A Three-Microphone Adaptive Noise Canceller for Minimizing Reverberation and Signal Distortion

Zayed M. Ramadan

Department of Network and Communications Engineering, College of Engineering and Information
Technology, Al Ain University of Science and Technology, P.O.Box 64141, Al Ain, UAE

Abstract: This paper introduces an adaptive noise canceller (ANC) to improve the system performance in the presence of signal leakage components. The proposed ANC consists of two adaptive filters and three microphones. The first adaptive filter cancels the signal leakage and the second filter cancels the noise. For best results, a least mean squares adaptive algorithm was also introduced and used in the proposed ANC. In this algorithm the step size was based on both error vector and data normalization. Simulation results, carried out using a real speech, demonstrate significant improvements of the proposed ANC over the conventional one in minimizing signal distortion and reverberation.

Keywords: crosstalk, noise canceller, reverberation, signal leakage.

INTRODUCTION

Adaptive noise cancellation techniques are usually applied in applications where a reference signal that is correlated with the noise at the primary signal is easily obtained. These applications include long distance telephone calls, adaptive antenna array processing, and adaptive line enhancement. Many two-microphone ANCs have been proposed in the literature of adaptive filtering using least mean-square (LMS)-based algorithms that alter the step-size of the update equation to improve the tracking ability of the algorithm and its speed of convergence as well[1-5].

A typical block diagram of a conventional adaptive noise canceller is shown in Fig.1, where the signal source provides the signal $S(n)$ which serves as the desired signal for the adaptive filter. The reference signal $v_2(n)$ is derived from a noise source $g(n)$ located within the noise field. This reference signal is correlated with the noise $v_1(n)$ that corrupts $S(n)$, and results in the noise corrupted observation, or the primary input signal, $d(n)$. The adaptive filter adjusts itself to produce an output $y(n)$ that best estimates $v_1(n)$ in the mean-square sense, and thus produces an error signal $e(n)= S(n)+ v_1(n)-y(n)$ that is as close as possible to $S(n)$.

In all conventional ANCs, it was assumed that there are no signal components leaking into the reference input. The presence of these signal components (also called signal crosstalk or signal leakage) at the reference input is a practical concern because it causes cancellation of a part of the original speech signal at the input of the ANC, and results in severe signal distortion and low signal to noise ratio at the output of the ANC. The magnitude of this distortion depends on the signal to noise ratios at the primary and reference inputs. Several techniques were proposed in the literature to enhance the system performance in this case of signal leakage[6-7]. High computational complexity is associated with these techniques and algorithms.

This paper introduces an ANC to improve the system performance in the presence of crosstalk. The proposed ANC consists of three microphones and two adaptive filters. Two microphones are used to represent the original speech signal and the reference noise input. The third microphone is used to provide a signal that is processed through the first adaptive filter to cancel the signal crosstalk leaking from the primary input into the reference input. The second adaptive filter is used to cancel the noise at its input and produce a signal that is as close as possible to the original speech.

The proposed ANC is simulated using different noise power levels for both stationary and nonstationary noise environments. Simulation results, carried out using a real speech, clearly demonstrate the significant

Corresponding Author: Zayed M. Ramadan, Department of Network and Communications Engineering, College of Engineering and Information Technology, Al Ain University of Science and Technology, P.O.Box 64141, Al Ain, UAE. Tel. +971-50-7534612. Fax. +971-3-7611198

achievements of the proposed ANC in minimizing the signal distortion and reverberation.

Proposed ANC: An adaptive noise canceller with signal leakage in the reference input is shown in Fig.2. The leakage signal is represented as an output of a low pass filter $\mathbf{h}_2$. Figure 3 shows a block diagram of the proposed ANC. The first microphone represents the speech signal and the second microphone represents a mixture of noise $g(n)$, and signal components leaking from the first microphone through a channel with impulse response $\mathbf{h}_3$. These signal components cause distortion in the recovered speech at the output of a conventional ANC.

