

후처리를 이용한 음향 반향 제거기의 성능향상

Performance Improvement of Acoustic Echo Canceller Using Post-Processor

박 장 식*, 김 현 태**, 손 경 식***
(Jang Sik Park*, Hyun Tae Kim**, Kyung Sik Son***)

요 약

본 논문에서는 크게 계산량이 증가하지 않으면서 음향 반향 제거기의 성능을 향상시킬 수 있는 잡음에 강건한 적응 알고리즘과 후처리를 제안한다. 제안하는 적응 알고리즘은 기준 입력신호와 마이크 입력신호의 전력 합으로 적응 상수를 정규화 한다. 근단화자의 음성신호나 주변 잡음이 마이크로 입력되면 그 전력만큼 적응 상수가 작아지기 때문에 이들 주변 잡음에 의한 계수 오조정을 줄일 수 있다. 잔여 반향을 제거하기 위해서 잡음에 강건한 알고리즘과 연동하는 새로운 후처리 방법을 제안한다. 제안하는 후처리 방법은 마이크 입력신호와 추정 오차신호의 상관도를 활용한다. 잔여 반향은 마이크 입력신호의 전력으로 정규화된 상관도에 의해서 감쇠되어 제거한다. 정규화 상관도는 잔여 반향에 대해서 Wiener 필터 역할을 한다. 동시통화시에는 추정 오차신호에는 근단화자의 음성신호가 대부분을 차지하므로 정규화 상관도는 거의 1이 되어 근단화자의 음성신호는 감쇠없이 전송된다. 반향 신호만 있을 때에는 잔여 반향은 후처리에 의해서 대부분 감쇠되어 제거된다. 제안하는 후처리를 이용한 음향 반향 제거기의 계산량은 NLMS 알고리즘에 비해서 크게 증가하지 않는다.

ABSTRACT

In this paper, a new robust adaptive algorithm and a post-processing method are proposed to improve the performance of AEC without computational burden. Its step-size is normalized by the sum of the powers of the reference input signal and the desired signal. When the near-end speaker's speech and noise are applied into the microphone, the step-size becomes small and the misalignment of coefficients are reduced. To reduce the residual echoes, a new post-processing method, which is co-operated with the proposed noise-robust adaptive algorithm, is proposed in this paper. The method is based on the correlation of the desired signal and the estimation error signal. The residual echoes are attenuated as proportional to the correlation normalized with the power of desired signals. The normalized correlation plays a role as Wiener filter for residual echoes. In the double-talk situation, the estimation error signals, that are residual echoes, dominantly include the near-end speaker's speech and the normalized correlation closes to 1. Therefore, the near-end speaker's speech can be transmitted without being attenuated. When the desired signals consists of only the acoustic echoes, the residual echoes are mostly attenuated and canceled by the proposed post-processor. The computation of AEC using the proposed post-processor is comparable to NLMS algorithm.

1. 서 론

통신시스템의 발달로 다양한 음성통신 서비스가 제공되고 있으며 편리하고 안전한 통화를 위해서 핸드프리(hand-free) 단말기를 이용한 음성 통신이 원격회의(teleconference), 차량용 핸드프리 전화기 등에 활용되고 있다[1-7]. 그러나 핸드프리 단말기를 이용한 음성 통신에서는 스피커의 출력 음성 신호가 마이크로 입력되는 음

향 반향(acoustic echo)이 생긴다[3]. 핸드프리 단말기를 이용한 음성 통신에서 음향 반향이 제거되지 않으면 통화가 부자연스럽게 된다. 핸드프리 단말기를 이용한 음성 통신에서 음향 반향을 제거하여 원활한 음성 통신을 할 수 있도록 하는 것이 음향 반향 제거(acoustic echo cancellation) 기술이다. 음향 반향 제거기는 시간에 따라 변하는 반향 경로를 추정하여 음향 반향 신호를 추정하는 적응필터(adaptive filter) 기법이 일반적으로 활용되고 있다[8,9].

구조가 간단하고 안정한 수렴을 하는 NLMS 알고리즘(normalized least mean square algorithm)이 음향 반향 제

* 동의공업대학 영상정보처리과

** 부산울산 지방중소기업청

*** 부산대학교 전자공학과

접수일자: 1999년 2월 2일

알고리즘은 입력 신호가 음성신호와 같이 유색 신호(colored signal)인 경우에는 수렴속도가 현저히 느려지고 [12], 근단화자의 음성신호를 포함한 주변 잡음신호에 의해서 계수가 오조정되어 반향 제거 성능이 저하된다[10]. 그리고 적응 필터를 이용한 음향 반향 제거기는 입력신호의 특성, 반향 경로의 변화 그리고 주변 잡음에 의해서 적응 필터가 반향 경로를 정확하게 추정하지 못해서 잔여 반향이 발생한다.

입력 음성신호에 대해서 수렴이 느려지지 않도록 하기 위해서 입력신호를 백색화하는 방법과 유색 특성을 갖는 입력에 대해서 수렴이 느려지지 않는 RLS(recursive least square)와 AP(affine projection) 알고리즘들이 연구되고 있다[12-20]. 계산량과 수렴 속도 면에서 성능을 향상시키기 위해서 서브밴드(subband) 적응 필터 기법이 연구 제안되고 있다[21-22]. 그리고 근단화자의 음성신호 등의 주변잡음에 의한 성능 저하를 해결하기 위해서 동시통화 검출기[10]를 이용하는 방법과 적응필터의 적응 상수를 조절하는 방법이 제안되고 있다. 일반적으로 동시통화 검출기를 이용하는 방법은 근단화자의 음성 신호가 검출되면 적응필터의 적응을 중단하는 방법이다. 통계적으로 비슷한 성질을 갖는 반향과 근단화자의 음성신호를 분리하여 근단화자의 음성신호의 발생 여부를 판단해야 하기 때문에 정확한 동시통화 검출기를 구현하는 것이 필요하다. Akihiro 등은 동시통화 검출기를 사용하지 않고 동시통화를 포함한 잡음 환경에 강건한 알고리즘을 제안했다 [5-6]. 이 알고리즘은 기본적으로 주변 잡음의 전력만큼 적응필터의 적응상수를 작게 함으로써 계수 오조정을 줄이는 방법이다. 그러나 수렴속도가 느리기 때문에 적응 초기에는 반향이 거의 제거되지 않는다.

