achieve fast convergence. The step-size control in this algorithm is effective for reduction of the reverberation since the step size is not affected by the speech power. However, audible distortion can still be detected in the beginning of an utterance. Since the reverberation has as significant an influence on the coded-speech quality as the background noise, it must always be reduced as much as possible.

This paper proposes a new ANC that reduces the reverberation in the processed speech continuously. In Section II, the basic concept of the ANC's with the normalized LMS (NLMS) algorithm [16] and the adaptive step-size algorithm

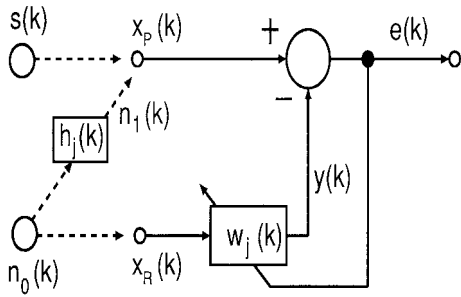


Fig. 1. Block diagram of classical ANC.

[15] are reviewed. Section III explains the proposed ANC in detail. Finally, in Section IV, experimental results including computer simulations and listening tests are shown.

## II. ADAPTIVE NOISE CANCELLATION

### A. Principle of ANC

Fig. 1 shows a block diagram of the classical ANC.  $s(k)$ ,  $n_0(k)$ , and  $n_1(k)$  are the signal, the noise, and the noise component in the primary microphone, all with a time index  $k$ .  $h_j(k)$  is the impulse response of the noise path from the noise source to the primary microphone. The primary signal  $x_P(k)$  and the reference signal  $x_R(k)$  can be written as

$$\begin{aligned} x_P(k) &= s(k) + n_1(k) \\ &= s(k) + \sum_{j=0}^{N-1} h_j(k) \cdot n_0(k-j) \end{aligned} \quad (1)$$

$$x_R(k) = n_0(k) \quad (2)$$

where  $N$  is the number of taps. The output of the ANC  $e(k)$  is given by

$$e(k) = s(k) + n_1(k) - y(k) \quad (3)$$

$$y(k) = \sum_{j=0}^{N-1} w_j(k) \cdot n_0(k-j) \quad (4)$$

where  $y(k)$  is the output of the adaptive filter, and  $w_j(k)$  is the  $j$ th filter coefficient of the adaptive filter.

Assuming that the noise path  $h_j(k)$  is estimated with the NLMS algorithm [16], the update of  $w_j(k)$  is given by

$$w_j(k+1) = w_j(k) + \frac{\mu \cdot e(k) \cdot n_0(k-j)}{\sum_{i=0}^{N-1} n_0^2(k-i)} \quad (0 \leq j \leq N-1) \quad (5)$$

where  $\mu$  is a step size. From (3), it can be seen that the error signal  $e(k)$  is nearly equal to  $s(k)$  when  $y(k)$  approximates  $n_1(k)$  after convergence. Therefore,  $w_j(k)$  is updated according to  $s(k)$ , resulting in signal-distortion such as the reverberation [14]. This distortion can be removed if the coefficient update is performed only in the absence of the speech  $s(k)$ . However, it requires an accurate speech detector. An alternative is to use a small step size  $\mu$  so that reverberation is kept minimum. An adaptive step-size algorithm should be

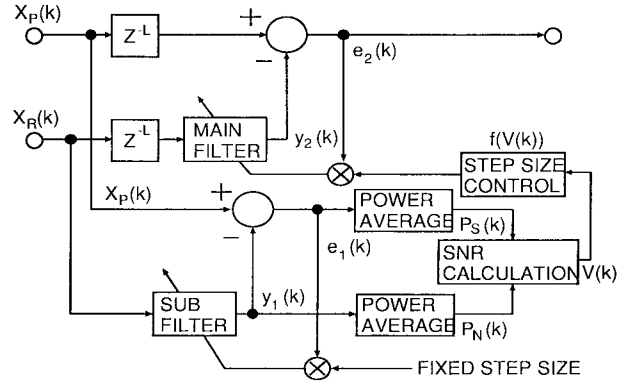


Fig. 2. Block diagram of proposed ANC.

considered to avoid slow convergence associated with the small  $\mu$ .

### B. Adaptive Step-Size Algorithm [14]

The estimation of speech power is an effective approach to reduce the influence of the speech. An adaptive step-size algorithm for fast convergence with a small misadjustment has been proposed [15]. This algorithm estimates the speech power and subtracts it to obtain a better estimate of misadjustment power for adaptation. The step size  $\Delta(k)$  is controlled by the estimate of the misadjustment power  $q(k)$  as

$$p(k) = \delta e^2(k) + (1 - \delta)p(k-1) \quad (6)$$

$$q(k) = \beta(e^2(k) - p(k)) + (1 - \beta)q(k-1) \quad (7)$$

$$\Delta(k) = q(k) \cdot \gamma \quad (8)$$

where  $\delta$  and  $\beta$  are leakage factors for first-order integrations in (6) and (7), and  $\gamma$  is a constant that determines a final misadjustment. Because the noise component is more stationary than speech,  $p(k)$  in (6) can be interpreted as a speech-power estimate extracted by averaging operation with a small time constant. The first term in the right-hand side of (7) gives an estimated power of the instantaneous error, which is then averaged with a large time constant. The averaged error power  $q(k)$  is scaled by  $\gamma$  so that it gives a good time-varying step size.

The  $j$ th filter coefficient of the adaptive filter is updated by

$$w_j(k+1) = w_j(k) + \Delta(k) \cdot e(k) \cdot n_0(k-j) \quad (0 \leq j \leq N-1) \quad (9)$$

where  $N$  is the number of taps for the adaptive filter. Thanks to the step-size adaptation, this algorithm can achieve fast convergence. However, audible distortion can be detected in the beginning of an utterance. This is because the estimation of the speech power is incorrect due to large leakage factors for the integrations.

## III. PROPOSED NOISE CANCELLER

The proposed noise canceller shown in Fig. 2 has two adaptive filters (MF and SF) that are operated in parallel to generate a noise replica. The SF is used to estimate the SNR of the primary input signal. The step size for the MF is controlled based on the estimated SNR.