To solve this problem we introduce a third microphone (sensor) to provide a signal that is correlated with the signal components leaked from the primary input. This signal is processed by the first adaptive filter ($\mathbf{w}_1$) to produce a crosstalk-free noise at its output. This noisy signal, with almost no leakage of the speech, is processed through the second adaptive filter to cancel the noise at its input, and accordingly produces the recovered speech at the output of the ANC. The transmission path between the third microphone and the first adaptive filter is represented by the impulse response $\mathbf{h}_2$. It is assumed that the third microphone is located farther away from the reference microphone such that there will no signal crosstalk leaking from the first into the second.

Figures 2 and 3 illustrate the fundamental concepts of the proposed crosstalk resistant ANC compared to the conventional ANC. In the ANC shown in Fig.2, the reference signal $v_2(n)$, which must be correlated with $v_1(n)$, is used to estimate the noise $v_1(n)$. A similar concept is used in the proposed crosstalk resistant ANC shown in Fig.3. The signal provided by the third microphone, $v_4(n)$, is correlated with the crosstalk signal that leaks from the primary microphone into the reference one. This signal (i.e., $v_4(n)$) is composed of both speech and noise and is almost an attenuated replica of the noise corrupted observation at the primary sensor.

The performance of an ANC can be measured in terms of a dimensionless quantity called misadjustment M, which is a normalized mean-square error defined as the ratio of the steady state excess mean-square error (EMSEss) to the minimum mean-square error.

$$M = \frac{\text{EMSE}_{ss}}{\text{MSE}_{min}} \qquad (1)$$

The EMSE at the nth iteration is given by

$$\text{EMSE}(n) = \text{MSE}(n) - \text{MSE}_{min} \qquad (2)$$

where

$$\text{MSE}(n) = E[|e(n)|^2] \qquad (3)$$

The MSE(n) in (3) is estimated by averaging $|e(n)|^2$ over I independent trials of the experiment. Thus, (3) can be estimated as:

$$\text{MSE}(n) \approx \overset{\wedge}{\text{MSE}}(n) = \frac{1}{I} \sum_{k=1}^{I} \left| e(n) \right|^2 \qquad (4)$$

From (2), we can write:

$$\text{EMSE}_{ss} = \text{MSE}_{ss} - \text{MSE}_{min} \qquad (5)$$

To obtain best results, an adaptive algorithm in which the step size depends on both error and data normalization is introduced and analyzed below. This algorithm is used in the computer simulations of both the proposed and the conventional ANCs. The simulations show performance superiority of the proposed ANC in decreasing signal distortion, reverberation and consequently, producing small values of EMSE.

Error-Data Normalized Step Size: Based on regularization Newton's recursion [11], we can write

$$\mathbf{w}(n+1) = \mathbf{w}(n) +$$
$$\mu(n) \left[\varepsilon(n) \mathbf{I} + \mathbf{R_X} \right]^{-1} \left[\mathbf{p} - \mathbf{R_X} \mathbf{w}(n) \right], \qquad (6)$$

where n is the iteration number, $\mathbf{w}$ is an $N \times 1$ vector of adaptive filter weights, $\varepsilon(n)$ is an iteration-dependent regularization parameter, $\mu(n)$ is an iteration-dependent step-size, and $\mathbf{I}$ is the $N \times N$ identity matrix. $\mathbf{p}(n) = E\left[d(n)\mathbf{x}(n) \right]$ is the cross-correlation vector between the desired signal $d(n)$ and the input signal $\mathbf{x}(n)$, and $\mathbf{R_X}(n) = E[\mathbf{x}(n)\mathbf{x}^T(n)]$ is the autocorrelation matrix of $\mathbf{x}(n)$.

Writing (6) in the LMS form by replacing $\mathbf{p}$ and $\mathbf{R_x}$ by their instantaneous approximation $d(n)\mathbf{x}(n)$ and $\mathbf{x}(n)\mathbf{x}^T(n)$, respectively, with appropriate proposed weights, we obtain

$$\mathbf{w}(n+1) = \mathbf{w}(n) +$$

$$\mu \times \left[\alpha \left\|\mathbf{e}_L(n)\right\|^2 \mathbf{I} + \gamma \mathbf{x}(n)\mathbf{x}^T(n)\right]^{-1} \mathbf{x}(n)e(n)$$

(7)

where μ is a positive constant step-size, α and γ are positive constants, $e(n)$ is the system output error defined by

$$e(n) = d(n) - \mathbf{x}^T(n)\mathbf{w}(n)$$

(8)

and

$$\left\|\mathbf{e}_L(n)\right\|^2 = \sum_{i=0}^{L-1} \left|e(n-i)\right|^2$$

(9)

Equation (9) is the squared norm of the error vector, $\mathbf{e}(n)$, estimated over its last L values.