본 논문에서 계산량의 큰 증가 없이 음향 반향 제거의 성능을 향상시키기 위해서 주변 잡음에 대해서 강건한 적응 알고리즘과 연동하는 후처리기를 적응필터의 후단에 삽입하여 잔여 반향을 제거하는 음향 반향 제거기를 제안한다. 제안하는 적응 알고리즘은 근단화자의 음성신호 등의 주변 잡음에 의해서 적응필터의 계수가 오조정되지 않도록 적응필터 입력신호의 전력과 마이크 입력신호 전력의 합으로 적응 상수를 정규화한다. 이 알고리즘은 근단화자의 음성 신호등의 주변 잡음 신호가 없는 경우에는 NLMS 알고리즘과 같이 동작하고 주변 잡음 신호가 있는 경우에는 주변 잡음 신호의 전력만큼 적응 상수가 작아져서 적응 필터의 계수 오조정이 작아진다. 따라서 주변 잡음 신호에 의해서 반향 제거 성능이 저하되는 것을 줄일 수 있다. 또 파인 평균 자승 오차(excess mean square error)를 통해서 제안하는 알고리즘이 주변 잡음신호에 대해서 강건한 특성이 있음을 보인다. 기존의 NLMS 알고리즘은 주변 잡음신호의 전력에 비례해서 과잉 평균 자승 오차가 증가하지만 제안하는 알고리즘은 잡음신호의 전력이 증가하더라도 일정한 값으로 수렴한다.

잔여 반향을 제거하기 위한 후처리는 잔여 반향신호 즉 적응필터의 추정 오차신호와 마이크 입력신호의 상관도를 이용한다. 근단화자의 음성신호가 없이 음향 반향신

호만 있을 때에는 적응필터에 의해서 추정 가능한 반향 성분이 제거되었기 때문에 마이크 입력신호와 추정 오차신호간의 상관도는 낮다. 반대로 근단화자의 음성신호가 있을 때에는 추정 오차신호에는 근단화자의 음성신호가 대부분을 차지하므로 상관도는 1에 가까워진다. 이러한 성질을 이용해서 상관도를 음향 반향 제거기의 출력인 추정 오차신호에 곱해서 전송신호를 구함으로써 근단화자의 음성신호의 손상없이 잔여반향을 제거한다. 따라서 제안하는 음향 반향 제거기는 계산량이 크게 증가하지 않으면서 효과적으로 음향 반향을 제거할 수 있다. 본 논문에서는 컴퓨터 시뮬레이션을 통해서 제안하는 음향 반향 제거기가 기존의 알고리즘에 비해서 우수함을 보였다.

II. 기존의 음향 반향 제거기

핸드프리 단말기를 이용한 음성통신에서는 상대방(원단화자, far-end speaker)의 음성신호 $x(k)$ 가 스피커로 출력되어 그림 1과 같이 특정한 음향 경로를 통해서 마이크로 입력되어 반향신호 $y(k)$ 가 발생한다. 반향 신호는 근단화자의 음성신호 $n(k)$ 와 함께 재전송되고 반향 경로와 통신 시스템의 지연에 의해 일정 시간 뒤에 원단화자는 음성 통신 중에 자신의 음성을 다시 듣게 되기 때문에 음성 통신이 부자연스럽게 된다. 따라서 자연스러운 통신을 위해서 음향 반향은 제거되어야 한다. 적응필터 기법을 이용한 음향 반향 제거기는 반향 신호 $\hat{y}(k)$ 를 적응 필터로부터 추정하여 마이크 입력신호 $d(k)$ 로부터 빼주므로써 반향을 제거한다.

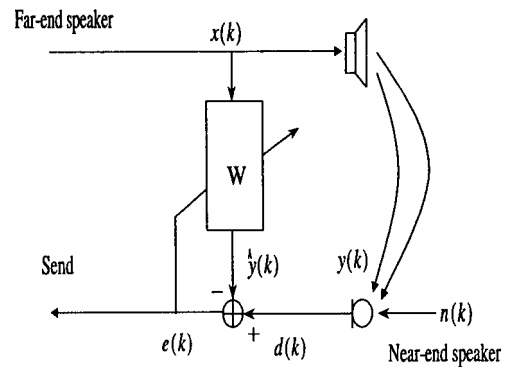


그림 1. 적응필터를 이용한 음향 반향 제거기
Fig. 1. Acoustic echo canceller using adaptive filtering.

음향 반향 제거기의 적응 알고리즘은 일반적으로 안정된 수렴을 하고 구조가 간단하여 구현이 용이한 NLMS 알고리즘을 활용한다[1]. NLMS 알고리즘은 LMS 알고리즘의 적응상수를 입력신호의 전력으로 정규화한 것이다. LMS 알고리즘은 정적인 환경(stationary)에서 유도된 알고리즘이지만[29] 비정적인 환경(nonstationary)에서도 잘 적용하는 것으로 알려져 다양한 분야에 응용되고 있다

[24-25]. 그림 1과 같은 구조에서 적응 필터를 적응시키는 NLMS 알고리즘은 식(1)~(4)과 같이 추정 오차신호와 입력신호로 적응필터의 계수를 적응시킨다.

$$d(k) = y(k) + n(k) = \mathbf{W}_0^T(k) X(k) + n(k) \quad (1)$$

$$e(k) = d(k) - \hat{y}(k) = d(k) - \mathbf{W}^T(k) X(k) \quad (2)$$

$$\mathbf{W}(k+1) = \mathbf{W}(k) + \mu(k)e(k)X(k) \quad (3)$$

$$\mu(k) = \frac{\alpha}{X^T(k)X(k)} \approx \frac{\alpha}{L\sigma_x^2} \quad (4)$$

$d(k)$ 는 음향 반향과 주변 잡음신호의 합으로써 마이크로 입력되어 샘플링된 신호이며, 적응필터의 주입력신호이다. $n(k)$ 는 근단화자의 음성신호 및 잡음을 포함하는 주변 잡음신호이다. $\mathbf{W}_0(k)$ 는 음향 반향 경로의 계수벡터로써 최적 계수벡터(optimal weight vector)이다. $X(k)$ 와 $\mathbf{W}(k)$ 는 각각 적응필터의 입력벡터와 계수벡터이다. $[*]^T$ 는 벡터의 전치(transpose)이다. 식(4)의 $\mu(k)$ 는 적응상수이고, α 는 적응필터의 정규화 적응상수이다. L 은 적응필터의 차수를 나타내며 σ_x^2 는 입력신호의 분산이다. $e(k)$ 는 추정 오차신호로써 상대방으로 전송되어질 신호이다. 추정 오차신호는 음향 반향 제거기에서는 잔여 반향신호이기도 하다.