### A. SNR Estimation by Subfilter

The SF works in the same way as the classical ANC. Filter coefficients are updated by the NLMS [16]. The step size  $\mu_{\text{SUB}}$  is set large and fixed for fast convergence and rapid tracking of the noise-path change. A small value of  $\mu_{\text{SUB}}$  results in more precise estimation; however, it is over the specifications required for a rough estimate of the SNR. For the estimation of SNR, an average power of the noise replica  $P_N(k)$  and the power of the error signal  $P_S(k)$  are calculated by

$$P_N(k) = \sum_{j=0}^{M-1} y_1^2(k-j) \quad (10)$$

$$\begin{aligned} P_S(k) &= \sum_{j=0}^{M-1} \{x_P(k-j) - y_1(k-j)\}^2 \\ &= \sum_{j=0}^{M-1} e_1^2(k-j) \end{aligned} \quad (11)$$

where  $y_1(k)$  and  $x_P(k)$  are the SF output and the primary signal.  $M$  is the number of samples used for estimation of  $P_N(k)$  and  $P_S(k)$ . From  $P_S(k)$  and  $P_N(k)$ , the SNR  $V(k)$  of the primary signal is calculated by

$$V(k) = 10 \log_{10} \{P_S(k)/P_N(k)\} \quad [\text{dB}]. \quad (12)$$

Implementation complexity for calculating an estimated SNR  $V(k)$  largely depends on its own implementation. As an example, use of a 32-bit floating-point digital signal processor (DSP) (the ADSP-21 020 [6]) is assumed, as it will actually be used in implementation. The division is performed with seven steps. A log of the quotient and the scaling by 10 are implemented by table look-up, which requires three steps. Therefore, the number of steps for calculation of (12) is 10. For a case where 64 taps are necessary for both the MF and the SF, as is the case in Section IV,  $64 \times 2 \times 2 = 256$  steps are spent on these adaptive filters. This means that the increase in computation is less than 4% of that used for adaptive filtering only. It should be noted that there are other steps to be consumed for the rest of the ANC. Therefore, the actual increase can be smaller than 4% of the total number of steps.

### B. Step-Size Control in Main Filter

The step size in the MF is controlled by the estimated SNR  $V(k)$ . If  $V(k)$  is low, the step size is set large for fast convergence because low  $V(k)$  means small disturbance for coefficient adaptation. Otherwise, the step size is set small for less signal distortion in the ANC's output. The following equation shows the function that determines the step size  $\mu_{\text{MAIN}}(k)$  based on  $V(k)$ .

$$\mu_{\text{MAIN}}(k) = \begin{cases} \mu_{\text{min}} & (V(k) > V_{\text{max}}) \\ f(V(k)) & (V_{\text{max}} \geq V(k) \geq V_{\text{min}}) \\ \mu_{\text{max}} & (V(k) < V_{\text{min}}) \end{cases} \quad (13)$$

where  $\mu_{\text{max}}$ ,  $\mu_{\text{min}}$ , and  $f(\cdot)$  are the maximum and the minimum step sizes and a function of  $V(k)$ , respectively. It is natural that  $f(\cdot)$  is a decreasing function since small step size

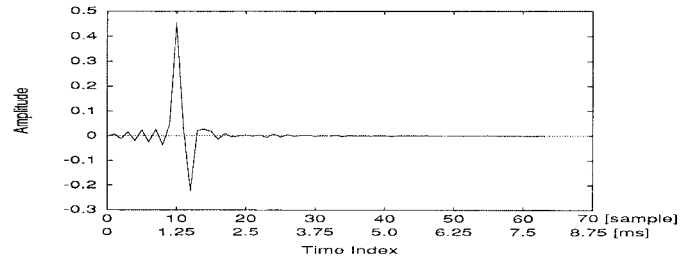


Fig. 3. Impulse response of noise path.

is suitable for a large SNR. For simplicity, let us assume that  $f(\cdot)$  is a first-order function of  $V(k)$ . Then, it may be given by

$$\mu_{\text{MAIN}}(k) = f(V(k)) = A \cdot V(k) + B \quad (14)$$

where  $A$  and  $B$  are constants.  $\mu_{\text{min}}$  determines the signal distortion in the utterance. If  $\mu_{\text{min}}$  is set to zero, the adaptation is skipped when  $V(k)$  is greater than  $V_{\text{max}}$ . In this case, the proposed algorithm works as the adaptation-stop method [7] with a speech detector.

### C. Delay Compensation for Main Filter

The estimated SNR  $V(k)$  is given with time delay, which depends on the number of samples  $M$  used for estimation of  $P_N(k)$  and  $P_S(k)$ . This time delay directly raises signal distortion in the processed speech because the step size remains large in the beginning of the utterance. To compensate for this delay, the delay unit  $z^{-L}$  is incorporated only in the MF side.  $L$  is set to  $M/2$  since the time delay is half of  $M$ .

## IV. EXPERIMENTAL RESULTS

Performance of the proposed ANC was evaluated by computer simulations in comparison with that of the conventional algorithm [15]. A diesel engine noise recorded in a special-purpose vehicle was used as a noise source. It is worthwhile to note that this recorded noise naturally has the uncorrelated noise component that exists in the acoustical environment used for the recording. Thus, the simulation results reflect the real situation in terms of the uncorrelated noise component. Fig. 3 shows a noise-path impulse response measured in a room with a dimension of 3.05 m (width)  $\times$  2.85 m (depth)  $\times$  1.80 m (height). In order to evaluate tracking capability, the polarity of the noise path was inverted at 12.5 s.<sup>1</sup> The noise component, which was generated by convolution of the noise source and the noise path, was added to the speech source to make the noise-corrupted signal. This speech contains the uncorrelated noise component, which should exist in the recording environment. The sampling frequency was 8 KHz.  $N, M, L, \mu_{\text{SUB}}, \mu_{\text{max}}, V_{\text{min}}, \mu_{\text{min}},$  and  $V_{\text{max}}$  are shown in Table I. A first-order function  $f(V(k))$  in Fig. 4 was used for step-size control.  $\beta$  and  $\delta$  of the conventional algorithm [15] were set to  $2^{-15}$  and  $2^{-14}$  for the fastest convergence, and  $\gamma$  was set to 40.0 so that the final misadjustment error is equivalent to that of the proposed algorithm.

<sup>1</sup>This is nothing more than an example. An abrupt polarity change was imposed as an extreme example. Path changes in the real environment are slower and less significant and, thus, easier to track.

TABLE I  
PARAMETERS FOR PROPOSED ANC

Parameter	Value
$N$	64
$M$	128
$L$	64
$\mu_{SUB}$	0.4
$\mu_{max}$	0.4
$V_{min}$	-50dB
$\mu_{min}$	$2^{-6}$
$V_{max}$	-10dB

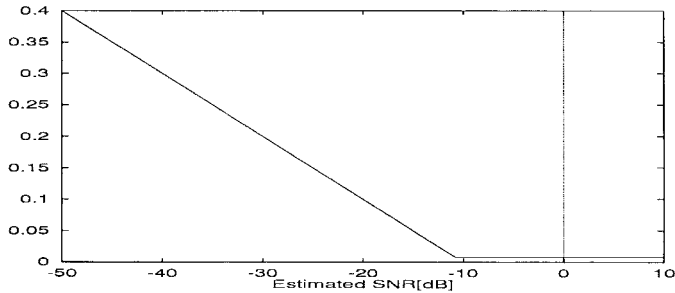


Fig. 4. Step-size function.

