ISCA Archive
http://www.isca-speech.org/archive

6th European Conference on
Speech Communication and Technology
(EUROSPEECH'99)
Budapest, Hungary, September 5-9, 1999

A Parameter-Based 2-Talker Detection Apparatus for Echo Cancellation

WEN-RONG WU , SHIH-CHEN LIN , PO-CHENG CHEN , CHUN-HUNG KUO
Advanced Technology Center, Computer and Communications Research Laboratories,
Industrial Technology Research Institute, Hsin-Chu, Taiwan, R.O.C.
jimmy@prosche.ccl.itri.org.tw

ABSTRACT

A new robust double talk detector (DTD) for acoustic echo cancellation is proposed. This detector uses the square norm of the adaptive filter weight vector as the detection criterion. A fast algorithm for the norm computation is also developed. Simulations show that the proposed DTD method can significantly outperform conventional algorithms in noisy environments.

Keywords: Double-talk detector, NLMS, robust DTD.

10.21437/Eurospeech.1999-568

I. INTRODUCTION

A typical acoustic echo cancellation (AEC) system employs an adaptive FIR filter to emulate the response of the echo path and cancel the far-end signal from the outgoing signal. However, during the double-talk (DT) period, the adaptive filter will attempt to cancel both the echo and near-end signals resulting the divergence of the filter weights. A common remedy to this problem is to use a DT detector (DTD) and stop the filter adaptation during this period. Many DT detection algorithms have been proposed. We can classify the algorithms into three approaches. The first approach is based on the comparison of the far-end signal power and the near-end signal or error signal power [1],[2],[3],[4]. The second approach is based on the comparison of the cross-correlation of the far-end and error-signal, and the auto-correlation of the far-end signal [5]. The last one is based on the comparison of the linear predict coefficients of near-end signal and the linear predict coefficients of far-end signal [6]. However, most of these algorithms are either computational extensive or sensitive to noise. To solve these problems, this paper presents a new double-talk detection algorithm using the square norm of the adaptive filter weight vector. This algorithm is efficient yet robust. It can significantly outperform conventional algorithms in noisy environments.

II. THE PROPOSED METHOD

Figure. 1 shows the traditional structure of AEC and DTD. The DTD detect the DT using far-end signal *x(n)*, near-end signal *d(n)* and residue *e(n)*. If a DT is detected, the AEC filter ***W(n)*** stops adaptation. A widely used adaptation algorithm for updating the filter weight vector is the normalized LMS (NLMS) algorithm [7]. The NLMS algorithm can be represented as

$$\begin{cases} y(n) = \boldsymbol{W}(n)\boldsymbol{X}^{T}(n) \\ e(n) = d(n) - y(n) \\ \boldsymbol{W}(n+1) = \boldsymbol{W}(n) + \tilde{\mu} e(n)\boldsymbol{X}(n) \\ \tilde{\mu} = \dfrac{\mu_0}{\|\boldsymbol{X}(n)\|^2} \end{cases} \qquad (1)$$

where ***W(n)***=[*w(n),w(n-1),…,w(n-L+1)*] is the adaptive filter weight vector, ***X(n)***=[*x(n),x(n-1),…,x(n-L+1)*] is the far-end input signal vector, *L* is the filter length, *y(n)* is filter output signal, *d(n)* is the near-end input signal, *e(n)* is the output error, and u_0 is step size that controls the stability and convergence rate.
We defined *p(n)* the square norm of the adaptive filter weight vector (SNFW) as

$$p(n) \equiv \|\boldsymbol{W}(n)\|^2 = \sum_{i=0}^{L-1} w_i^2(n) \qquad (2)$$

and the SNFW difference between two consecutive instants as

$$\Delta p(n) = p(n+1) - p(n) \qquad (3)$$

Since the filter weight vector models the echo path ***H(n)***, its norm remains stable when the echo path is time-invariant. Even the echo path response is time-varying, the variation is expected to be small in many cases. When the DT occurs, the filter weight vector will soon diverge from its optimal solution. Thus, the weight-vector norm increases rapidly. We can use this property to detect the DT. In this paper, we use the SNWF difference as the detection criterion.