Expanding (7) and applying the matrix inversion formula

$$\left[\mathbf{A} + \mathbf{B}\mathbf{C}\mathbf{D}\right]^{-1} =$$
$$\mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\left[\mathbf{C}^{-1} + \mathbf{D}\mathbf{A}^{-1}\mathbf{B}\right]^{-1}\mathbf{D}\mathbf{A}^{-1}$$

(10)

with

$$\mathbf{A} = \alpha \left\|\mathbf{e}_L(n)\right\|^2 \mathbf{I}, \mathbf{B} = \mathbf{x}(n), \mathbf{C} = \gamma, \text{and } \mathbf{D} = \mathbf{x}^T(n)$$

we obtain

$$\left[\alpha\left\|\mathbf{e}_L(n)\right\|^2 \mathbf{I} + \gamma\mathbf{x}(n)\mathbf{x}^T(n)\right]^{-1} =$$
$$\alpha^{-1}\left\|\mathbf{e}_L(n)\right\|^{-2}\mathbf{I} - \alpha^{-1}\left\|\mathbf{e}_L(n)\right\|^{-2}\mathbf{I}\,\mathbf{x}(n)\times$$
$$\frac{\mathbf{x}^T(n)\alpha^{-1}\left\|\mathbf{e}_L(n)\right\|^{-2}}{\gamma^{-1} + \mathbf{x}^T(n)\alpha^{-1}\left\|\mathbf{e}_L(n)\right\|^{-2}\mathbf{x}(n)}.$$

(11)

Multiplying both sides of (11) by $\mathbf{x}(n)$ from the right, and rearranging the equation, we have

$$\left[\alpha\left\|\mathbf{e}_L(n)\right\|^2 \mathbf{I} + \gamma\mathbf{x}(n)\mathbf{x}^T(n)\right]^{-1} \mathbf{x}(n) =$$
$$\frac{\mathbf{x}(n)}{\alpha\left\|\mathbf{e}_L(n)\right\|^2 + \gamma\left\|\mathbf{x}(n)\right\|^2}$$

(12)

Substituting (12) in (7), we obtain a new error-data normalized step-size (EDNSS) algorithm:

$$\mathbf{w}(n+1) = \mathbf{w}(n) +$$
$$\frac{\mu}{\alpha\left\|\mathbf{e}_L(n)\right\|^2 + (1-\alpha)\left\|\mathbf{x}(n)\right\|^2}\mathbf{x}(n)e(n)$$

(13)

where γ is replaced by $(1-\alpha) \geq 0$ in (13) without loss of generality. It should be noted that α and μ in this equation are different than those in the preceding equations. However, since these are all constants, α and μ are reused in (13).

The fractional quantity in (13) may be viewed as a time-varying step-size, $\mu(n)$, of the proposed EDNSS algorithm. Clearly, $\mu(n)$ is controlled by normalization of both error and input data vectors. This algorithm is dependent on normalization of both data and error. It differs from the NLMS algorithm in the added term $\left\|\mathbf{e}_L(n)\right\|^2$ with a proportional constant. For the case when L=n, this added term will increase the denominator of the time-varying step-size $\mu(n)$ (the fractional quantity of (13)), and hence a larger value of μ than that of the NLMS should be used in this algorithm to assure fast rate of convergence at the early stages of adaptation. As n increases (with L=n), $\mu(n)$ decreases except for possible up and down variations due to statistical changes in the input signal energy $\left\|\mathbf{x}(n)\right\|^2$. This indicates that the EDNSS algorithm with L=n performs well in stationary environments. In a nonstationary environment, the length of the error vector, L, should be constant to improve the tracking ability of the algorithm. In this case, as n increases, $\left\|\mathbf{e}_L(n)\right\|^2$ decreases, and $\mu(n)$ increases to a maximum value of $\dfrac{\mu}{(1-\alpha)\left\|\mathbf{x}(n)\right\|^2}$,