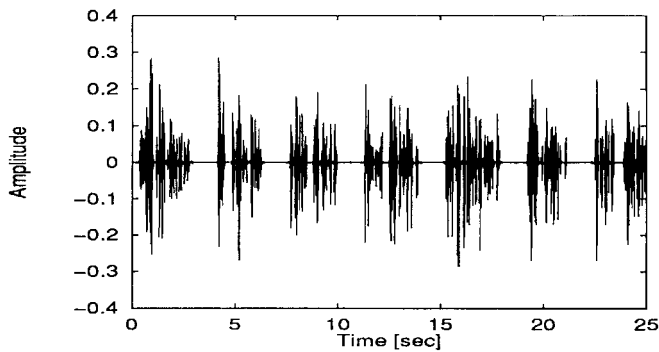


Fig. 5. Original speech.

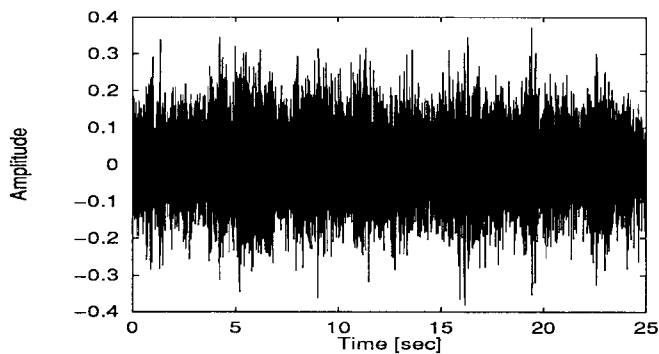


Fig. 6. Noise-corrupted speech.

#### A. Basic Characteristics

Figs. 5–7 illustrate the original speech, the noise-corrupted speech, and the noise-cancelled speech, respectively. The SNR in the primary signal was around 0 dB in the utterance. The proposed ANC successfully cancels the noise and tracks the noise-path change at 12.5 s.

The original SNR and the SNR estimated by the SF are compared in Figs. 8 and 9. Since the peaks of the estimated

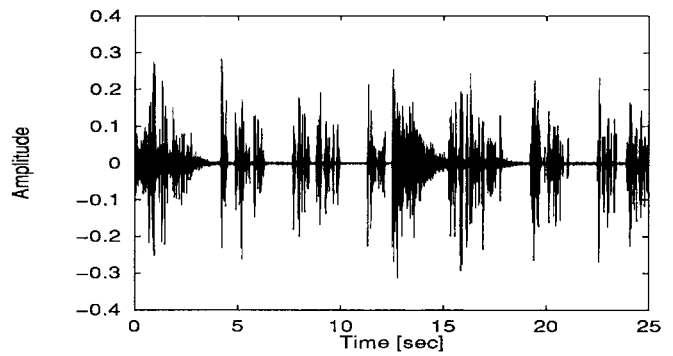


Fig. 7. Noise-cancelled speech.

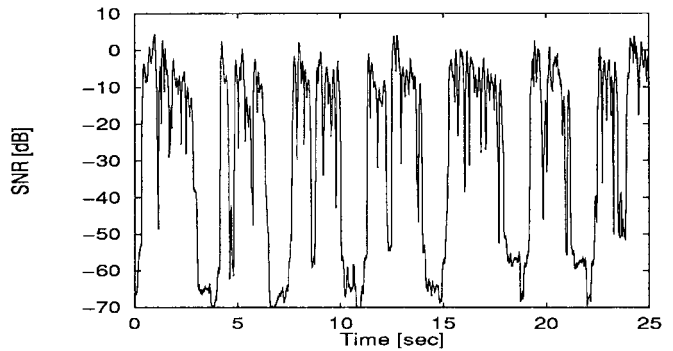


Fig. 8. Original SNR.

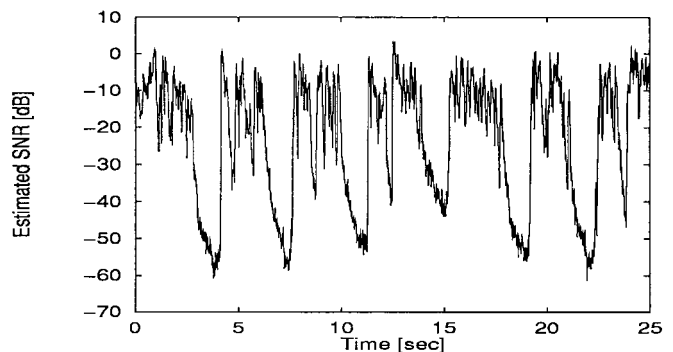


Fig. 9. Estimated SNR.

SNR well approximate those of the original SNR, it is considered reliable. Fig. 10 exhibits the step-size behaviors in the proposed ANC and in the adaptive step-size algorithm [15]. The step size in the proposed ANC remains small in the utterance, whereas the other becomes larger in the beginning of the utterance.

Fig. 11 shows the residual noise  $R(k)$  defined by

$$R(k) = 10 \log_{10} \{P_E(k)/P_X(k)\} \quad [\text{dB}] \quad (15)$$

$$P_E(k) = \sum_{j=0}^{Q-1} e_2^2(k-j) \quad (16)$$

$$P_X(k) = \sum_{j=0}^{Q-1} x_P^2(k-j) \quad (17)$$

where  $Q$ ,  $e_2(k)$ , and  $x_P(k)$  are the number of samples used for estimation of  $P_E(k)$  and  $P_X(k)$ , which are the output of the

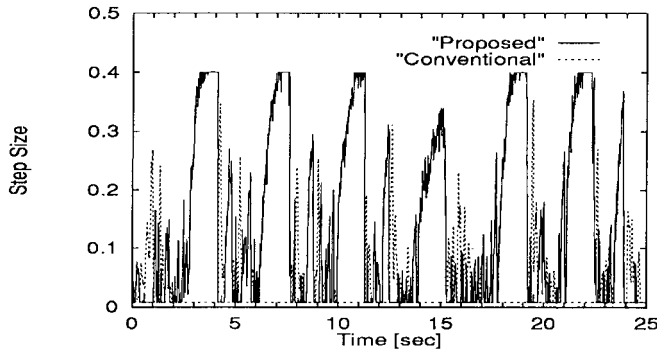


Fig. 10. Step-size behavior.