One problem associated with this approach is that calculation of the filter norm requires extensive computations making the straightforward application of this algorithm impractical. We then derive a fast algorithm to solve the problem.

Substituting (1) into (2), we can obtain a recursive equation for the SNFW.

$$\begin{aligned}
\sum_{i=0}^{L-1} w_i^2(n+1) &= \| \boldsymbol{W}(n+1) \|^2 \qquad (4)\\
&= \boldsymbol{W}^{\mathrm{T}}(n+1)\boldsymbol{W}(n+1)\\
&= (\boldsymbol{W}^{\mathrm{T}}(n) + \tilde{\mu}e(n)\boldsymbol{X}^{\mathrm{T}}(n))(\boldsymbol{W}(n) + \tilde{\mu}e(n)\boldsymbol{X}(n))\\
&= \| \boldsymbol{W}(n) \|^2 + \tilde{\mu}e(n)\boldsymbol{W}^{\mathrm{T}}(n)\boldsymbol{X}(n)\\
&\quad + \tilde{\mu}e(n)\boldsymbol{X}^{\mathrm{T}}(n)\boldsymbol{W}(n) + [\tilde{\mu}e(n)]^2 \| \boldsymbol{X}(n) \|^2\\
&= \| \boldsymbol{W}(n) \|^2 + 2\tilde{\mu}e(n)y(n) + [\tilde{\mu}e(n)]^2 \| \boldsymbol{X}(n) \|^2
\end{aligned}$$

Thus, we have

$$\begin{aligned}
&p(n+1)\\
&= p(n) + 2\tilde{\mu}e(n)y(n) + [\tilde{\mu}e(n)]^2 \| \boldsymbol{X}(n) \|^2 \qquad (5)
\end{aligned}$$

Substituting (5) into (3), the SNFW difference can be efficiently calculated as follow:

$$\begin{aligned}
\Delta p(n) &= p(n+1) - p(n)\\
&= 2\tilde{\mu}e(n)y(n) + [\tilde{\mu}e(n)]^2 \| \boldsymbol{X}(n) \|^2 \qquad (6)
\end{aligned}$$

Using the normalized LMS (NLMS) as the adaptive algorithm, we are able to compute the SNFW difference using only three multiplications and one addition disregard the filter length.

The detailed structure of the proposed DTD, which is referred to as SNFW-DTD, is shown in Figure 2. Assume that the switch SW1 is turned on at the beginning. The SNFW-DTD then calculates the SNFW difference. If it is less then a threshold (ie. $|\Delta p(n)|$ < threshold), a single-talk (ST) is claimed and the SNFW-DTD turns on the switches SW1 and SW2, and turn off SW3. When the SNFW difference of the AEC is larger then a threshold (ie. $|\Delta p(n)|$ > threshold), a DT is claimed and the SNFW-DTD turns off the switches SW1 and SW2, and turn on SW3. During a DT period, the AEC filter ***W(n)*** stops adaptation. As a consequence, the filter norm no longer reflects the status of the DT. To solve the problem, we introduce a shorter auxiliary NLMS filter ***W1(n)*** during a DT period. This filter starts adaptation when a DT period is claimed. Its SNFW difference is then calculated to detect the end of a DT. When the SNFW difference of the auxiliary filter is smaller than a threshold for a certain time, the DT is considered finished and the main AEC filter resumes adaptation.

III. SIMULATION RESULTS

We evaluated the performance of the proposed SNFW-DTD algorithm using real speech signals and compared the results with conventional DTD methods. All of the speech signals were sampled using a 8 K Hz sampling rate. The signals were recorded in a 4m x 4m square room. The far-end signal was received from a mobile phone, the echo signal and the local signal were received separately form a microphone. The near-end signal was a mixed signal containing the echo and local signals .

Figure 3(a)~3(e) summarize the simulation results. Figure 3(a) shows the far-end signal *x(n)*, Figure 3(b) shows the local source signal *s(n)*, Figure 3(c) shows the echo signal *c(n)*, Figure 3(d) shows the mixed near-end signal *d(n)*, and Figure 3(e) shows the DT interval. In the figure 3(e), the lower level signal denotes that the ST status and a higher level signal the DT status. The DT occurs in the intervals of [6.51~7.83], [8.28~9.41], [10.46~12.96] and [13.36~13.62]. The unit here is second.

