Adaptive Step-Size Control Using Orthogonality Principles for Acoustic Echo Cancellation

Jang Sik Park*, Jin Youl Lee and Kyung Sik Son

* Dept. of Visual Technologies, Dongeui Institute of Technology
Dept. of Electronics Eng., Pusan National University

ABSTRACT

This paper presents a new time-varying step-size LMS algorithm for acoustic echo cancellation. The step-size is controlled by the cross-correlation of primary input signals and estimation error signals of adaptive filter. The cross-correlation is estimated by low-pass filtered instantaneous gradient estimate of LMS algorithm. And the step-size is normalized by sum of the input signal and error signal powers. The proposed algorithm has good performance at double-talk situation and faster convergence speed. As results of computer simulations, it is shown that the performance of the proposed algorithm is better than conventional ones.

1. INTRODUCTION

Acoustic echo cancellers are required in many voice communication systems such as teleconference and mobile hands-free telephones, as it makes natural and safety conversations possible[1]. For acoustic echo cancellers, NLMS algorithm is widely used due to its simplicity and stability[2]. However, NLMS algorithm with fixed step-size is not robust to ambient noises including near-end speech signals. Recently, several NLMS based algorithms were introduced employing time-varying step-size to have good convergence as well as smaller steady state error[3-7]. Although some algorithms of previous works are robust to ambient noises, the performance of acoustic echo cancellation is degraded due to slower convergence.

In this paper, a new time-varying step-size algorithm is proposed to improve the performance of acoustic echo cancellation at double-talk situation. The step-size is determined by the cross-correlation of the input signal and the error signal. As coefficients of adaptive filter converge to acoustic echo path, the correlation becomes close to zero. This is orthogonality principles of optimal filter[8]. The step-size will be smaller as the adaptive filter converges to acoustic echo path which has to be estimated to cancel echo. As the step-size is directly obtained from the cross-correlation, the performance is severely degraded since it is so fast decreased that the convergence is slower. At the initial state of adaptation, it is required large step-size to converge fast. To preserve large step-size, low-pass filtered cross-correlation is introduced. And the step-size is normalized with sum of the input and error power to be robust to noises.

Computer simulations are carried out to verify the performance of the proposed algorithm in this paper. As results, the proposed algorithm is robust to ambient noises and the misalignment of adaptive filter coefficients is smaller than other noise-robust adaptive algorithm.

2. ACOUSTIC ECHO CANCELLATION

Typical acoustic echo cancellation scheme is illustrated in Fig. 1. Near-end speech signal, \( n(k) \) is picked up by microphone and sent to far-end receiver. In a hands-free speech communication system the essential problem is that the microphone picks up the echo of far-end speech signal \( x(k) \) through acoustic echo path. \( y(k) \) and \( d(k) \) are echo and input signal of the microphone, respectively. Echo signal has to be eliminated for comfortable speech communication.

![Fig. 1. The concept of acoustic echo cancellation in a hands-free communication system.](image)

Acoustic echo canceller is the same as system identification problem as shown in Fig. 1. Echo signal is estimated by modeling the acoustic path with FIR filter, \( W(k) \). Estimated echo signal, \( \hat{y}(k) \) is subtracted from \( d(k) \) to remove echo signal. The FIR filter is updated by an adaptive algorithm to minimize error signal, \( e(k) \). NLMS algorithm described as equation (1)-(4) is widely used.

\[
\begin{align*}
    d(k) &= y(k) + n(k) = W^T(k)X(k) + r(k) \\
    e(k) &= d(k) - \hat{y}(k) = d(k) - W^T(k)X(k) \\
    W(k+1) &= W(k) + \mu(k)e(k)X(k)
\end{align*}
\]
\[
\mu(k) = \frac{\alpha}{X^T(k)X(k)} \approx \frac{\alpha}{L\sigma_x^2}
\]  
(4)