NLMS 알고리즘은 LMS 알고리즘의 적응상수를 식(4)와 같이 입력신호의 전력으로 정규화함으로써 시간에 따라 σ_x^2 가 변하는 비정적인 입력 음성신호에 대해서 LMS보다 빠르고 안정된 수렴을 한다. 적응필터의 계수벡터 $\mathbf{W}(k)$ 는 $e(k)$ 와 $X(k)$ 를 이용해서 $E[e^2(k)]$ 가 최저가 되도록 적응한다. 계수벡터 $\mathbf{W}(k)$ 가 최적 계수 벡터 $\mathbf{W}_0(k)$ 를 추정해감에 따라서 적응필터의 출력 $\hat{y}(k)$ 가 반향신호 $y(k)$ 를 추정한다. 식(2)와 같이 마이크 입력신호 $d(k)$ 로부터 $\hat{y}(k)$ 를 빼면 전송되어야할 근단화자의 음성신호 $n(k)$ 를 분리하여 전송할 수 있다.

NLMS 알고리즘은 동시통화시 근단화자의 음성신호에 의해서 적응필터의 계수가 크게 오조정되어 음향 반향 제거 성능이 저하되는 단점이 있다. 동시통화시에는 추정 오차신호에 이들 신호가 포함되어 식(1)을 (2)에 대입하면

$$\begin{aligned} e(k) &= y(k) + n(k) - \mathbf{W}^T(k) X(k) \\ &= \mathbf{W}_0^T(k) X(k) - \mathbf{W}^T(k) X(k) + n(k) \end{aligned} \quad (5)$$

이 되고 식(5)를 식(3)에 대입하면

$$\mathbf{W}(k+1) = \mathbf{W}(k) + \mu(k)(y(k)X(k) - \mathbf{W}^T(k)X(k)X(k) + \mu(k)n(k)X(k)) \quad (6)$$

이 된다. 적응필터 계수벡터 $\mathbf{W}(k)$ 는 상관관계(correlation)가 없는 $n(k)$ 에 의해서 오조정된다. 그리고 $n(k)$ 에 에너지가 큰 근단화자의 음성신호가 포함되면 오조정은 더욱 커진다. 그리고 음성 통신을 하는 환경에 따른 주변 잡음과 측정 잡음신호에 의해서 적응필터가 정확하게 반향경로를 추정하는 것은 어렵기 때문에 상당한 잔여 반향이 남게된다.

III. 제안하는 음향 반향 제거기

본 논문에서는 간단한 적응 알고리즘을 이용하여 반향을 제거하고 남은 잔향은 적응 필터의 추정 오차신호와 마이크 입력신호의 상관도를 이용하여 잔향을 감쇠시켜 제거하는 방법을 제안한다. 음향 반향 제거 시스템은 그림 2과 같다. 그림 2의 적응 필터는 마이크 입력신호 $d(k)$ 를 적응 필터의 출력 $\hat{y}(k)$ 로 빼주므로써 반향을 제거한다. 그리고 적응필터의 추정 오차신호이며 잔여 반향인 $e(k)$ 를 감쇠기로 감쇠시켜 $\hat{n}(k)$ 를 출력한다.

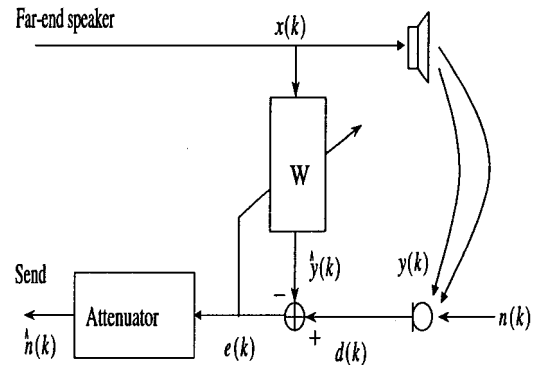


그림 2. 후처리기를 포함한 제안하는 음향 반향 제거기 구조
Fig. 2. Structure of acoustic echo canceller including a post-processor.

3.1 적응 알고리즘

본 논문에서는 주변잡음에 의해서 적응필터의 계수 오조정되는 것을 줄이기 위해서 잡음에 대해서 강건한 음향 반향 제거 알고리즘을 제안한다. 제안하는 알고리즘은 적응필터의 계수를 적응시킬 때 식(4)의 NLMS 알고리즘 적응상수 대신 식(7)과 같이 가중값을 곱한 적응상수를 이용하여 잡음에 의한 계수 오조정을 감소시킨다.

$$\mu(k) = \frac{\sigma_x^2(k)}{\sigma_x^2(k) + \sigma_n^2(k)} \frac{\alpha}{X^T(k)X(k)} \quad (7)$$

$\sigma_x^2(k)$ 와 $\sigma_n^2(k)$ 는 각각 입력신호와 마이크로폰 입력신호의 분산(variance), 즉 전력이다. 가중값은 근단화자나 주변잡음이 적은 경우에는 1에 가까워져서 NLMS와 같이 적용하고 주변 잡음신호가 있는 경우에는 그 전력만

음 적응상수 값을 감소시켜 계수오조정을 줄일 수 있다. 그리고 식(7)은 입력신호 벡터의 내적을 입력신호의 전력으로 표현할 수 있기 때문에 식(8)로 표현할 수 있다.

$$\mu(k) = \frac{\sigma_x^2(k)}{\sigma_x^2(k) + \sigma_n^2(k)} \frac{\alpha}{L\sigma_x^2(k)} = \frac{\alpha}{L(\sigma_x^2(k) + \sigma_n^2(k))} \quad (8)$$

식(8)과 보는 바와 같이 제안하는 알고리즘은 마이크 입력신호의 전력을 추정하여 입력신호의 전력과 더하는 계산만 추가되기 때문에 NLMS 알고리즘에 비해서 크게 계산량이 증가하지는 않는다.

