An Echo Canceller Robust to Noise and Residual Echo

Hyun-Tae Kim, Jang-Sik Park, *Member, KIMICS*

***Abstract*— When we talk with hands-free in a car or noisy lobby, the performance of the echo canceller degrade because background noise added to echo caused by the distance from mouth to microphone is relatively long. It gives a reason for necessity of noise-robust and high convergence speed adaptive algorithm. And if acoustic echo canceller operated not perfectly, residual signal going through the echo canceller to far-end speaker remains residual echo, which degrade quality of talk. To solve this problem, post-processing needed to remove residual echo ones more. In this paper, we propose a new acoustic echo canceller, which has noise robust and high convergence speed, linked with linear predictor as a post-processor. By computer simulation, it is confirmed that the proposed algorithm shows better performance from acoustic interference cancellation (AIC) viewpoint.**

***Index Terms*— Noise-robust adaptive algorithm, double-talk detector, linear prediction, Post-processing.**

I. INTRODUCTION

NLMS family algorithms are simple and numerically robust. But these algorithms have drawback of converging slowly, especially when input signal is colored. On the other hand, RLS algorithm exhibits fast convergence in colored or strongly correlated input signal, but it has heavy computational burdens. To solve these problems, affine projection (AP) algorithm was suggested [1].

The AP algorithm is a generalization of the NLMS algorithm. This algorithm lies somewhere between NLMS algorithm and RLS algorithm from a performance and computational complexity point of view. But, when a projection is performed, noise amplification problem arises and this phenomenon degrades the performances of the AP algorithm [2][3].

And when the source of the acoustic echo signal is speech it is difficult to remove residual echo perfectly, because residual signal come from the output of the echo canceller has the same character of the speech signal [4].

In this paper, we propose a new modified AP algorithm to reduce noise amplification problem of AP algorithm. And it is linked with linear predictor as a post-processor, which whitens residual speech characteristics.

Manuscript received October 5, 2010; revised October 23, 2010; accepted October 31, 2010

Hyun-Tae Kim is with the Department of Multimedia Engineering, Dongeui University, Busan, 614-714, Korea (Email: htaekim@deu.ac.kr)

Jang-Sik Park is with the Department of Electronics Engineering, Dongeui Institute of Technology, Busan, Korea (Email: jsipark@dit.ac.kr)

II. PROPOSED ECHO CANCELLER

The block diagram of the proposed echo canceller is shown in fig. 1. The proposed system consists of two stages. The first stage consists of adaptive echo path modeling. And second stage consists of linear prediction error filter, as a post-processor, which whitens residual speech characteristics.

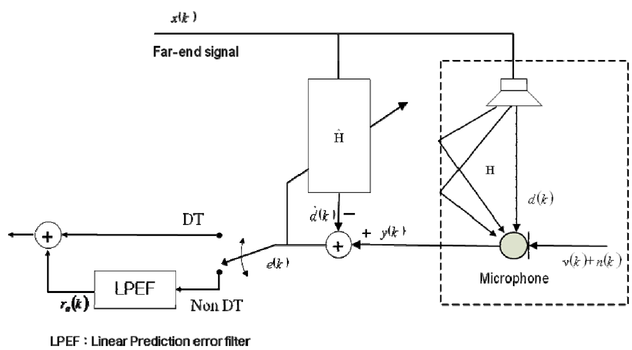

Fig. 1. Block diagram of the proposed total echo canceller.

A. Modified AP algorithm

A modified AP algorithm is proposed to reduce noise amplification problem of AP algorithm. The proposed modified AP algorithm normalizes the update equation to reduce noise amplification of AP algorithm, by adding the multiplication of error power and projection order to auto-covariance matrix of input signal.

The proposed modified AP algorithm, in general form, is summarized by Table. **I**

B. Post-Processing

1) Double-talk detector

An important characteristic of a good echo canceller is a double-talk detector. But in some applications, near-end noise may be continuously present and then the use of a double-talk detector becomes futile. Robustness to double-talk may be established by taking into account the near-end signal characteristics [5]. In proposed system, normalized cross correlation (NCR) algorithm [6] as a double-talk detector, is based on a comparison of the variances of the estimated and measured microphone signals $\hat{y}(t)$ and $y(t)$. In its standard form, the NCR algorithm is computationally infeasible, but fortunately one may form a computationally cheap version of the algorithm by assuming that the adaptive filter has converged to the true echo path. Here, we use the forgetting factor version of the cheap NCR algorithm, forming the decision variable [7].