Figure 4 is the detected DT status of the proposed SNFW-DTD method. Here, the horizontal line is the SNFW difference, $|\Delta p(n)|$. Figure 4(a) is the results for the no noise environment and 4(b) is for the white noise environment (SNR=20dB). We can see that the SNFW-DTD method can detect DT correctly.

We compare the performance of SNFW-DTD with conventional DTD methods in noisy environments. The simulation results are shown in Table 1. Here, α represents DT detection rate defined as the ratio of the correctly detected DT interval length and the real DT interval length, and β the DT error rate defined as the ratio of the wrongly detected DT interval length and the total speech interval length. Three conventional methods were compared; POWR denotes the power based DTD method [1], CORR denotes the cross-correlation based DTD method [5], and LPC denotes the linear-predict based DTD method [7]. Table 1(a) corresponds to the results of the no noise case. Table 1(b)-(e) are the results of the noisy cases (white noise). The SNR in Table 1(b) is 20 dB, in Table(c) is 15dB, in Table 1(d) is 10dB, and in Table 1(e) is 5dB. As we can see that in the noise-free case, except for CORR, all methods have a similar performance. However, if noise is present and the SNR is not too low (10-20 dB), the DT error rate of SNFW-DTD is much lower than other methods.

IV. CONCLUSIONS

We proposed a new and computationally efficient yet robust double-talk detector named SNFW-DTD for echo cancellation. The SNFW-DTD uses the square norm of filter weight vector as the detection criterion. One problem associated with the DTD is that the computational complexity of is high. Using the NLMS as the adaptive algorithm, we have derived a fast algorithm that significantly reduces the computations. We are able to compute the norm using only three multiplications and one addition disregard the filter length. Simulations show that the proposed SNFW-DTD is very effective and robust. In noisy environments, it significantly outperforms the conventional DTDs.

REFERENCES

[1] Otakar A. Horna and Bethesda, “Double Talk Detector For Echo Cancellers,” *no.4360712 in U.S. Patent*, Nov.23,1982.

[2] Ryoichi Miyamoto; Yoshio Itoh; Yoshikazu Nakano et al, “Double-Talk Detection In An Echo Canceller,” *no.4894820 in U.S. Patent*, Jan 16,1990.

[3] James P. Ashley and Lee M. Proctor, “Method and Apparatus For Echo Canceling With Double-Talk Immunity,” *no.5535194 in U.S. Patent* ,Jul. 9,1996.

[4] Shawn R. McCaslin; Nariankadu D. Hemkumar and Bheeshmar Redheendran, “Double-Talk Detector For Echo Canceller,” *no.5631900 in U.S. Patent*, May. 20,1997.

[5] Hua Ye and Bo-Xiu Wu, “A New Double-Talk Detection Algorithm Based on the Orthogonality Theorem,” IEEE Trans. on Communications, vol.39 no.11,pp. 1542-1545, Nov 1991.

[6] Jae Ha Yoo and Sung Ho Cho, “A New Double Talk Detector Using The Lattice Predictors For An Acoustic Echo Canceller,” Proceeding of IEEE TENCON'97 Speech and Image Technologies for Computing and Telecommunications *vol. 2, 1997, pp483-486, New York*, USA..

[7] J. Nagumo and A. Noda, "A leaning Method for System Identification," *IEEE Trans. on Automatic Control*, vol.12,pp. 282-287, June 1967.