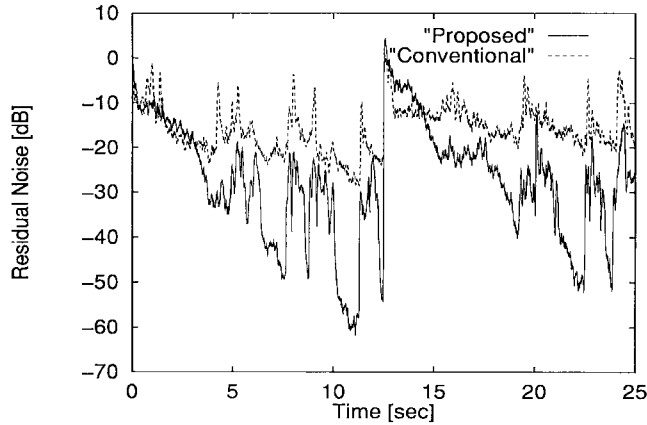


Fig. 11. Residual noise.

ANC and the primary signal, respectively.  $P_E(k)$  and  $P_X(k)$  were calculated with  $Q = 512$ . The proposed ANC reduces the residual noise by up to 10 dB in the utterance compared with the conventional algorithm.

Fig. 12 gives the signal distortion  $D(k)$  in the output defined by

$$D(k) = 10 \log_{10} \{P_D(k)/P_O(k)\} \quad [\text{dB}] \quad (18)$$

$$P_D(k) = \sum_{j=0}^{Q-1} \{e_2(k-j) - s(k-j)\}^2 \quad (19)$$

$$P_O(k) = \sum_{j=0}^{Q-1} s^2(k-j) \quad (20)$$

where  $s(k)$  is the original speech. The proposed ANC reduces the signal distortion by up to 15 dB compared with the conventional algorithm in the utterance.

Fig. 13 shows the impulse response estimation ratio (IRER)  $ER(k)$  defined by

$$ER(k) = 10 \log_{10} \frac{\sum_{j=1}^N h_j^2(k)}{\sum_{j=1}^N \{h_j(k) - w_j(k)\}^2} \quad [\text{dB}] \quad (21)$$

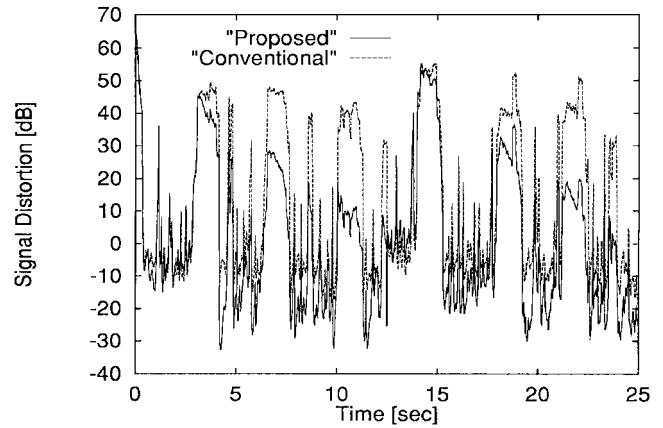


Fig. 12. Signal distortion of output.

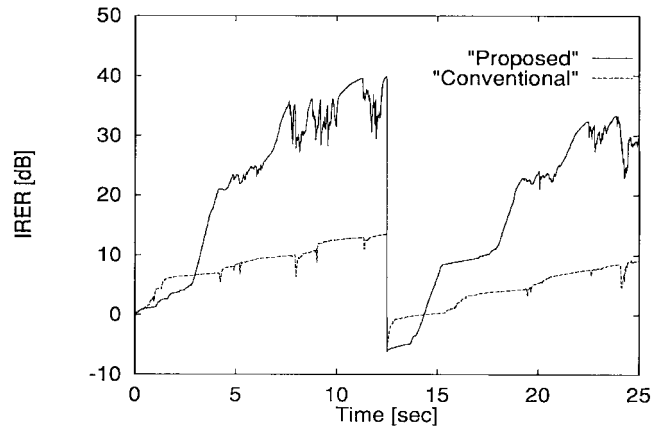


Fig. 13. IRER.

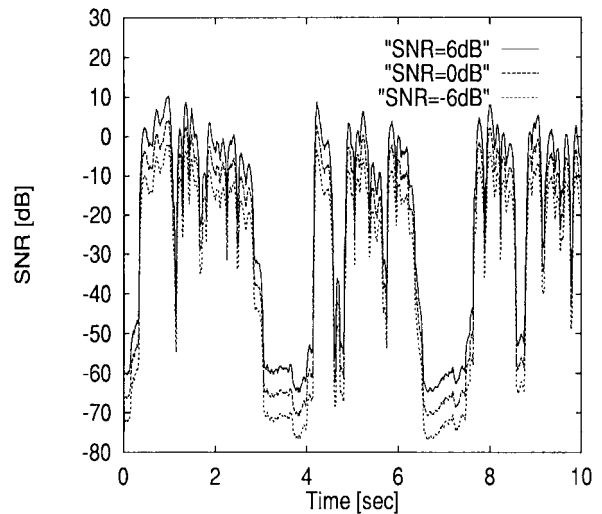


Fig. 14. Original SNR (SNR = 6 dB, 0 dB, -6 dB).

where  $h_j(k)$  and  $w_j(k)$  are the noise-path impulse response and the filter coefficient of the adaptive filter, respectively. The increase of IRER in the proposed ANC is slower than that in the conventional ANC because the step size in the proposed ANC remains small during the utterance. However, a large step size during the silent intervals speeds up the IRER

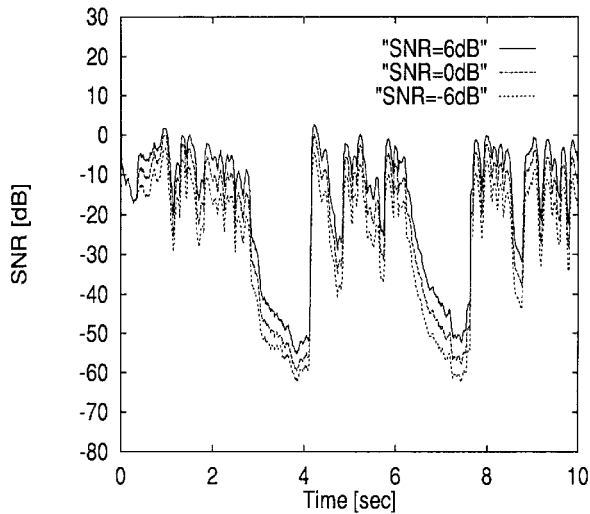


Fig. 15. Estimated SNR (SNR = 6 dB, 0 dB, -6 dB).