which is the time-varying step-size of the NLMS algorithm. A small positive constant could be added to the denominator of (13) to insure stability of the algorithm when the denominator is close to zero. Note that setting $\alpha=0$ in this equation results in the standard NLMS algorithm.
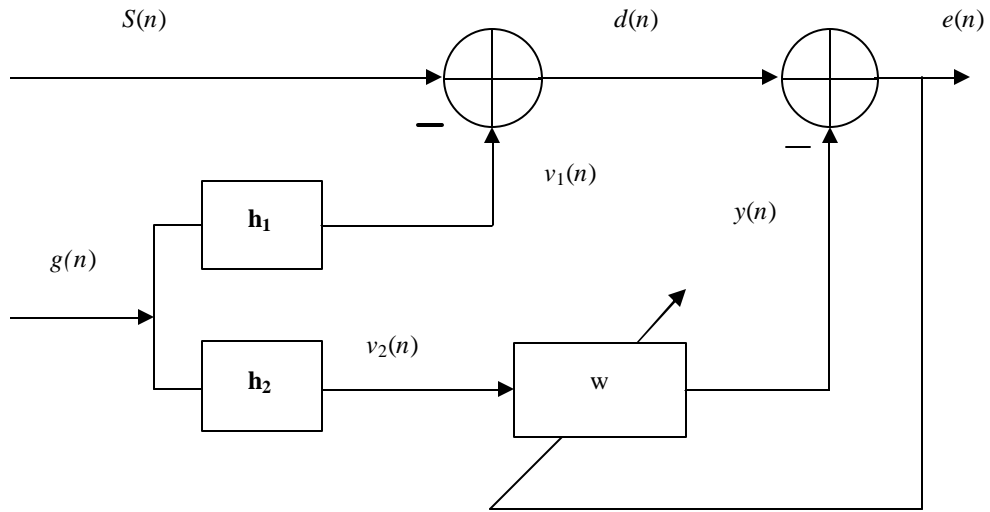
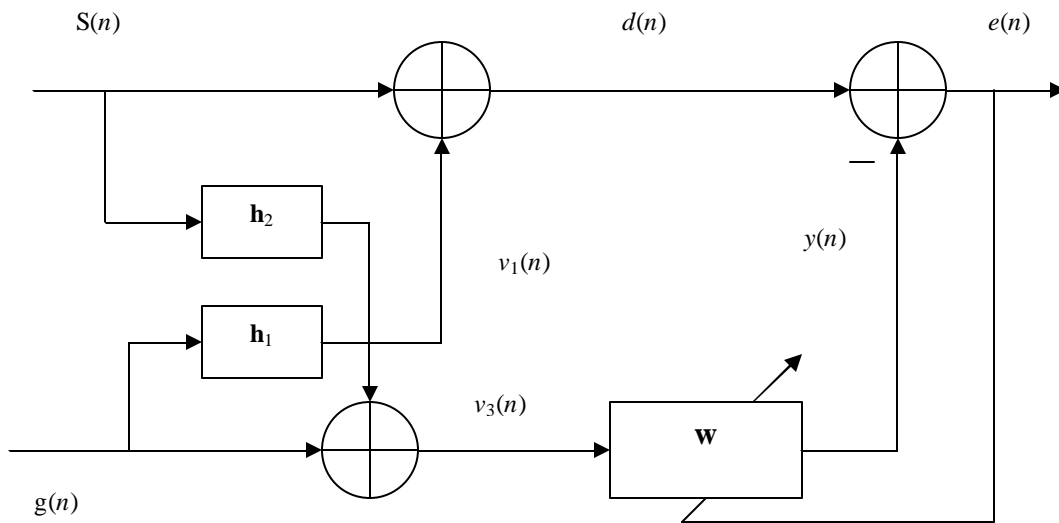Fig. 1: A conventional ANC with no signal leakage