\(\mu(k)\) is time-varying step-size of NLMS algorithm and \(\alpha\) is the adaptation constant. \(L\) is the order of the filter and \(\sigma_x^2\) is variance of input signal of adaptive filter. NLMS algorithm effectively performs to remove echo at non-stationary environment. However, coefficients of adaptive filter are misadjusted by ambient noises including near-end speech, \(n(k)\) as denoted as following equation (5).

\[
W(k+1) = W(k) + \mu(k)(W^T(k)X(k) - y(k)) + \mu(k)X(k)n(k)
\]  
(5)

In equation (5), the third term is the cause that coefficients are misaligned. As small value is assigned to \(\mu(k)\), the misadjustment can be decreased. However, the convergence speed is slower due to small step-size. To overcome this problem, it is required that \(\mu(k)\) is intellectually adapted. This paper presents a time-varying step-size using orthogonality principles.

3. STEP-SIZE CONTROL USING ORTHOGONALITY PRINCIPLES

The basic idea is derived from orthogonality principles, which is that the cross-correlation of the input and error signal becomes zero as the adaptive filter converges to target system which has to be estimated. Equation (6) represents orthogonality principles

\[
E[e(k)X(k)] = 0
\]  
(6)

As the step-size obtains from cross-correlation, the trade-off of steady state error and convergence speed can be dismissed. The cross-correlation is estimated by instantaneous gradient estimate of LMS algorithm. The estimation can be disturbed by ambient noises since it is the same as instantaneous gradient estimate. Bad estimation will cause severe performance degradation because behaviors of adaptive filters depend on the step-size. It is normalized with sum of input variance and error variance to be robust to noises.

The idea proposed in this paper is similar to the variable step-size adaptive algorithm proposed by Hwang and Mathew[9]. The computation for obtaining the step-size determined by the inner product of current and previous gradient estimate vector is too increased. To reduce computational burden, the estimation of cross-correlation are performed with input signal corresponding to the first coefficient. The estimate of cross-correlation is computed by (7).

1) estimation of cross-correlation

\[
\delta(k) = \beta\delta(k-1) + (1 - \beta)\frac{e(k)x(k)}{\sigma_x^2 + \sigma_e^2}
\]  
(7)

\(\delta(k)\) is the estimate of cross-correlation and \(\beta\) is forgetting factor. \(\sigma_x^2\) and \(\sigma_e^2\) are variances of input and error, respectively. Since \(\sigma_x^2 + \sigma_e^2 > e(k)x(k)\), the range of \(\delta(k)\) is generally \(-1 < \delta(k) < 1\).

\(\delta(k)\) has large value at the begin of adaptation since \(e(k)\) is highly correlated with \(x(k)\). \(\delta(k)\) is smaller as the adaptive filter converges to acoustic path. At the double-talk situation, \(\delta(k)\) will become close to zero because near-end speech signal uncorrelated with input signal is dominant component of the residual error, \(e(k)\). The estimated cross-correlation is depicted in Fig. 2. It is exponential decreased with converging the adaptive filter to acoustic echo path, which leads to small step-size. Therefore, the steady state error of the proposed algorithm can be smaller and the adaptive filter can be robust to ambient noises.

![Cross-correlation of the input signal and error signal of adaptive filters](image)

Fig. 2. The cross-correlation of the input signal and error signal of adaptive filters

However, there is a problem that the estimate of correlation is decrease so fast that the adaptive filter can not estimate the acoustic path. It is required for adaptive filter to preserve the step-size larger at beginning of adaptation. And behaviors of adaptive filters are sensitive to their step-size. It is undesirable that the step-size is perturbed with impulse noise. The estimate of cross-correlation obtain from (7) is low-pass filtered by (8).

2) low-pass filtered cross-correlation(linear smoothing)

\[
\mu'(k) = \beta\mu'(k-1) + (1 - \beta)\delta(k)
\]  
(8)

\(\mu'(k)\) is low-pass filtered cross-correlation. The effect of low-pass filtering(linear smoothing) is well illustrated in Fig. 3. Low-pass filtered estimate is smooth and delayed a little. Finally, the step-size is computed by (9).

