암묵신호분리를 이용한 동시통화 음향반향제거기

이행우* · 윤현민**

An Acoustic Echo Canceller for Double-talk by Blind Signal Separation

Haeng-Woo Lee* · Hyun-Min Yun**

요 약

본 논문은 암묵신호분리방법을 이용하여 동시통화를 가능하게 하는 음향반향제거기에 관한 것이다. 음향반향제거기는 동시통화 구간에서 성능이 저하되거나 발산하게 된다. 그래서 근단화자신호를 추정해서 잔차신호로부터 차감하기 위하여 암묵신호분리방법을 사용한다. 암묵신호분리방법은 이중 마이크를 가지고 2차 통계적 성질을 이용한 반복적인 계산에 의해 근단화자신호를 추정해낸다. 그런데 폐쇄된 반향환경에서 암묵신호분리의 혼합모델은 다채널이기 때문에 분리계수를 직접 계산하지 않고 반향제거기의 계수를 복사하여 그대로 사용한다. 많은 시뮬레이션을 통하여 제안한 음향반향제거기의 성능을 검증하였다. 시뮬레이션 결과, 이 방법을 사용한 음향반향제거기는 동시통화의 유무에 상관없이 안전하게 동작하고, 일반적인 LMS 알고리즘에 비해 ERLE가 평균 20dB 향상되는 것으로 나타났다.

ABSTRACT

This paper describes an acoustic echo canceller with double-talk by the blind signal separation. The acoustic echo canceller is deteriorated or diverged in the double-talk period. So we use the blind signal separation to estimate the near-end speech signal and to eliminate the estimated signal from the residual signal. The blind signal separation extracts the near-end signal with dual microphones by the iterative computations using the 2nd order statistical character. Because the mixture model of blind signal separation is multi-channel in the closed reverberation environment, we used the copied coefficients of echo canceller without computing the separation coefficients. By this method, the acoustic echo canceller operates irrespective of double-talking. We verified performances of the proposed acoustic echo canceller by simulations. The results show that the acoustic echo canceller with this algorithm detects the double-talk periods thoroughly, and then operates stably in the normal state without the divergence of coefficients after ending the double-talking. And it shows the ERLE of averagely 20dB higher than the normal LMS algorithm.

키워드

암묵신호분리, 음향반향제거기, 동시통화, 다채널, 적응알고리즘

Key word

Blind Signal Separation, Acoustic echo canceller, Double-talk, Multi-channel, Adaptive algorithm

* 정회원 : 남서울대학교(교신저자, hwlee@nsu.ac.kr)
** 정회원 : 재능대학교

접수일자 : 2011. 09. 14
심사완료일자 : 2011. 10. 04

Ⅰ. 서 론

핸즈프리 등 스피커-폰 시스템의 사용이 증가함에 따라서 음향반향제거기술이 중요한 문제로 대두하게 되었다. 음향반향신호는 원단화자로부터의 수신신호가 스피커를 통해 방사된 후 벽면 등에 반사되어 마이크에서 근단화자의 송신신호와 결합되므로서 원단에서 근단화자의 음성청취를 방해받게 된다. 반향신호를 제거하기 위하여 적응 필터를 이용한 음향반향제거기를 사용하게 되는데 이 회로는 원단화자신호만 존재하는 일방통화 때에만 정상적으로 동작하고 근단화자신호도 존재하는 동시통화 구간에서는 반향신호를 제거하는 능력이 갑자기 저하되는 현상이 발생한다. 따라서 이 동시통화문제를 처리할 수 있는 기능이 포함되어야 한다.[1][2]

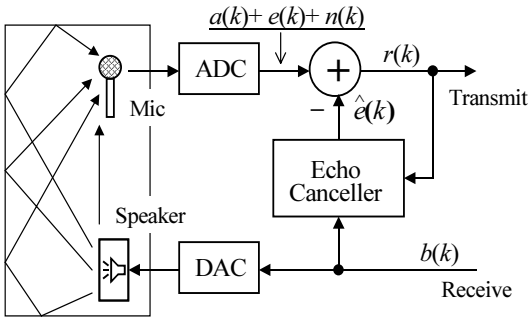

그림 1. 음향반향제거기
Fig. 1 Acoustic echo canceller

일반적으로 동시통화상태를 검출하여 반향제거기의 계수 갱신을 중지시킴으로서 계수의 발산을 방지하는 방법[3]이 있다. 그러나 이 방법은 검출시간이 비교적 길어 검출이 되기 전에 반향제거기의 계수가 발산하는 경우가 있다. 따라서 본 논문에서는 이러한 문제를 해결하기 위하여 새로운 음원분리기술을 사용한다.

음원분리기술은 관찰된 혼합음으로부터 각 음원신호를 분리해낼 수 있다. 특히 입력신호의 통계적인 특성 등 사전 정보를 알지 못하는 상태에서 신호들을 분리하는 블라인드 음원분리(Blind Signal Separation; BSS) 방법이 있다. 이것은 음성간의 통계적 특성이 독립이라는 가정하에 여러 음원을 분리하는 방법이다.[4]

Jutten과 Herault가 제안한 BSS 알고리즘은 선형 혼합된 신호를 간단히 분리할 수 있다고 발표하였으며 이후 여러 후속 연구가 이루어졌다.[5] Jutten과 Herault는 시간영역에서 잡음과 신호를 분리하는 BSS 알고리즘을 제안하였다. 그러나 이 알고리즘은 수천 개의 많은 탭을 사용하므로 큰 메모리가 필요하고 계산량도 많아진다. Bell과 Sejnowski는 독립성분분석(Independent Component Analysis; ICA)을 통하여 지연을 추정하는 방법을 제안하였다.[6] 그러나 음성신호는 주기적인 특성 때문에 지연을 추정하는데 있어 초기값에 따라서는 최적해로 수렴하지 못하는 경우도 있다. 이를 해결하기 위해 Torkkoa는 정보량 최대화 알고리즘을 이용하여 지연을 추정하였다.[7]