3.2 후처리 알고리즘

후처리는 마이크 입력신호로 정규화된 추정 오차신호와 마이크 입력신호 간의 상관도, $E[e(k)d(k)]/E[d^2(k)]$ 를 이용하여 식(9)와 같이 추정 오차신호를 감쇠시켜 송신신호 $\hat{n}(k)$ 를 출력한다.

$$\hat{n}(k) = \rho(k)e(k). \quad (9)$$

$\rho(k)$ 는 식(1)을 이용하여 식 (10)과 같이 표시할 수 있다.

$$\rho(k) = \frac{E[e(k)d(k)]}{E[d^2(k)]} = \frac{E[e(k)y(k)] + E[e(k)n(k)]}{E[d^2(k)]} \quad (10)$$

그리고, 반향신호 $y(k)$ 와 주변 잡음신호 $n(k)$ 가 서로 상관 없이 있다고 가정하면 $E[y(k)n(k)] = 0$ 이다. 따라서 식 (10)에서 분모항은

$$E[d^2(k)] = E[y^2(k) + n^2(k) + 2y(k)n(k)] = E[y^2(k)] + E[n^2(k)] \quad (11)$$

이 되고 식(10)는

$$\rho(k) = \frac{E[e(k)d(k)]}{E[d^2(k)]} = \frac{E[e(k)y(k)] + E[e(k)n(k)]}{E[y^2(k)] + E[n^2(k)]} \quad (12)$$

이 된다. 정규화 상관도 $\rho(k)$ 는 반향신호나 주변 잡음신호의 크기 그리고 수렴 정도에 따라 값이 변한다. 본 논문에서는 적응필터가 수렴하지 않은 상태와 적응필터가 수렴된 상태에서 반향신호만 있는 경우와 근단화자 신호가 있는 경우로 나누어서 각각의 경우에 $\rho(k)$ 의 변화에 대해서 분석했다.

1) 적응필터가 수렴하지 않은 경우

적응필터가 수렴해 가는 과정에서 후처리의 동작을 다음과 같이 분석한다. 식(12)에 식(2)를 대입하면

$$\rho(k) = \frac{E[(y(k) - \hat{y}(k) + n(k))y(k)] + E[(y(k) - \hat{y}(k) + n(k))n(k)]}{E[y^2(k)] + E[n^2(k)]} \quad (13)$$

이 된다. 주변 잡음신호 $n(k)$ 가 음향 반향신호 $y(k)$ 그리고 반향 추정신호 $\hat{y}(k)$ 와 상관도가 없다는 가정을 적용하면 식(13)은

$$\rho(k) = \frac{E[y^2(k) - \hat{y}(k)y(k)] + E[n^2(k)]}{E[y^2(k)] + E[n^2(k)]} \quad (14)$$

와 같이 간단하게 표시된다. 적응필터가 충분히 적응하기 전의 수렴 초기에는 반향신호와 적응필터의 출력신호의 상관도가 낮다. 따라서 근단화자의 음성 신호를 포함한 주변 잡음 신호의 유무에 상관없이 $\rho(k)$ 는 1에 가까운 값이 된다. 그리고 적응필터가 음향 반향 경로를 추정해 감에 따라 $\hat{y}(k)$ 는 $y(k)$ 에 가까워져서 $\rho(k)$ 는 주변 잡음신호의 전력에 의존하게 된다.

2) 적응필터가 수렴한 경우

적응필터가 음향 반향 경로를 거의 추정할 정도로 수렴한 경우, 적응 필터의 출력 $\hat{y}(k)$ 는 음향 반향 $y(k)$ 를 추정한다. 따라서 식(14)는

$$\rho(k) = \frac{E[n^2(k)]}{E[y^2(k)] + E[n^2(k)]} \quad (15)$$

이 된다. 식 (15)의 결과는 적응필터의 직교원리를 이용하여 $\rho(k)$ 의 변화에 대해서 분석하는 결과와 같다. 식(1)을 식(12)에 대입하면

$$\rho(k) = \frac{E[e(k)X^T(k)W^*] + E[e(k)n(k)]}{E[y^2(k)] + E[n^2(k)]} = \frac{E[e(k)X^T(k)]W^* + E[e(k)n(k)]}{E[y^2(k)] + E[n^2(k)]} \quad (16)$$

이 된다. 분자의 첫 번째 항은 직교원리(orthogonal principle)에 의해서 적응필터가 수렴한 경우 추정 오차신호와 입력신호와 상관도, $E[e(k)X^T(k)] = 0$ 이 된다. 따라서 식(16)은

$$\rho(k) = \frac{E[e(k)n(k)]}{E[y^2(k)] + E[n^2(k)]} = \frac{E[(y(k) - \hat{y}(k) + n(k))n(k)]}{E[y^2(k)] + E[n^2(k)]} \quad (17)$$

이 되고, $n(k)$ 가 $y(k)$ 그리고 $\hat{y}(k)$ 와 상관도가 없다는 가정을 적용하여 다시 정리하면 식(17)는 식(15)와 같아진다. 식(15)의 우변을 보면 $n(k)$ 가 입력신호이고 $y(k)$ 가 잡음인 경우에 대한 Wiener 필터와 같은 역할을 한다.

반향신호가 근단화자의 음성신호보다 크면 $\rho(k)$ 값은 작아지고 근단화자의 음성신호가 반향신호보다 크면 $\rho(k)$ 는 거의 1이 된다.