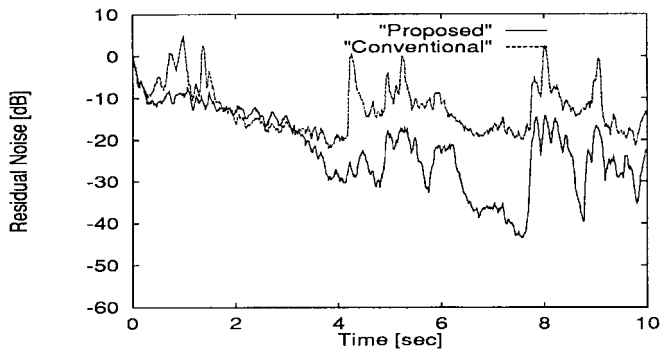


Fig. 16. Residual noise (SNR = 6 dB).

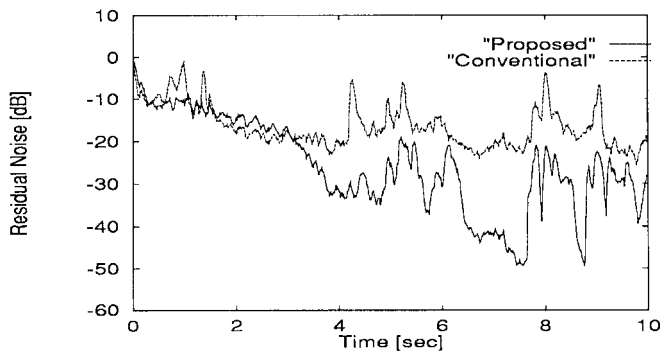


Fig. 17. Residual noise (SNR = 0 dB).

growth. The final IRER in the proposed ANC is larger by 30 dB than that in the conventional ANC.

*B. Robustness Against Different SNR's*

To demonstrate robustness of the proposed ANC, three different SNR values of the input signal were evaluated. These evaluations were carried out using 10 s of signals without the noise-path change. Fig. 14 illustrates the original SNR in the three cases where SNR in the utterance was set to 6, 0, and -6

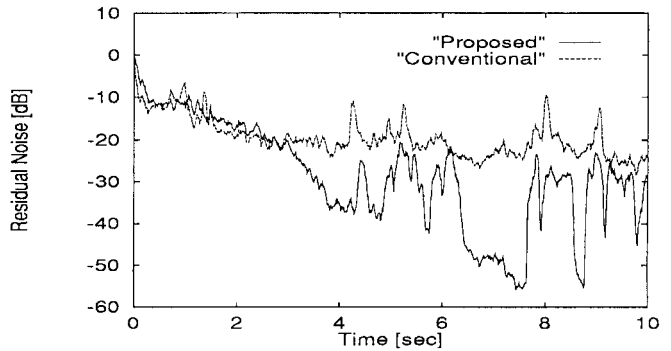


Fig. 18. Residual noise (SNR = -6 dB).

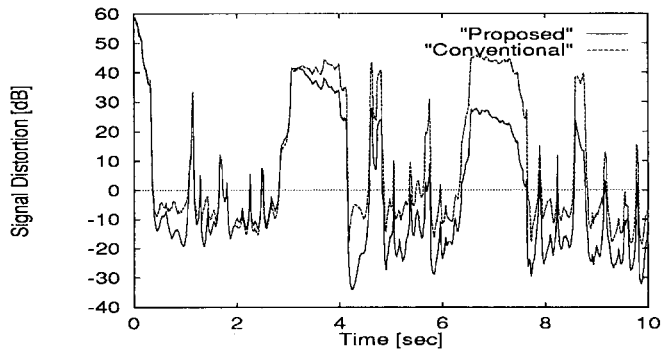


Fig. 19. Signal distortion (SNR = 6 dB).

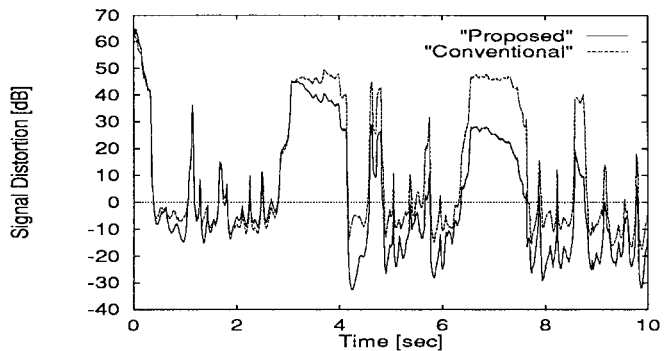


Fig. 20. Signal distortion (SNR = 0 dB).

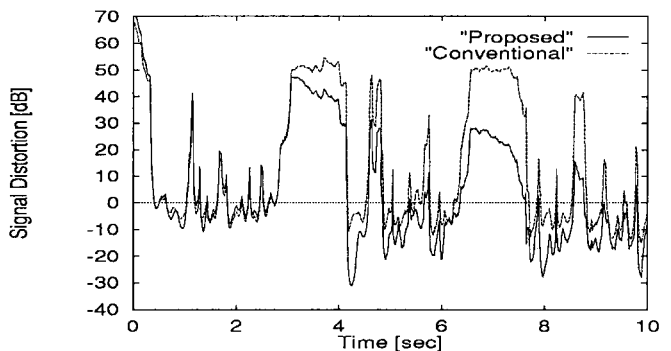


Fig. 21. Signal distortion (SNR = -6 dB).

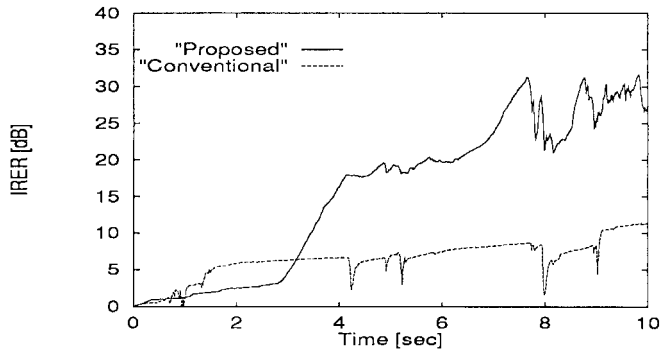


Fig. 22. IRER (SNR = 6 dB).