TABLE I
SUMMARY OF THE MODIFIED AP ALGORITHM

$$\mathbf{e}_n = \mathbf{d}_n - \mathbf{X}_n^t \mathbf{w}_n \quad (1)$$

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu \mathbf{X}_n \left[\mathbf{X}_n^t \mathbf{X}_n + P \cdot L \cdot \sigma_{e,n}^2 \mathbf{I} \right]^{-1} \mathbf{e}_n \quad (2)$$

where L is the adaptive filter length, P is the projection order of modified AP algorithm, and the following definitions are made:

x_n : excitation signal of n[th] instant

$\mathbf{x}_n = \left[x_n \quad x_{n-1} \quad \cdots \quad x_{n-L+1} \right]^t$: excitation vector (3)

$\mathbf{X}_n = \left[\mathbf{x}_n \quad \mathbf{x}_{n-1} \quad \cdots \quad \mathbf{x}_{n-P+1} \right]$: P excitation matrix (4)

$\mathbf{w}_n = \left[w_{0,n} \quad w_{1,n} \quad \cdots \quad w_{L-1,n} \right]^t$: adaptive coefficient vector, $w_{i,n}$: i[th] adaptive coefficient at time n (5)

y_n : measurement noise signal.

$\mathbf{y}_n = \left[y_n \quad y_{n-1} \quad \cdots \quad y_{n-P+1} \right]^t$: measurement noise vector (6)

$\mathbf{d}_n = \left[d_n \quad d_{n-1} \quad \cdots \quad d_{n-P+1} \right]^t$: system output vector (7)

$e_n = d_n - \mathbf{x}_n^t \mathbf{w}_n$: *a priori* error signal (8)

$\mathbf{e}_n = \left[e_n \quad e_{n-1} \quad \cdots \quad e_{n-P+1} \right]^t$: *a priori* error vector (9)

μ : step size parameter

$\sigma_{e,n}^2 = \beta \sigma_{e,n-1}^2 + (1-\beta) e_n^2$: running power estimate of a *priori* error signal (10)

$$d_{CN}(k) = \sqrt{\frac{1}{\sigma_y^2} \hat{\mathrm{r}}_{\mathrm{xy}}^{\mathrm{T}}(k) \hat{\mathrm{H}}(k)} \quad (11)$$

where $\hat{\mathrm{r}}_{\mathrm{xy}}^{\mathrm{T}}(k) = \lambda \hat{\mathrm{r}}_{\mathrm{xy}}^{\mathrm{T}}(k-1) + (1-\lambda)\mathrm{x}(k) y(k)$ (12)

$$\hat{\sigma}_{\mathrm{y}}^2(k) = \lambda \hat{\sigma}_{\mathrm{y}}^2(k-1) + (1-\lambda) y^2(k) \quad (13)$$

with λ denoting a forgetting factor. Double-talk is deemed to occur when $d_{CN}(k)$ is below some predetermined threshold, *T*, i.e.,

$$Decision = \begin{cases} d_{CN} \geq T, & DT\ not\ present. \\ d_{CN} < T, & DT\ present. \end{cases} \quad (14)$$

2) Linear prediction error filter

The residual echo signal is whitened by using linear prediction error filter with P[th] order. During non-double-talk states, estimated error signal $e(k)$ in adaptive filter includes residual echo $r(k)$ and background noise $n(k)$. Residual echo $r(k)$ which has speech characteristics removed by whitening process using P[th] order linear prediction error filter, just like (15).

$$r_e(k) = e(k) - \sum_{i=1}^{P} a_i(k) e(k-i) \quad (15)$$

In equation (15) $a_i(k)$ and $r_e(k)$ denote coefficients of linear predictor and error signal from whitened acoustic echo canceller respectively. The coefficient of the linear predictor calculated from the solution of Wiener-Hopf equation, which is earned from Levinson-Durbin algorithm [8].

III. COMPUTER SIMULATION AND RESULTS

We apply to acoustic echo canceller with the proposed algorithm in hands-free environments. Far-end signal recorded with 8 kHz sampling rate, 16 bits quantization-level, and 10 second-long man and woman alternate pronounced English sentences. For considering double-talk situation, near-end signal recorded with 2 second-long man pronounced the other English sentences. Far-end signal to background noise ratio set to 30 dB and 20dB by assuming low level and high level white-Gaussian background noise, respectively. And acoustic echo path impulse response measured in small size office room with 512[th] order lengths. Adaptive filter has the same length, step-size, $\mu = 0.125$ and projection order $P = 2$. For performance evaluation with proposed algorithm, AIC(acoustic interference cancellation) were used[4].