적용필터가 충분히 수렴하고 주변 잡음 신호가 없는 경우 즉 $E[n^2(k)] \approx 0$ 일 때는 식(15)에서 $\rho(k) \approx 0$ 가 되고 잔여반향은 후처리에 의해서 대부분 감쇠되어 제거 된다.

원단화자의 음성신호가 발생하지 않아서 음향 반향이 없고 근단화자의 음성신호만 있는 경우에는

$$\rho(k) \approx \frac{E[n^2(k)]}{E[n^2(k)]} = 1 \quad (18)$$

이 된다. 근단화자의 음성 신호만 있는 경우에는 추정 오차신호에 포함되어 있는 근단화자의 음성신호가 거의 감쇠없이 전송된다.

근단화자의 음성신호와 음향 반향신호가 동시에 발생했을 때에는 식(15)과 같은 비율로 근단화자의 음성신호와 음향 반향신호의 비율에 따라 감쇠된다. 실험하는 과정에서 음향 반향신호는 원단화자의 음성신호에 비해서 약 20dB 감쇠되는 것을 확인했다. 따라서 근단화자의 음성신호와 음향 반향신호가 동시에 발생한 동시통화시에는 $E[n^2(k)] \gg E[y^2(k)]$ 가 되어 식(18)과 같아지고 추정 오차신호는 거의 감쇠되지 않고 전송된다.

그런데 NLMS 알고리즘과 같이 주변 잡음에 의해서 계수 오조정이 심한 알고리즘은 적응 필터의 출력이 반향 신호를 제대로 추정하지 못하기 때문에 잔여 반향이 많이 생길 뿐만 아니라 전송되어야 할 근단화자의 음성신호가 왜곡된다. 이 경우 마이크 입력신호와 추정 오차신호 간의 상관도가 낮아지기 때문에 근단화자의 음성신호가 왜곡된다. 따라서 제안하는 후처리가 효과적으로 동작하기 위해서는 잡음에 대해서 강건한 알고리즘과 연동하는 것이 바람직하다. 본 논문에서는 식(8)과 같이 적응상수를 적응필터의 입력신호와 마이크 입력신호의 전력합으로 정규화함으로써 후처리에 의한 근단화자의 음성신호 왜곡을 줄인다.

본 논문에서는 후처리의 간단한 구현을 고려하여 running estimate를 이용해서 상관도와 전력을 추정한다. 그림 3은 추정 오차신호와 마이크 입력신호의 상관도를 프레임 단위로 평균하여 추정하는 것과 running estimate로 추정하는 것을 비교했다. 제안한 알고리즘으로 적응필터를 적용시키고 적응필터의 탭수는 256탭으로 하고, 정규화 적응상수는 α 는 0.2로 두었다. 원단화자신호와 근단화자 신호는 남자 음성신호를 8kHz로 샘플링하고 16비트로 양자화했다. 15000샘플에서 22000샘플 사이에 근단화자의 음성신호를 반향과 함께 동시녹음했다. 30msec를 8kHz로 샘플링하여 한 프레임(240샘플)에 대해서 평균하여 상관도를 구하고 running estimate로는 망각지수를 0.998로 두고 추정한다. 추정한 결과 프레임 단위로 추정한 것과 running estimate의 결과가 유사한 결과를 나타낸다.

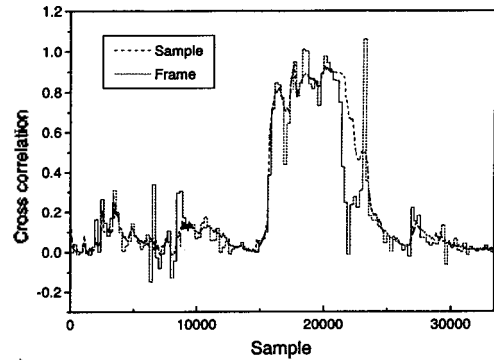


그림 3. 추정 오차신호와 마이크 입력신호의 상관도 추정
Fig. 3. Estimating correlation of estimate error signal and microphone input signal.

3.3 계산량 비교

음향 반향 제거기 알고리즘을 정리하면 표 1과 같다. 제안하는 음향 반향 제거기는 기존의 NLMS 알고리즘을 사용하고 동시통화 검출기를 이용한 방법에 비해서 계산량이 크게 증가하지 않는다. 표 2는 NLMS 알고리즘만을 이용한 음향 반향 제거기와 제안하는 방법과 계산량을 비교했다.

제안하는 음향 반향 제거기는 NLMS 알고리즘보다 나눗셈이 1회 증가하고 곱셈과 덧셈이 각각 7회, 5회가 증가한다. 그러나 적응 FIR 필터를 이용해서 음향 반향 제거기를 구현에 있어 중요한 요소는 필터 탭수이다. 적응 필터의 탭수를 256으로 하는 경우에 대해서 곱셈 계산량을 비교하면 NLMS 알고리즘이 516회이고 제안하는 알고리즘이 523회이다. 계산량의 증가는 1.34% 이다.

기존의 음향 반향 제거기는 일반적으로 동시통화 검출기를 동반하게 되는데 동시통화 검출기에 대한 계산량을 포함하지 않은 NLMS 알고리즘에 비해서 크게 계산량이 증가하지 않는다. 그리고 8kHz로 음성신호를 샘플링하고 곱셈과 덧셈을 1회 연산에 수행한다면 약 4.2MIPS 정도면 256탭의 적응필터를 적용시킬 수 있다.

표 1. 제안하는 후처리를 이용한 음향 반향 제거 알고리즘
Table 1. The proposed AEC algorithm using a post-processing.

$$\begin{aligned} e(k) &= d(k) - W^T(k) X(k) \\ P_x(k) &= \beta P_x(k) + (1 - \beta)x^2(k) \\ P_d(k) &= \beta P_d(k) + (1 - \beta)d^2(k) \\ P_{ed}(k) &= \beta P_{ed}(k) + (1 - \beta)e(k)d(k) \\ \mu(k) &= \frac{\alpha}{L(P_x(k) + P_d(k))} \\ W(k+1) &= W(k) + \mu(k)e(k)X(k) \\ \rho(k) &= \frac{P_{ed}(k)}{P_d(k)} \\ \hat{n}(k) &= \rho(k)e(k) \end{aligned}$$

표 2. 계산량 비교

Table 2. Comparison of computations.