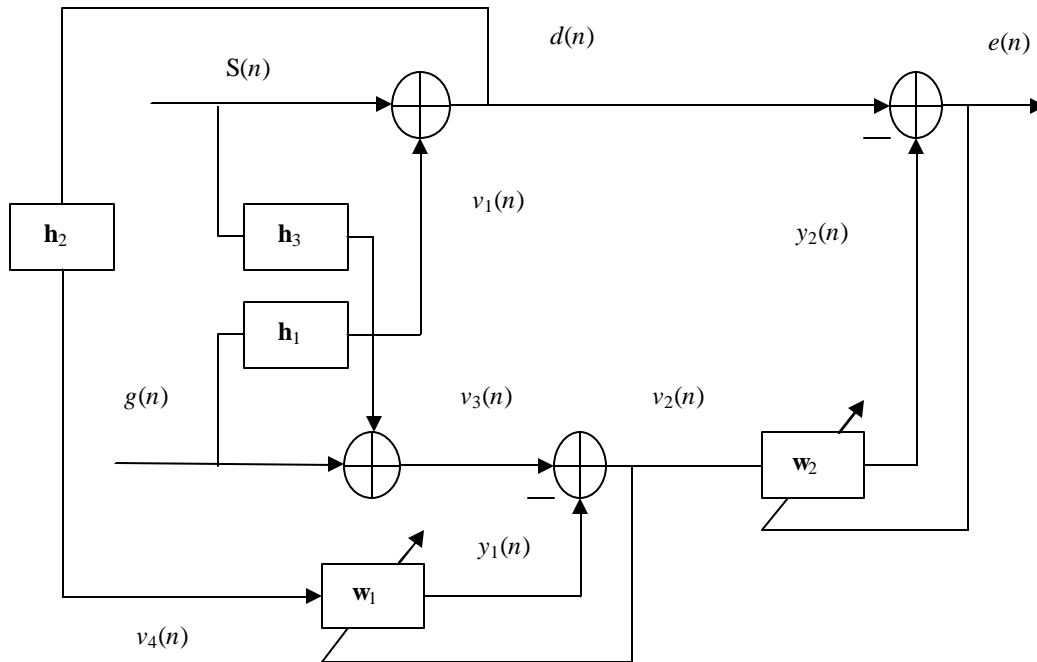Fig. 2: A conventional ANC with signal leakage

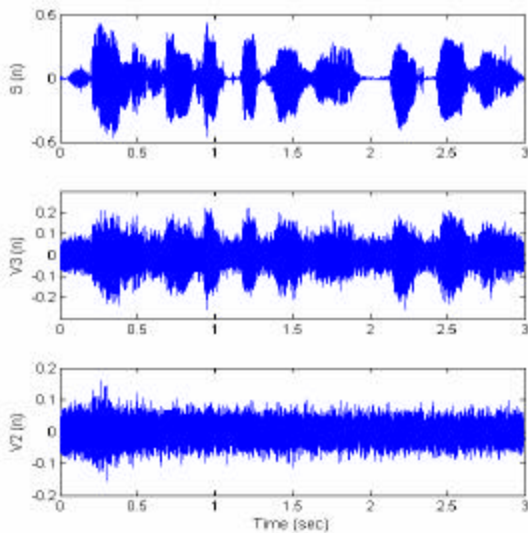Fig. 3: Proposed ANC for solving signal leakage problem

Fig. 4: Cancellation of crosstalk at the output of the first adaptive filter of the proposed ANC in stationary environment (σ_g^2 = 0.001, Table 1). From top to bottom: Original clean speech S(n), noise corrupted with crosstalk $v_3(n)$, and crosstalk-free noise, $v_2(n)$. See Fig.3.

SIMULATION RESULTS

The simulations are carried out using a male native speech saying sampled at a frequency of 11.025 kHz. The number of bits per sample is 8, and the total number of samples is 33000 or 3 sec of real time. Simulation results are presented for stationary and nonstationary environments. For the stationary case, the noise $g(n)$ was assumed to be zero-mean white Gaussian with three different variances as shown in Table 1. For the nonstationary case, the noise was assumed to be zero-mean white Gaussian with a variance that increases linearly from $\sigma_{g_{\min}}^2$ =0.001 to three different maximum values $\sigma_{g_{\max}}^2$ as demonstrated in Table 2. The impulse responses of the three autoregressive (AR) low pass filters used in the simulations, are $\mathbf{h}_1$ =[1.5 −0.5 0.1], $\mathbf{h}_2$ =[2 −1.5 0.3], and $\mathbf{h}_3$ =[3 −1.2 0.3].