(2.25)

$$AIC(k) = 10\log_{10} \frac{E\{y^2(k)\}}{E\{\hat{z}^2(k)\}} \quad (16)$$

(2.26)

$$(2.27) \quad = 10\log_{10} \frac{E\{y^2(k)\}}{E\{y^2(k) - \hat{i}^2(k)\}} \quad [dB]$$

Where, $\hat{i}(k) = \hat{d}(k) + \hat{n}(k)$ means estimated interference signal including estimated echo and background noise. Equation (16) means power ratio of microphone input signal $y(k)$ (including acoustic echo and background noise) and transmitted signal $\hat{z}(k)$ (or residual error signal). Therefore as acoustical echo is eliminated more by acoustic echo canceller, AIC has larger value.

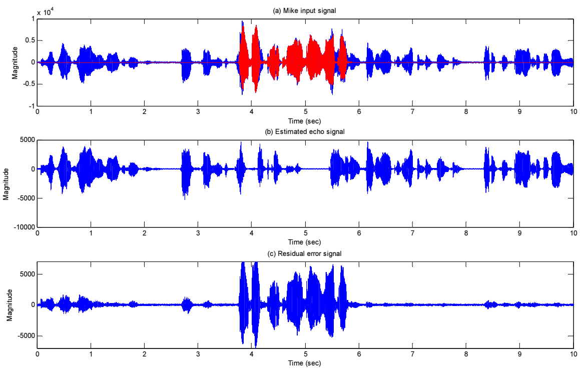

Fig. 2. A comparison of the echoes before and after cancellation (NLMS, 30dB white Gaussian noise).

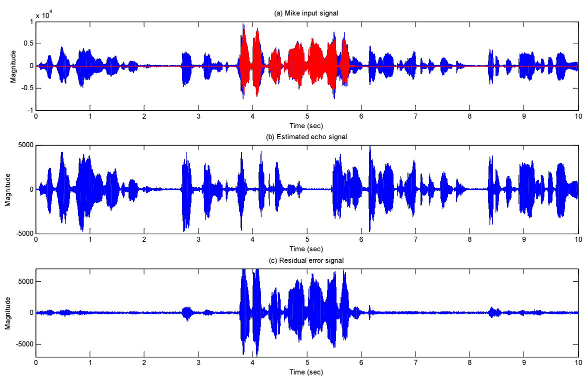

Fig. 3. A comparison of the echoes before and after cancellation (AP (P=2), 30dB white Gaussian noise).

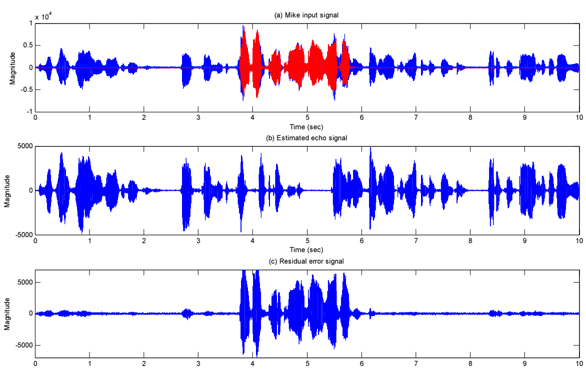

Fig. 4. A comparison of the echoes before and after cancellation (Proposed modified AP (P=2), without Post-processing, 30dB white Gaussian noise).

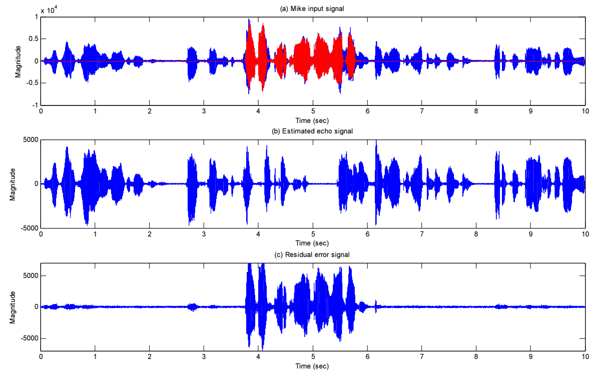

Fig. 5. A comparison of the echoes before and after cancellation (Proposed modified AP (P=2), 30dB white Gaussian noise).