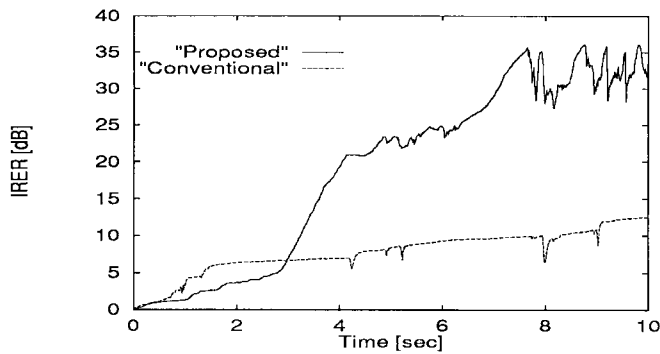


Fig. 23. IRER (SNR = 0 dB).

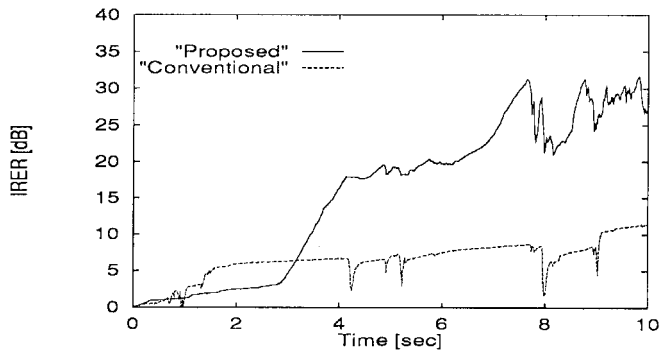


Fig. 24. IRER (SNR = -6 dB).

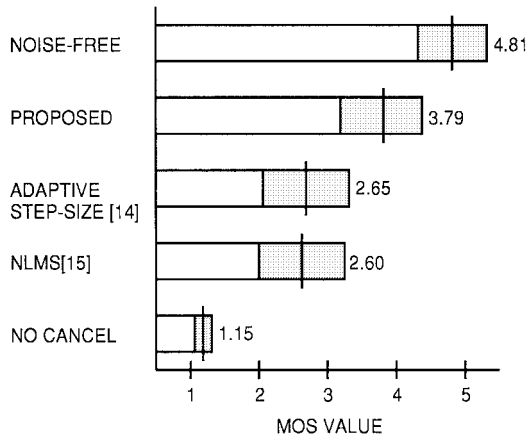


Fig. 25. MOS result in Case I (no coded speech).

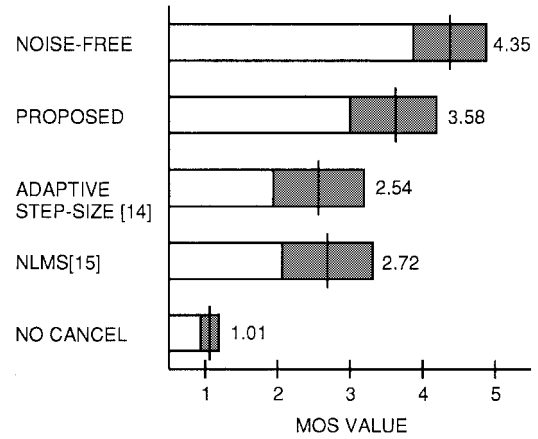


Fig. 26. MOS result in Case II (coded speech).

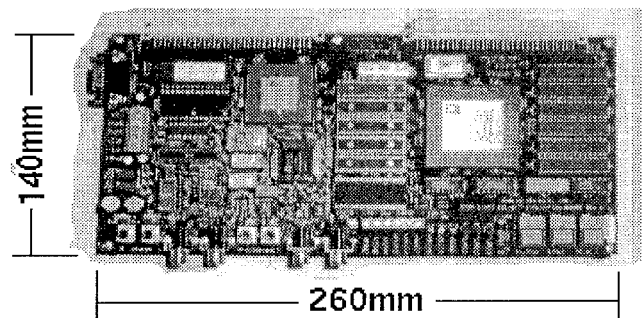


Fig. 27. Implemented ANC.

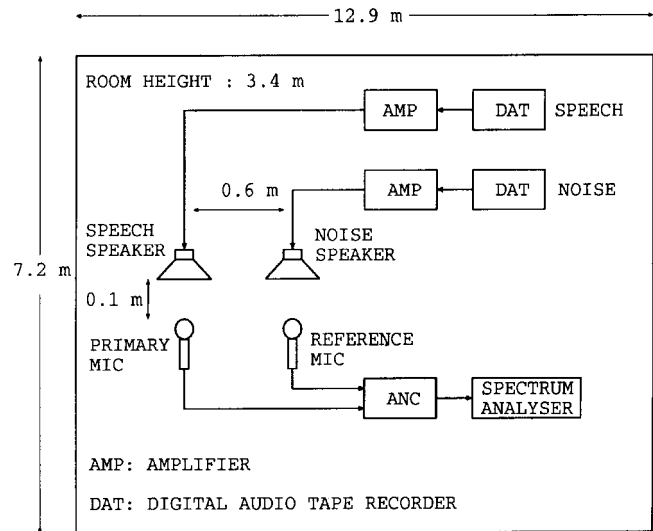


Fig. 28. Layout of microphones and speakers.

dB. The estimated SNR by the SF in the three cases is given in Fig. 15. The SF well estimates the original SNR. Figs. 16–18 and 19–21 show the residual noise and the signal distortion. The performance in the proposed ANC does not change for different SNR's. The IRER's are shown in Figs. 22–24. The IRER's of the proposed ANC reach 40 dB, whereas those of the conventional ANC achieve 12 dB. The proposed ANC