In all simulations, the EDNSS algorithm, represented in (13), is used with the following values of parameters:

Proposed ANC:

$N_1=N_2=12$, $L_1=L_2=20N$, $\mu_1=0.15$, $\mu_2=0.03$, $\alpha_1=0.9$, and $\alpha_2=0.7$, where N_1, L_1, μ_1, and α_1 are filter length, error vector length, step-size and proportional constant, respectively, of the EDNSS algorithm used in the first adaptive filter ($\mathbf{w}_1$) shown in Fig.3. Similarly, N_2, L_2, μ_2, and α_2 are the corresponding parameters of the EDNSS algorithm used in the second adaptive filter ($\mathbf{w}_2$) shown in Fig.3.

Conventional ANC: $N=12$, $L=20N$, $\mu=0.03$, and $\alpha=0.7$, where N, L, μ, and α are the corresponding parameters of the EDNSS algorithm and used in the adaptive filter ($\mathbf{w}$) shown in Fig.2.

Figure 4 illustrates the performance of our proposed ANC in canceling the signal leakage at the output of the first adaptive filter for the case in which $\sigma_g^2 = 0.001$ as shown in Table 1. From top to bottom, Fig.4 shows the original speech $S(n)$, combination of noise and signal leakage $v_3(n)$, and the error signal of the first adaptive filter $v_2(n)$ which is the noise free version of signal leakage.

A comparison of the proposed ANC with the conventional ANC for both stationary and nonstationary noise environments is shown in Tables 1 and 2. The adaptation constants of the algorithm used in both ANCs were selected to achieve a compromise between small EMSE and high initial rate of convergence for a wide range of noise variances. From these tables, improvements of up to 23dB in $EMSE_{ss}$ of the proposed ANC over the conventional one were achieved. It is worthwhile to note that if the noise variance increases, the performance of the conventional ANC becomes a little better as illustrated in Tables 1 and 2. This is expected because increasing noise power results in a less significant effect of the signal leakage at the reference input. That is, the SNR at the output of a conventional ANC with signal leakage is inversely proportional to the SNR at its reference input [12].

Figure 5 illustrates the performance of proposed ANC in a stationary noise environment ($\sigma_{g_{\max}}^2 = 0.001$, Table 1). The figure shows the original clean speech $S(n)$, speech corrupted with noise $d(n)$, recovered speech $e(n)$, and the excess error $e(n)-S(n)$. Figures 6 and 7 provide more illustrations of the significant achievements of the proposed ANC over the conventional one in the nonstationary noise case in which $\sigma_{g_{\max}}^2 = 0.01$ (Table 2). Figure 6 shows the speech signal $S(n)$, the noise corrupted observation $d(n)$, and the excess error ($e(n)-S(n)$) of both the proposed and conventional ANCs. The effect of

increasing the variance of the noise on the processed speech is clearly shown in the second plot of the figure. Figure 7 shows the plot of EMSE in decibels for both ANCs.

Figure 8 repeats Fig.7 but in a stationary noise environment with $\sigma_g^2 = 0.01$ as shown in Table 1. The EMSE in this case starts decreasing during the initial stage of convergence. After that, it keeps fluctuating depending on the tracking capability of the adaptive algorithm. However, the EMSE in the nonstationary noise case is different from that in stationary case during the early stages of adaptation, since the variance of the noise is increasing linearly in this case.
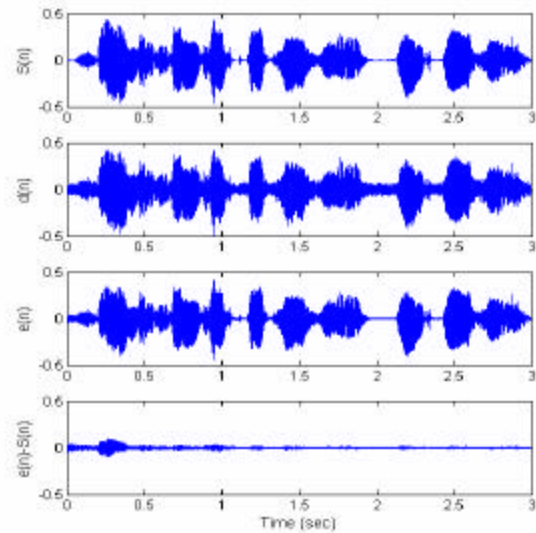