Algorithm	Multiplication	Addition	Division
NLMS	2L+4	L+4	1
Proposed	2L+11	L+9	2

IV. 시뮬레이션 및 검토

본 논문에서는 멀티미디어 PC의 사운드 카드를 전이 중 방식(full-duplex)으로 설정하고 원단화자신호를 스피커로 출력하면서 동시에 마이크로 입력된 반향신호를 녹음하여 반향을 제거하는 시뮬레이션을 한다. 적응필터의 차수는 256차로 지정하고 정규화 적응상수는 α 는 0.2로 두었다. 적응필터의 입력신호의 전력 $P_x(k)$ 과 마이크로 입력신호의 전력 $P_y(k)$ 는 running estimate로 추정한다. 각 전력을 추정하기 위한 망각지수 β 는 0.998로 두고 실험했다. 원단화자의 음성 신호는 8kHz로 샘플링하고 16비트로 양자화했다.

성능평가는 잔여 반향신호와 음향 반향 제거 성능 평가에 일반적으로 사용되는 ERLE(echo return loss enhancement)를 비교한다. ERLE는 식(19)과 같이 계산한다.

$$ERLE(k) = 10 \log_{10} \frac{\sum_{i=0}^{L-1} d^2(n-i)}{\sum_{i=0}^{L-1} (d(n-i) - \hat{y}(n-i))^2} \quad (19)$$

4.1 동시통화가 발생한 경우

그림 4와 그림 5 원단화자의 음성신호와 근단화자의 음성신호가 동시에 있는 경우에 대한 시뮬레이션 결과이다. 원단화자의 음성신호를 스피커로 출력하면서 남성 근단화자의 음성신호를 동시녹음하여 마이크로 입력신호를 만들어 시뮬레이션 한다.

그림 4의 (a)는 반향 신호와 근단화자의 음성신호가 합해진 마이크로 입력신호를 나타낸다. 약 16000~22000 샘플사이 근단화자의 음성신호가 나타난다. (b)는 적응필터를 이용해서 반향을 제거한 결과이다. 근단화자의 음성신호에 비해서 작지만 잔여 반향이 근단화자의 음성신호가 있는 구간 전후에 남아있는 것을 볼 수 있다(남아있는 잔향은 귀로 들을 수 있을 정도의 크기이다). (c)에 후처리가 잔여 반향을 제거하기 위한 감쇠율을 표시하였다. 근단화자의 음성신호가 있는 구간에서는 감쇠율이 거의 1에 가까운 값이고 반향만 있는 구간에서는 감쇠율이 거의 0에 가까운 값을 나타낸다. 따라서 근단화자의 음성신호의 감쇠없이 잔여 반향을 제거할 수 있다. (d)는 (c)의 감쇠율로 후처리를 한 결과이다. (b)와 비교하면 잔여 반향은 거의 제거가 되고 근단화자의 음성신호는 거의 왜곡이 생기지 않았다.

그림 5에서는 후처리 전후에 대한 ERLE를 나타낸다. 점선과 실선은 각각 후처리 전과 후처리 후에 대한 ERLE이다. 후처리를 한 후 ERLE가 근단화자의 음성신호가 있는 구간을 제외한 영역에서 약 25dB 정도 개선되었다. 근단화자의 음성신호가 있는 구간에서는 후처리 전이나 후에도 ERLE가 0dB에 가까운 값으로 근단화자의 음성신호가 거의 손상없이 출력되고 있음을 알 수 있다.

따라서 제안하는 음향 반향 제거기는 별도의 동시통화 검출기를 사용하지 않고 간단한 후처리를 추가하여 근단화자의 음성신호 손상이 없이 잔여 반향을 효율적으로 제거할 수 있음을 시뮬레이션을 통해서 확인했다. 주변 잡음에 대해서 강건한 적응 알고리즘으로 적용할 때 후처리를 이용해서 효율적으로 잔향을 제거할 수 있다.

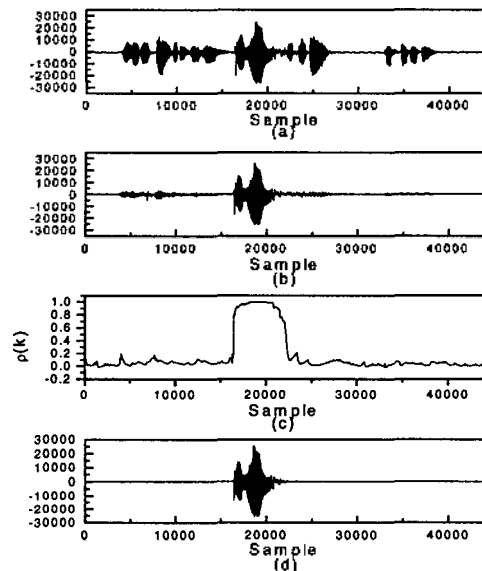


그림 4. 잔여 반향 제거를 위한 후처리 과정; (a) 반향 (b) 잔여 반향 (c) 감쇠율 (d) 후처리 결과
Fig. 4. Post-processing for reducing residual echo; (a) echo, (b) residual echo, (c) ratio of attenuation, (d) result of post-processing.

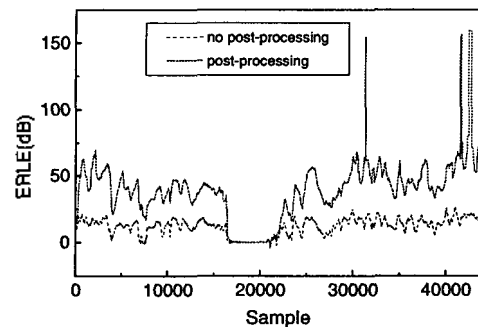


그림 5. ERLE 비교
Fig. 5. ERLE comparison.