Fig. 2 ~ 6 show the results for relatively high SNR, 30dB including double-talk intervals. In fig. 2 ~ 5 (a), red color signal means near-end signal for double-talk intervals. As shown in Fig. 5 (c), the echo and residual signal well removed by the proposed adaptive filter before and after double-talk intervals.

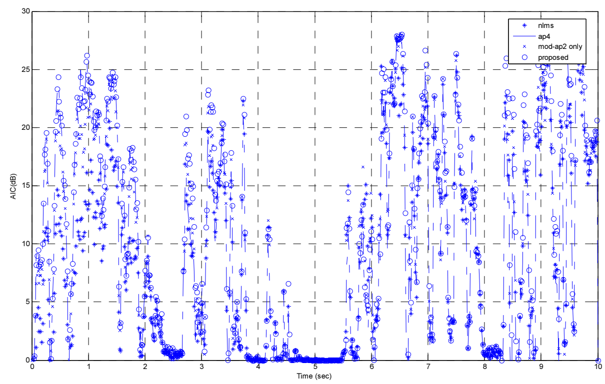

Fig. 6. AIC comparison with 30dB white Gaussian noise (NLMS, AP, proposed modified AP without post-processing, and proposed modified AP).

The results at relatively high SNR, based on AIC with NLMS, AP, proposed modified AP and proposed modified AP without post-processing are shown in Fig. 6. The Proposed algorithm has about 3 dB gain over the existing method at relatively high SNR.

Fig. 7 ~ 11 show the results for relatively low SNR, 20dB including double-talk intervals. In fig. 7 ~ 10 (a), red color signal means near-end signal for double-talk intervals. As shown in Fig. 10 (c), the echo and residual signal well removed by the proposed adaptive filter before and after double-talk intervals at relatively low SNR.

The results at relatively low SNR, based on AIC with NLMS, AP, proposed modified AP and proposed modified AP without post-processing are shown in Fig. 11.

The Proposed algorithm has more than 3 ~ 5dB gain over the existing method at relatively low SNR.

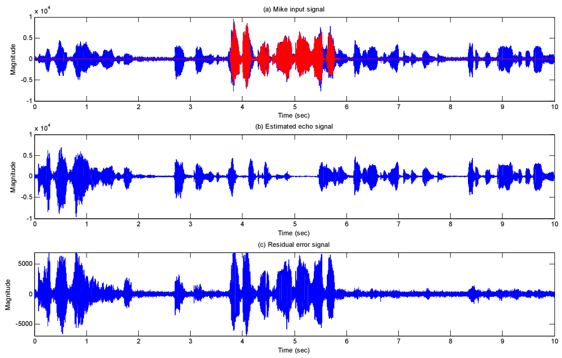

Fig. 7. A comparison of the echoes before and after cancellation (NLMS, 20dB white Gaussian noise).

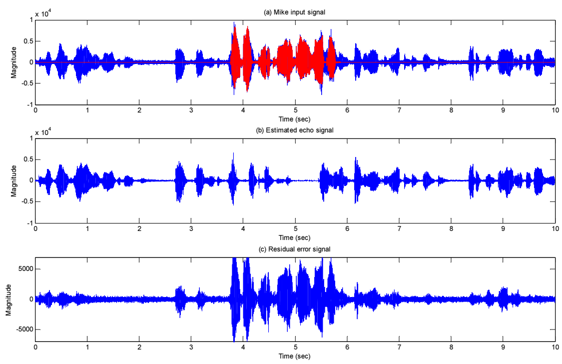

Fig. 8. A comparison of the echoes before and after cancellation (AP (P=2), 20dB white Gaussian noise).

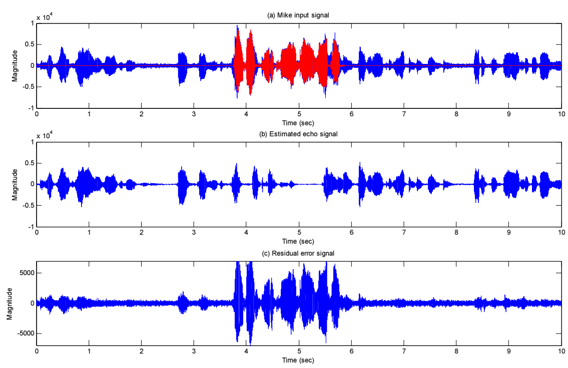

Fig. 9. A comparison of the echoes before and after cancellation (Proposed modified AP (P=2), without Post-processing, 20dB white Gaussian noise).