Fig. 5: The performance of proposed ANC in a stationary noise environment ($\sigma_g^2 = 0.001$, Table 1). From top to bottom: Original clean speech $S(n)$, speech corrupted with noise $d(n)$, recovered speech $e(n)$, and excess error $e(n)-S(n)$.

CONCLUSION

In this paper, a crosstalk resistant adaptive noise canceller is proposed to improve the performance in the presence of signal leakage or crosstalk at the reference input. This ANC uses three microphones and two adaptive filters. The third microphone is used to provide a signal that is correlated with the leaking signal components. This signal is processed by the first adaptive filter to produce a noise that is free from crosstalk at its output. The second adaptive filter cancels the noise at it input and produces a recovered speech that is as close as possible to the original speech. Computer simulations, using a new proposed adaptive

algorithm based on normalization of both error and data, show performance superiority of the proposed ANC in decreasing signal distortion and reverberation in the resulted speech, and consequently, producing small values of EMSE.
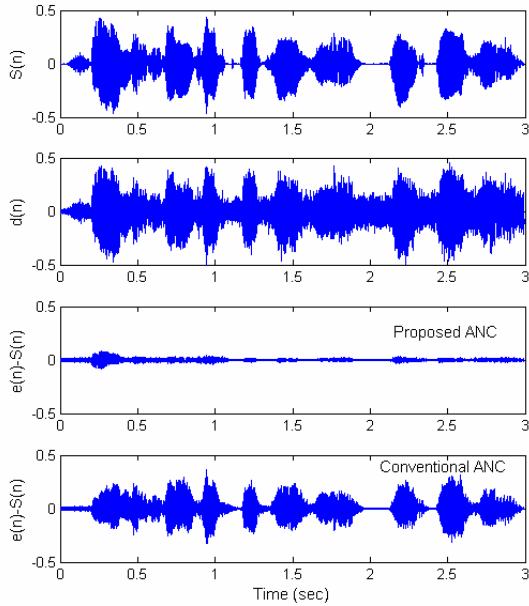

Fig. 6: Performance comparison between the proposed and conventional ANCs in a nonstationary noise environment ($\sigma_{g\,max}^2 = 0.01$, Table 2). From top to bottom: Original clean speech S(n), speech corrupted with noise $d(n)$, excess error of the proposed ANC, and excess error of the conventional ANC.
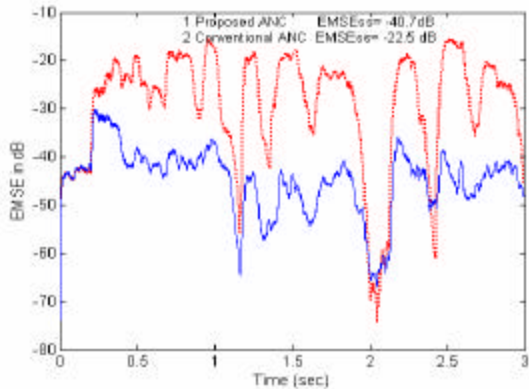

Fig. 7: EMSE in dB of the proposed and conventional ANCs in a nonstationary noise environment ($\sigma_{g\,max}^2 = 0.01$, Table 2).
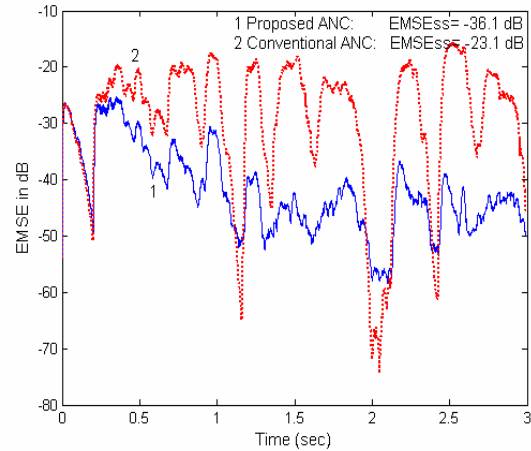