4.2 반향 경로가 변하는 경우

반향 경로가 변하는 경우에 대한 시뮬레이션은 약 20000~30000샘플사이 에 스피커와 마이크 사이를 장애물로 막아 반향 경로가 변화하도록 하고 전이중 방식으로 녹음하여 시뮬레이션했다. 그림 6과 그림 7에서는 반향 경로가 변하는 상황에서 후처리의 감쇠율과 후처리 결과를 나타낸다. 그림 6의 (a)는 반향 신호를 그린 것이다. (b)는 적응필터의 출력으로 잔여 반향을 나타낸다. 그림 4의 (b)와 비교하면 20000~30000샘플 사이에서 잔여 반향의 에너지가 다소 높아진 것을 볼 수 있다. 이는 반향 경로의 변화를 적응 필터가 추정해 가는 과정이기 때문에 정확하게 추정하지 못한 결과이다. (c)는 후처리의 감쇠율을 나타낸다. 전 구간에 걸쳐서 거의 0에 가까운 값이다. (d)는 후처리를 한 결과이다. 잔여 반향이 대부분 감쇠되어 제거되었다.

그림 7은 반향 경로가 변하는 경우에 대한 시뮬레이션 결과를 ERLE로 표시한 것이다. 점선은 후처리하지 않은 경우이고 실선은 후처리한 결과이다. 반향 경로가 변하는 경우에도 전체적으로 약 25dB 정도 ERLE 가 개선되었다. 제안하는 후처리는 반향경로가 변화에 영향을 받지 않고 잔여 반향을 제거할 수 있다는 것을 확인했다.

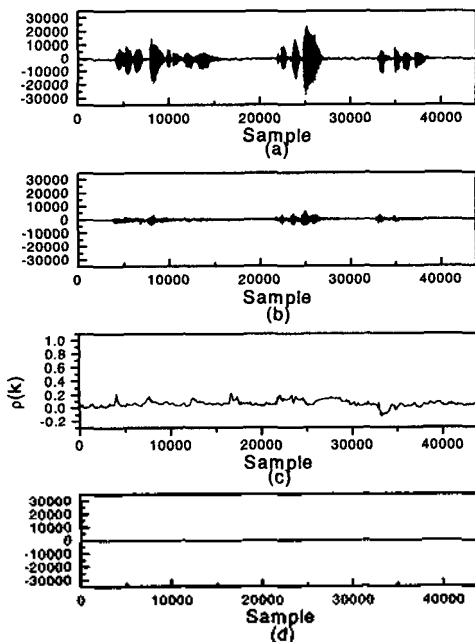


그림 6. 잔여 반향 제거를 위한 후처리 과정
 (a) 반향 (b) 잔여 반향 (c) 감쇠율 (d) 후처리 결과
 Fig. 6. Post-processing for reducing residual echo;
 (a) echo, (b) residual echo, (c) ratio of attenuation,
 (d) result of post-processing.

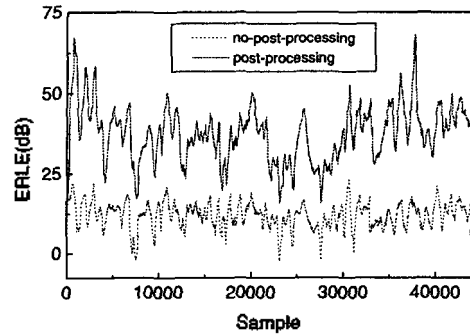


그림 7. ERLE 비교
 Fig. 7. Comparison ERLE.

V. 결 론

음향 반향 제거기는 핸드프리 단말기를 이용한 음성 통신에서 필수적인 요소이다. 음향 반향 제거기는 주로 적응필터 기법을 이용하고 있으나 근단화자의 음성신호를 포함한 주변 잡음신호에 의해서 적응필터의 계수가 오조정되어 음향 반향 제거 성능이 저하된다. 그리고 반향 경로를 정확하게 추정하지 못해서 상당한 잔여 반향이 남는다.

본 논문에서는 잡음에 강건한 알고리즘과 연동되는 후처리를 이용해서 음향 반향 제거기의 성능을 향상시키는 방법을 제안한다. 제안하는 강건한 알고리즘은 LMS 알고리즘의 적용 상수를 입력신호와 마이크 입력신호 전력의 합으로 정규화한다. 마이크에 근단화자의 음성신호가 없다면 제안하는 알고리즘은 기존의 NLMS 알고리즘과 같이 동작하고 근단화자의 음성신호가 있는 경우에는 근단화자의 음성신호의 전력만큼 적용 상수가 작아진다. 따라서 적응필터의 계수가 근단화자의 음성신호에 의해서 오조정되는 것을 줄일 수 있다. 또한 잡음에 강건한 알고리즘과 연동하여 동작하는 후처리를 통해서 잔여 반향을 제거하는 방법을 제안했다. 후처리는 추정 오차신호와 마이크 입력신호의 상관도를 마이크 입력신호로 정규화하여 추정 오차신호를 감쇠시킴으로써 잔여 반향을 제거한다. 적응필터가 적절히 수렴한 상태에서 마이크 입력으로 반향 음향신호만 있다면 추정 오차신호와 마이크 입력신호 사이의 상관도는 작아진다. 그리고 근단화자의 음성신호가 발생했다면 추정 오차신호에는 근단화자의 음성신호가 대부분을 차지함으로써 상관도는 1에 가까워진다. 따라서 잔여 반향은 감쇠되어 제거되고 근단화자의 음성신호는 감쇠없이 전송된다.

제안하는 음향 반향 제거기는 계산량이 기존의 NLMS 알고리즘에 비해서 크게 증가하지 않기 때문에 구현이 용이하다. 제안하는 음향 반향 제거기를 바탕으로 스테레오 음향 반향 제거기를 음성 CODEC 과 함께 구현하는 연구를 계속 하고자 한다.