3) time-varying step-size

\[
\mu(k) = \frac{|\mu'(k)|}{L(\sigma_x^2 + \sigma_e^2)}
\]  
(9)
The step-size is obtained from normalizing the low-pass filtered cross-correlation with sum of input and error signal power to be robust to target speech signal and ambient noises. Stability is confirmed by taking absolute value of $\mu(k)$.

The step-size (solid-line) and the estimated cross-correlation of input and error signal (dot-line) are plotted in Fig. 3. Assuming double-talk, white Gaussian noise is generated between 2000 and 2500 sample. The cross-correlation is a little perturbed, but the step-size keeps smoothness. Therefore, it is expected that the proposed algorithm is insensitive to noises. To demonstrate for acoustic echo path change, the impulse response of path is changed at 3500 sample. It is shown that the step-size tracks the change of path.

Fig. 3 The proposed step-size and estimated cross-correlation of input and error signal. The variance of input signal is 1 and the variance of noise is 1 between 2000 and 2500 sample.

The step-size can be applied to other stochastic adaptive algorithm, for example, affine projection algorithm. Therefore, the performance of acoustic echo cancellation will be more improved.

4. SIMULATIONS AND CONSIDERATIONS

Compute simulations were carried out to verify the performance of proposed algorithm. Far-end speech and near-end speech signals are recorded at a sample rate of 8 kHz and quantized into 16 bits.

To demonstrate the robustness at double-talk situation, proposed algorithm is compared with NLMS and a noise-robust algorithm proposed by Greenberg[7] (we called it “SumLMS”). The step-size of SumLMS algorithm is normalized with sum of the input and error signal power. It was effective applied hearing-aids to cancel acoustic feedback under noisy environment including target speech signal. For the simulation, the adaptation constant sets to $\alpha = 0.2$ and the forgetting factor set to $\beta = 0.998$. The impulse response of acoustic echo path was measured at a small room. The order of adaptive filter sets to $L = 256$.

Recorded speech signals and simulation results are illustrated in Fig. 4 (a) and (b) are far-end speech and near-end speech, respectively. (c) and (d) are results of NLMS and SumLMS algorithm and (e) is error signal of the proposed algorithm. Although residual echo is remained a little, SumLMS and the proposed algorithm are well performed under double-talk situation. NLMS algorithm is completely misconverged after double-talk.

Fig. 4 Recorded speech signals and simulation results of proposed algorithm and conventional algorithms. (a) far-end speech signal (b) near-end speech signal (c) error signal of NLMS (d) error signal of SumLMS (e) error signal of proposed algorithm

EREL is compared in Fig. 5. EREL (echo return loss enhancement) is defined by equation (10).

$$EREL(k) = 10 \log_{10} \left( \frac{\sum_{n=0}^{L} d^2(n-i)}{\sum_{n=0}^{L} (d(n-i) - \hat{y}(n-i))^2} \right)$$

89
In Fig. 5, dot-line and dash-line denote NLMS and SumLMS and solid line represents the proposed algorithm. It is reliable that the performances of proposed algorithm and SumLMS algorithm are better than NLMS algorithm.

\[ \mathbf{H}_{sp}(k) \] means the impulse response of acoustic echo path. The proposed algorithm (solid-line) more accurately converges than SumLMS algorithm (dash-line) as well as NLMS algorithm (dot-line).

As results of simulations, it is certain that the proposed algorithm is more accurately converged to acoustic echo and robust to near-end speech signal.

5. CONCLUSIONS

In this paper, we presented a new time-varying step-size of adaptive algorithms to cancel acoustic echo at noisy environments. The step-size is controlled by the cross-correlation of input signal and error signal according to orthogonality principles. As results of computer simulations, miscalignments are improved than conventional ones. It is expected that the proposed step-size is well applied to affine projection algorithm which is fast converged for colored input signal like speech and the performance of acoustic echo cancellation will be more improved. We will implement the algorithm with fixed-point DSP.

REFERENCES