Fig. 8: EMSE in dB of the proposed and conventional ANCs in stationary noise environment ($\sigma_g^2 = 0.01$, Table 1).

Table 1: Comparison of the EMSE$_{ss}$ and M of the proposed and conventional ANCs for stationary noise case

Gaussian white zero-mean noise, g(n)	Conventional ANC		Proposed ANC	
	Steady-State EMSE (dB)	M %	Steady-State EMSE (dB)	M %
$\sigma_g^2 = 0.001$	−21.5	69.9	−41.7	0.7
$\sigma_g^2 = 0.01$	−23.1	49.0	- 36.1	2.4
$\sigma_g^2 = 0.1$	−25.4	28.9	−33.4	4.6

Table 2: Comparison of the EMSE$_{ss}$ and M of the proposed and conventional ANCs for nonstationary noise case

Nonstationary Gaussian noise g(n) $\sigma_{g\,min}^2 = 0.0001$	Conventional ANC		Proposed ANC	
	Steady-State EMSE (dB)	M %	Steady-State EMSE (dB)	M %
$\sigma_{g\,max}^2 = 0.001$	−21.4	72.5	−44.8	0.3
$\sigma_{g\,max}^2 = 0.01$	−22.5	56.3	- 40.7	0.8
$\sigma_{g\,max}^2 = 0.1$	−24.5	35.5	−35.2	3

REFERENCES

1. Boll, S. F., & D. C. Publisher, 1980. Suppression of acoustic noise in speech using two microphone adaptive noise cancellation. IEEE Transactions on Acoustics, Speech, and Signal Processing, ASSP-28, 752-753

2. Ikeda, S., & A. Sugiyama, 1999. An adaptive noise canceller with low signal distortion for speech codecs, IEEE Transactions on Signal Processing, 47 (3), 665-674

3. Greenberg, J. E., 1998. Modified LMS algorithms for speech processing with an adaptive noise canceller, IEEE Transactions on Speech and Audio Processing, 6, 338-351

4. Harrison, W. A., J. S. Lim, & E. Singer, 1986. A new application of adaptive noise cancellation, IEEE Transactions on Acoustics, Speech, and Signal Processing, 34 (1), 21-27

5. Al-Kindi, M. J., & J. Dunlop, 1987. A low distortion adaptive noise cancellation structure for real time applications. Proc. IEEE ICASSP, 2153-2156

6. G. Mirchandani, R. L. Zinser, and J. B. Evans. 1992. A new adaptive noise cancellation scheme in the presence of crosstalk. IEEE Trans. Circuits Syst.,Vol. 39, pp. 681-694

7. V. Parsa, and P. Parker, 1994. Constrained crosstalk resistant adaptive noise canceller. IEE Electronic Letters, Vol. 30, pp. 1276-1277

8. V. Parsa, P. A. Parker, and R. N. Scott, 1996. Performance analysis of a crosstalk resistant adaptive noise canceller. IEEE Trans. Circuits Syst.,Vol. 43, pp. 473-482

9. M.J. Al-Kindi and J. Dunlop, 1987. A low distortion adaptive noise cancellation structure for real time applications. Proc. IEEE ICASSP, pp. 2153-2156

10. S. Ikeda and A. Sugiyama, 1999. An adaptive noise canceller with low signal distortion for speech codecs. IEEE Trans. on Signal, Processing, Vol. 47, pp. 665-674.

11. A. H. Sayed, 2003. Fundamentals of Adaptive Filtering, Wiley Interscience, Hoboken, N.J

12. B. Widrow, and S.D. Stearns, 1985. Adaptive Signal Processing. Prentice–Hall, Englewood Cliffs, N.J