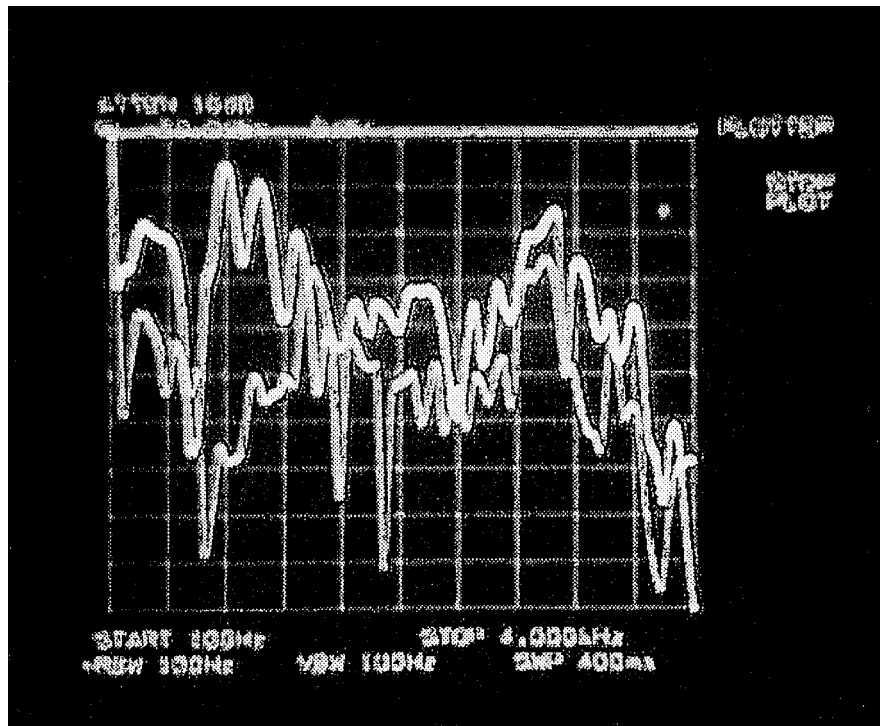


Fig. 29. Power spectra of noise-corrupted speech and noise-cancelled speech (diesel-engine noise).

achieves a good IRER independent of the SNR demonstrating its robustness.

### C. Subjective Listening Test

To evaluate the subjective performance of the proposed ANC, a listening test was carried out. The NLMS [16] algorithm, the adaptive step-size algorithm [15], and the proposed algorithm were compared for the case where the SNR in the primary signal is 0 dB. The mean opinion scores (MOS's) by 20 listeners were evaluated. The MOS tests were performed in two cases. In the first case (Case I) and the second case (Case II), samples of the ANC's output and samples of the coded speech of the ANC's output were evaluated, respectively. In both cases, a noise-free speech sample and a noisy speech sample before noise cancellation were included as the highest and the lowest anchors. The same speech source as in Sections IV-A and B was employed for the subjective test. A 4-kb/s CELP [17] was used as the algorithm of the speech codec. Figs. 25 and 26 show the MOS results in Cases I and II. The vertical line centered in the shaded area and the numeral are the mean value of the MOS. The width of the shaded area corresponds to the standard deviation. In both cases, the mean values of the MOS for the proposed ANC are higher than the adaptive step-size algorithm and the NLMS by about one point.

## V. IMPLEMENTATION BY DSP

The proposed ANC has been implemented by using a 32-bit floating-point DSP: the ADSP-21 020 [6]. Fig. 27 is a picture of the implemented ANC that is controlled by a personal computer. The number of taps of the MF and the SF are both

up to 460 with a sampling frequency of 8 KHz. Sixteen-bit analog-to-digital (A/D) and digital-to-analog (D/A) converters are used.

The performance of the implemented ANC was evaluated in a meeting room with a dimension of 12.9 m (width)  $\times$  7.2 m (depth)  $\times$  3.4 m (height). As a personal computer controls the ANC board, there was noise generated by its cooling fan. This noise served as the uncorrelated noise components for the primary and the reference microphone. Room noise was masked by this noise and was negligible. Fig. 28 illustrates the layout of microphones and speakers in the room. Two one-directional microphones were used as the primary and the reference microphones. These microphones were located in parallel 0.6 m apart. Two speakers were used for speech and noise generation. The speaker for the speech and that for the noise were located in front of the primary microphone and in front of the reference microphone, respectively. Both of the distances between the primary microphone and the speech speaker and between the reference microphone and the noise speaker were 0.1 m. There were uncorrelated noise components and *crosstalk* in this environment.

Fig. 29 compares the power spectra of the noise-corrupted speech with a diesel-engine noise and that of the noise-cancelled speech measured by a spectrum analyzer. The horizontal axis is frequency with 400 Hz per scale. The vertical axis is power spectrum with 5 dB per scale. The upper line is the power spectrum of the noise-corrupted speech, and the other is that of the noise-cancelled speech. The significant peaks of the power spectrum for the diesel-engine noise are between 800 and 1200 Hz. From this figure, it is apparent that the noise spectra for this frequency range are reduced by up to 30 dB, and that for the other frequency range are



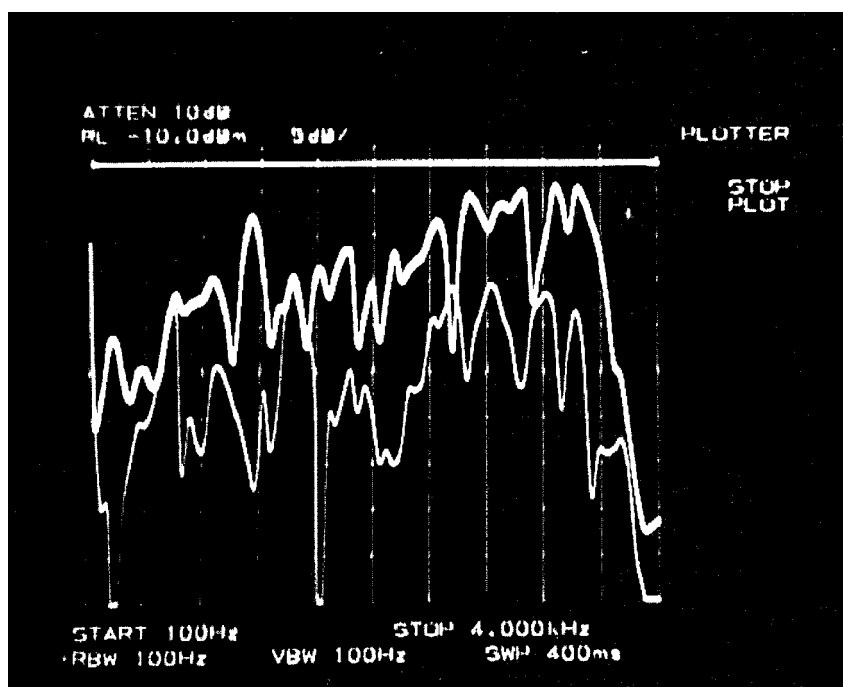


Fig. 30. Power spectra of noise-corrupted speech and noise-cancelled speech (white Gaussian noise).

cancelled by 10 dB. Fig. 30 illustrates the power spectra in a white-Gaussian noise case. In this case, the noise spectra are uniformly reduced by up to 10 dB for the entire frequency range. Since the room used for the experiment is echoic, average noise reduction is not as good as that in the simulation results presented in the previous section.

## VI. CONCLUSION

A new adaptive noise canceller (ANC) with low signal distortion for speech codecs has been proposed. It has two adaptive filters: a main filter (MF) and a subfilter (SF). The SNR of the input signals is estimated using the SF. The step size for coefficient update in the MF is controlled by the estimated SNR. Results of computer simulations show that the proposed ANC reduces signal distortion by up to 15 dB compared with the conventional ANC. Results of subjective listening tests show that the mean opinion scores (MOS's) for the proposed ANC are one point higher than the conventional ANC for the coded speech and the noncoded speech at the ANC's output.