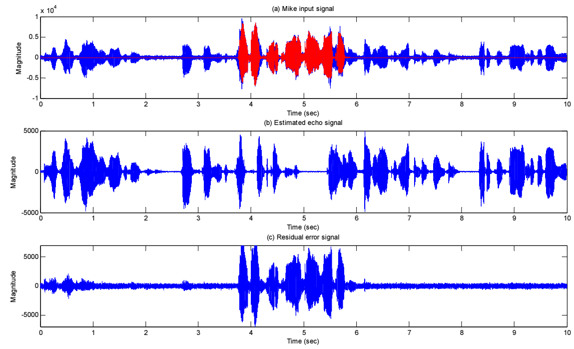

Fig. 10. A comparison of the echoes before and after cancellation (Proposed modified AP (P=2), 20dB white Gaussian noise).

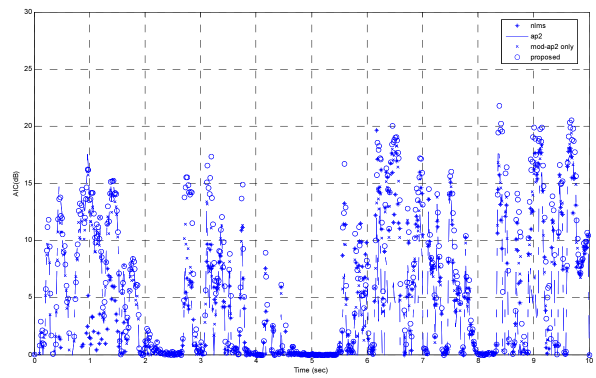

Fig. 11. AIC comparison with 20dB white Gaussian noise (NLMS, AP, proposed modified AP without post-processing, proposed modified AP).

IV. CONCLUSIONS

By computer simulation, the proposed algorithm has better performance over the existing method at relatively low SNR. Consequently, the proposed algorithm is more efficient to large background noise and residual echo.

REFERENCES

[1] S. L. Gay, "A fast converging, low complexity adaptive filtering algorithm," *App. of Signal Processing to Audio and Acoustics*, 1993. Final Program and Paper Summaries., 1993 IEEE Workshop on, Sep. 1993.

[2] M. Ferrer, A. Gonzalez, "Fast Affine Projection Algorithms for Filtered-x Multichannel Active Noise Control," *IEEE Trans. On Audio, Speech, & Language Processing*, Vol. 16, No. 8, Nov. 2008.

[3] Heping Ding, "Fast Affine projection Adaptation Algorithms with Stable and Robust Symmetric Linear System Solvers," *IEEE Trans. On Signal Processing*, Vol. 55, No. 5, May. 2007.

[4] Seon Joon Park, Chom Gun Cho, Chungyong Lee, Dae Hee Youn, "Integrated Echo and Noise Canceler for Hands-free applications" *IEEE Trans. Analog and Digital Signal Processing*, vol.49, no.3,p.188-195 March 2002.

[5] T. Waterschoot, G. Rombouts, " Double-Talk-Robust Prediction Error Identification Algorithms for Acoustic Echo Cancellation," *IEEE Trans. On Signal Processing*, Vol. 55, No. 3, March. 2007.

[6] J. Benesty, D. R. Morgan, "A new class of doubletalk detectors based on cross-correlation," *IEEE Trans. Speech Audio Processing,* vol. 8, no. 2, pp. 168-172, Mar. 2000.

[7] P. Ahgren, *On system identification and acoustic echo cancellation*, Ph.D. thesis, Uppsala University, 2004.

[8] Haykin, Adaptive Filter Theory, Prentice Hall, N. J., 1995.

Hyun-Tae Kim
received the B.S., the M.S. and the Ph.D. degree in the Electronics Eng. from Pusan National University, Korea in 1989, 1995 and 2000, respectively. He joined the Dongeui University in Korea as professor in the Multimedia Engineering Department since March 2002. He was a visiting professor at Georgia Institute of Tech. in USA at 2008.

Jang-Sik Park
received the B.S., the M.S. and the Ph.D. degree in the Electronics Eng. from Pusan National University, Korea in 1992, 1994 and 1999, respectively. He joined the Dongeui Institute of Technology in Korea as professor in the Electronics Engineering Department since March 1997.