참고문헌

1. S. Makino, "The Past, Present, and Future of Audio Signal Processing-Acoustic Echo Cancellation", IEEE Signal Processing Magazine, pp.39-41, Sep., 1997.
2. 조영민, 이원철. "고속 Affine Projection 알고리즘을 이용한 스테레오 음향 반향 제거", 한국음향학회 논문지 17권 1호, pp. 86-97, 1998년 1월.
3. S. H. Jensen, "Acoustic Echo Canceller for Hands-free Mobile Radiotelephony", Proc. EUSIPCO, Belgium, pp.1629-1632, 1992.
4. S. M. Kuo, Y. C. Huang, and Z. Pan, "Acoustic Noise and Echo Cancellation Microphone System for Videoconferencing", IEEE Trans. Consumer Electronics, VOL. 41, NO. 4, pp. 1195-1150, Nov., 1995.
5. A. Hirano and A. Sugiyama, "A Noise-Robust Stochastic Gradient Algorithm with An Adaptive Step-Size Suitable for Mobile Hands-Free Telephones", Proc. of ICASSP95, pp.1392-1395, 1995.
6. Y. Hiratani, A. Hirano and M. Kanazawa, "A Noise-Robust Echo Canceller on V830 Multimedia RISC Processor Intergrated into a Car Navigation System", Proc. of ICASSP 98, pp.1314-1317, 1998.
7. B. Ayad, G. Faucon and R. Le Bouguin-Jeannes, "Optimization of a Noise Reduction Preprocessing in An Acoustic Echo and Noise Controller", Proc. ICASSP 96, pp.953-956, 1996.
8. C. Antweiler, J. Grunwald and H. Quark, "Approximation of Optimal Step Size Control for Acoustic Echo Cancellation", Proc. of ICASSP97, pp.295-298, 1997.
9. P. Heitkamper, "An Adaptation Control for Acoustic Echo Cancellers", IEEE Signal Processing Letters. VOL. 4, NO. 6, pp. 170-172, June, 1997.
10. H. Ye and B. X. Wu, "A New Double-Talk Detection Algorithm Based on the Orthogonality Theorem", IEEE Trans. Communications, VOL. 39, NO 11, pp. 1542-1545, Nov., 1991.
11. B. Widrow and S. D. Stern, Adaptive Signal Processing, Prentice-Hall, p.26, 1985.
12. K. Ozeki and T. Umeda, "An Adaptive Filtering Algorithm Using an Orthogonal projection to an affine subspace and its property", Electron. Comm. Japan, VOL. J67-A, NO. 2, pp.126-132, 1984.(Japanese)
13. S. L. Gay and S. Tavathia, "The Fast Affine Projection Algorithm", Proc. ICASSP 95, pp. 3023-3026, 1995.
14. Y. Guelou, Abdelkrim Benamar and Pascal Scalart, "Analysis of Two Structure for Combined Acoustic Echo Cancellation and Noise Reduction", ICASSP 96, Atlanta Georgia, VOL. 2, pp. 637-640, May, 1996.
15. T. Wigren, "Fast Converging and Low Complexity Adaptive Filtering Using an Averaged Kalman Filter", IEEE Trans. Signal Processing, VOL. 46, NO. 2, pp. 515-518, Feb., 1998.
16. M. Tanaka, Y. Kaneda, S. Makino and J. Kojima, "Fast Projection Algorithm and Its Step Size Control", Proc. ICASSP 95, pp. 945-948, 1995.
17. S. Haykin, Adaptive Filter Theory, Prentice Hall, N.J., 1995
18. E. Hansler, "Adaptive Echo Compensation Applied to the Hands-free Telephone Problem", Proc. ISCAS, pp.279-282, 1990.
19. I. Furukawa, "A Design of Canceller of Broad Band Acoustic Echo", Int. Teleconference Symposium, Tokyo, pp. 1-8, 1984.
20. K. Maouche and D. T.M. Slock, "A Fast Instrumental Variable Affine Projection Algorithm", ICASSP 98, pp. 2386-2389, 1998.
21. W. Kellermann, "Analysis and Design of Multirate Systems for Cancellation of Acoustical Echoes", Proc. ICASSP 88, pp. 2570-2573, 1988.
22. A. Gilloire and M. Vetterli, "Adaptive Filtering in Subbands with Critical Sampling: Analysis, Experiments and Application to Acoustic Echo Cancellation", IEEE Trans. Signal Processing, VOL 40., NO. 8, pp. 1862-1875, Aug., 1992
23. M. M. Sondhi and D. A. Berkley, "Silencing Echoes on the Telephone Network", proc. IEEE, VOL. 68, pp.948-963, 1980.
24. B. Widrow, et. al, "Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter", Proc. IEEE, VOL. 64, NO. 8, pp.1151-1162, Aug., 1976.
25. B. Widrow, et. al, "Adaptive Noise Cancelling: Principles and Applications", Proc. IEEE, VOL. 63, NO. 12, pp.1692-1716, Dec., 1975.

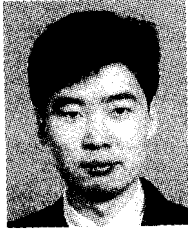
▲박장식(Jang Sik Park)



1992년 2월 : 부산대학교 공과대학
전자공학과(공학사)
1994년 2월 : 부산대학교 대학원 전
자공학과(공학석사)
1999년 2월 : 부산대학교 대학원 전
자공학과(공학박사)
1997년 3월~현재 : 동의공업대학 영상
정보처리과 조교수

※주관심분야: 디지털신호처리, 적응신호처리, 음향신호처리

▲김 현 태(Hyun Tae Kim)



1989년 2월 : 부산대학교 공과대학
전자공학과(공학사)

1995년 2월 : 부산대학교 대학원 전
자공학과(공학석사)

1998년 8월 : 부산대학교 대학원 전
자공학과(박사수료)

1990년 3월 ~ 1995년 7월 : 공군 기술
장교 근무

1995년 1월 ~ 1997년 6월 : 현대전자 산업전자연구소 연구원

1995년 6월 ~ 1997년 1월 : 경남지방중소기업청 공업연구사

1997년 2월 ~ 현재 : 부산·울산지방중소기업청 공업연구사

※주관심분야: 디지털신호처리, 적응신호처리, 음향신호처리

▲손 경 식(Kyung Sik Son)

한국음향학회지 1996년 15권, 2E호 참조

현재 : 부산대학교 전자공학과 교수