## ACKNOWLEDGMENT

The authors would like to thank N. Abe and S. Kondo for their help in the experiment and to J. Takizawa for providing the speech-codec program; they are all with the Communications Systems Engineering Department, Radio Application Division, NEC Corporation. They are also indebted to Dr. K. Ozawa, Senior Manager of Signal Processing Research Laboratory, Information Technology Research Laboratories, and Dr. T. Nishitani, Deputy General Manager of Information Technology Research Laboratories, for their guidance and encouragement. This research was initiated by T. Taguchi, Senior Manager, and has been managed by T. Ishikawa, Engineering

Manager, both of Communications Systems Engineering Department, Radio Application Division, NEC Corporation. Their contributions are also acknowledged.

## REFERENCES

- [1] M. R. Sambur and N. S. Jayant, "LPC analysis/synthesis from speech inputs containing quantization noise or additive white noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, pp. 488-494, 1976.
- [2] C. F. Teacher and D. Coulter, "Performance of LPC vocoders in a noisy environment," in *Proc. IEEE ICASSP*, 1979, pp. 216-219.
- [3] K. K. Paliwal and A. Basu, "A speech enhancement method based on Kalman filtering," in *Proc. IEEE ICASSP*, 1987, pp. 177-180.
- [4] J. D. Gibson, B. Koo, and S. D. Gray, "Filtering of colored noise for speech enhancement and coding," *IEEE Trans. Signal Processing*, vol. 39, pp. 1732-1742, 1991.
- [5] B. Widrow *et al.*, "Adaptive noise cancelling: Principles and applications," *Proc. IEEE*, vol. 63, pp. 1692-1716, Dec. 1975.
- [6] "ADSP-21020/21010 User's Manual," Analog Devices, Inc., 1993.
- [7] W. A. Harrison, J. S. Lim, and E. Singer, "A new application of adaptive noise cancellation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 21-27, Jan. 1986.
- [8] M. J. Al-Kindi and J. Dunlop, "A low distortion adaptive noise cancellation structure for real time applications," in *Proc. IEEE ICASSP*, 1987, pp. 49.15.1-49.15.4.
- [9] J. Dunlop and M. J. Al-Kindi, "Application of adaptive noise cancelling to diver voice communication," in *Proc. IEEE ICASSP*, 1987, pp. 40.6.1-40.60.4.
- [10] H. Kubota, T. Furukawa, and H. Itakura, "Pre-processed noise canceller design and its performance," *IEICE Trans.*, vol. J69-A, pp. 584-591, 1986 (in Japanese).
- [11] T. Taniguchi, Y. Tsukahara, T. Obara, and S. Minami, "A study on reducing distortion for acoustic noise canceller," in *Proc. IEICE Fall Conf.*, 1994, pp. 126 (in Japanese).
- [12] G. Mirchandani, R. L. Zinsler, and J. B. Evans, "A new adaptive noise cancellation scheme in the presence of crosstalk," *IEEE Trans. Circuits Syst.*, vol. 39, pp. 681-694, 1992.
- [13] V. Parsa, P. A. Parker, and R. N. Scott, "Performance analysis of a crosstalk resistant adaptive noise canceller," *IEEE Trans. Circuits Syst.*, vol. 43, pp. 473-482, 1996.
- [14] S. F. Boll and D. C. Pulsipher, "Suppression of acoustic noise in speech using two microphone adaptive noise cancellation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 752-753, 1980.

- [15] A. Sugiyama, M. N. S. Swamy, and E. I. Plotkin, "A fast convergence algorithm for adaptive FIR filters," in *Proc. IEEE ICASSP*, 1989, pp. 892–895.
- [16] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction and Control*. Englewood Cliffs, NJ: Prentice-Hall, 1985.
- [17] K. Ozawa, M. Serizawa, T. Miyano, T. Nomura, M. Ikekawa, and S. Taumi, "M-LCELP speech coding at 4kb/s with multi-mode and multi-codebook," *IEICE Trans. Commun.*, vol. E77-B, pp. 1114–1121, 1994.



**Shigeji Ikeda** was born in Hokkaido, Japan, on September 10, 1962. He received the B.E. degree in electrical engineering from Hokkaido University, Sapporo, Japan, in 1985.

He joined the Communications Systems Engineering Department, Radio Application Division, NEC Corporation, Tokyo, Japan, in 1985, where he had been engaged in research and development on speech signal processing, especially from narrow and medium bands coding and adaptive signal processing. Since 1998, he has been an Assistant

Manager of the Development Department, General Audio Division, Personal A&V Products Company, Sony Corporation, Tokyo, Japan.

Mr. Ikeda is a member of the Institute of Electronics, Information, and Communication Engineers of Japan.



**Akihiko Sugiyama** (M'85) received the B.Eng., M.Eng., and Dr.Eng. degrees in electrical engineering from Tokyo Metropolitan University, Tokyo, Japan, in 1979, 1981, and 1998, respectively.

He joined NEC Corporation, Kawasaki, Japan, in 1981 and has been engaged in research on signal processor applications to transmission terminals, subscriber loop transmission systems, adaptive filter applications, and high-fidelity audio coding. In 1987, he was on leave at the Faculty of Engineering and Computer Science, Concordia University, Montreal, P.Q., Canada, as a Visiting Scientist. From 1989 to 1994, he was involved in the activities of the Audio Subgroup ISO/IEC JTC1/SC29/WG11 (known as MPEG/Audio) for international standardization of high-quality audio data compression as a member of the Japanese delegation. His current interests lie in the area of signal processing and circuit theory.

Dr. Sugiyama is a member of the Institute of Electronics, Information and Communication Engineers of Japan. He served as an Associate Editor for the *IEEE TRANSACTIONS ON SIGNAL PROCESSING* from 1994 to 1996. He is also a member of the Technical Committee for Audio and Electroacoustics. He is currently serving as an Associate Editor for the *Transactions of the Institute of Electronics, Information, and Communication Engineers* (IEICE), vol. A. He received the 1988 Shinohara Memorial Academic Encouragement Award by the Institute from Electronics, Information, and Communication Engineers of Japan. He is a coauthor of *International Standards for Multimedia Coding* (Yokohama, Japan: Maruzen, 1991) and *MPEG/International Standards for Multimedia Coding* (Yokohama, Japan: Maruzen, 1994) *MPEG* (Tokyo, Japan: Ohmusha, 1996), and *Digital Broadcasting* (Tokyo, Japan: Ohmusha, 1996).