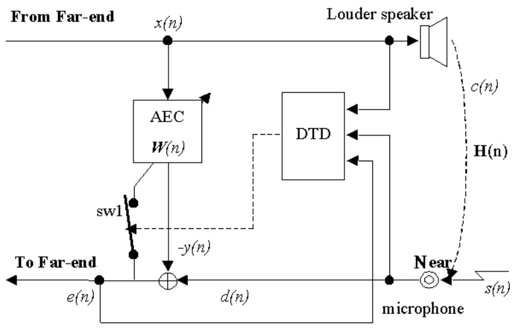

Figure 1.The conventional structure of AEC & DTD

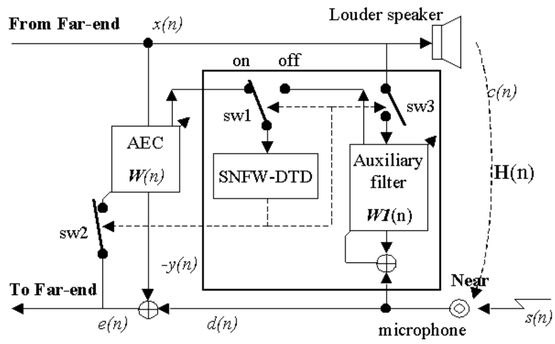

Figure 2. The structure of AEC & SNWF-DTD

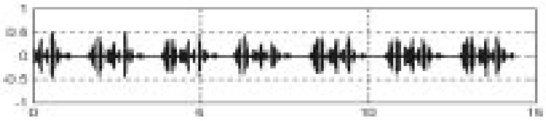

Figure 3(a).Far-end signal.

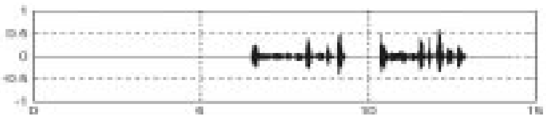

Figure 3(b).Local source signal.

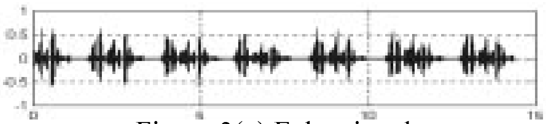

Figure 3(c).Echo signal.

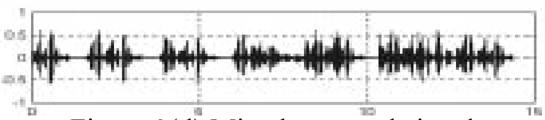

Figure 3(d).Mixed near-end signal.

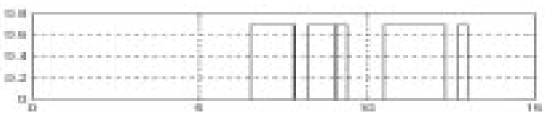

Figure 3(e).Double-talk status.

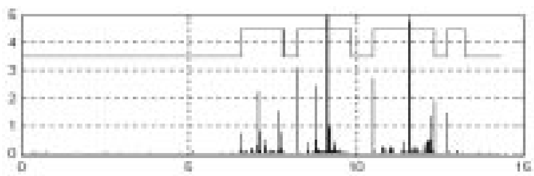

Figure 4(a). Detected DT status of SNFW-DTD for no noise

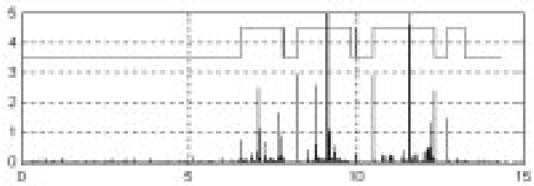

Figure 4(b). Detected DT status of SNFW-DTD for 20dB white noise.

Type \ Rate	POWR	CORR	LPC	SNFW
α	98.38%	83.46%	99.41%	99.34%
β	6.21%	9.31%	5.87%	5.76%

Table 1(a). No noise

Type \ Rate	POWR	CORR	LPC	SNFW
α	99.67%	83.56%	99.54%	99.34%
β	21.56%	20.11%	22.87%	5.78%

Table 1(b) SNR=20dB White Noise

Type \ Rate	POWR	CORR	LPC	SNFW
α	100%	97.08%	100%	99.30%
β	26.16%	23.16%	50.62%	13.32%

Table 1(c) SNR=15dB White Noise

Type \ Rate	POWR	CORR	LPC	SNFW
α	100%	97.36%	100%	99.34%
β	37.33%	36.51%	51%	35.56%

Table 1(d) SNR=10dB White Noise

Type \ Rate	POWR	CORR	LPC	SNFW
α	100%	97.92%	100%	100%
β	48.49%	48.89%	51%	47.59%

Table 1(e) SNR=5dB White Noise

Table 1 DTD efficiency comparison for four DTDs in noise environments: α is DT detection rate, β is DT error rate. POWR is power based DTD, CORR is correlation based DTD, LPC is linear-predict based DTD, and SNFW is the proposed SNFW-DTD.