그러나 계산량이 많고 주변잡음이 있으면 정확하게 지연을 추정하지 못한다. 따라서 최근에는 주파수영역에서 블라인드 음원분리를 하는 연구가 많이 이루어지고 있다.[8] 또한 폐쇄된 공간에서 반향신호가 존재하는 다채널 환경에서 신호를 분리하는 문제는 다차원의 많은 계산량을 요구한다.[9][10][11][12]

제안하는 음원분리기술은 마이크 간에 신호의 크기와 도달시간의 차이를 이용한다. 다시 말해서 이 알고리즘은 여러 비선형 수학적 계산식을 구현한다. 복수의 음향신호는 3가지 차원에서 중첩됨으로 인하여 서로 혼합되고 간섭된다. 세 차원은 시간, 주파수, 그리고 도착 방향을 말한다. 지금까지 시간영역에서 중첩하는 신호는 주파수 스펙트럼이나 도착 방향을 이용하여 구별하였다. 그러나 본 기술은 시간, 스펙트럼, 도착 방향이 동시에 중첩된 신호들도 구분할 수 있다. 즉 2개의 음원으로부터 발생된 혼합음이 존재하는 환경에서 이중 마이크를 사용하여 각 음원을 분리해낸다. 이는 각 마이크에 유입되는 신호의 크기와 도달시간이 다르기 때문에 이러한 특성을 활용하여 반복적인 계산을 통해 음원을 구하게 된다.

본 논문은 다음과 같이 구성되었다. Ⅱ장에서 암묵신호분리 알고리즘의 개념을 논하였고, Ⅲ장에서는 다채널이 존재하는 음향반향제거기에서 암묵신호분리 알고리즘의 적용방법에 대하여 살펴보았다. 그리고 모의실험 결과를 Ⅳ장에서 기술하였고, Ⅴ장에서 결론을 도출하였다.

Ⅱ. 암묵신호분리 알고리즘

폐쇄된 공간에서 핸즈프리 통화는 음향반향을 발생시킨다. 그래서 원하지 않는 반향신호를 제거하기 위한 회로를 사용한다. 그런데 원단화자와 근단화자가 동시에 말을 하는 동시통화의 경우에는 마이크로 유입되는 근단화자신호로 인하여 반향제거기의 계수가 발산할 수 있다. 이를 방지하기 위하여 계수를 갱신하는데 사용되는 잔차신호에서 근단화자신호성분을 제거하고자 한다. 마이크로 유입되는 반향신호와 근단화자신호는 암묵신호분리를 통하여 근단화자신호를 구할 수 있다. 동시통화시 dual 마이크를 사용한 암묵신호분리는 그림 2와 같은 구조로 이루어진다.

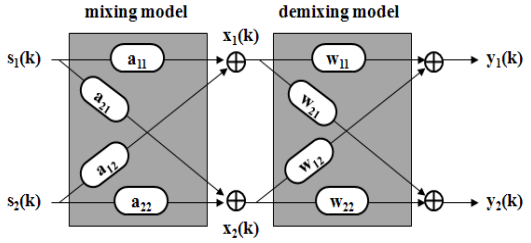

그림 2. 이중 마이크를 사용한 암묵신호분리 모델
Fig. 2 Blind signal separation model with dual microphones

각 음원으로부터 발생된 음향신호는 크기와 시간을 달리하여 마이크로 유입된다. 따라서 음원을 $s(k)$, 마이크 입력신호를 $x(k)$라 하면 각 마이크로 유입되는 신호의 혼합모델은 식 (1), (2)과 같이 표현된다.

$$x_1(k) = a_{11}(k)s_1(k) + a_{12}(k)s_2(k-d_2) \tag{1}$$

$$x_2(k) = a_{21}(k)s_1(k-d_1) + a_{22}(k)s_2(k) \tag{2}$$

여기서 $a(k)$는 각 음원으로부터 마이크까지의 거리에 반비례하는 혼합계수이다. 샘플링 주파수가 8kHz인 경우 두 마이크 간의 거리가 약 4㎝이내이면 지연시간은 무시할 수 있다. 그리고 반향신호는 다채널 신호의 합이다. 즉

$$s_2(k) = \sum_{n=1}^{N} c_n(k)s_2(k-n) \tag{3}$$

여기서 $c_n(k)$는 계산하지 않고 반향제거기의 계수를 복사하여 사용한다. BSS 출력에서 분리된 음향의 크기는 음원의 크기와 같지 않으므로 각 음원을 $s'_1(k) = s_1(k)/a_{11}(k)$, $s'_2(k) = s_2(k)/a_{22}(k)$로 정규화하면 수식이 간략화된다. 따라서

$$x_1(k) = s'_1(k) + a_{12}(k)s'_2(k) \tag{4}$$

$$x_2(k) = a_{21}(k)s'_1(k) + s'_2(k) \tag{5}$$

두 입력신호를 구하는 형태로 수식을 정리하면

$$\tilde{s}_1(k) = x_1(k) - \tilde{a}_{12}(k)\tilde{s}_2(k) \tag{6}$$

$$\tilde{s}_2(k) = x_2(k) - \tilde{a}_{21}(k)\tilde{s}_1(k) \tag{7}$$

여기서 미지수 $\tilde{a}_{12}(k)$ 및 $\tilde{a}_{21}(k)$는 반복적인 계산에 의하여 해를 구할 수 있다. 즉 음향신호 $s_1(k)$, $s_2(k)$가 서로 독립이라 할 때 각 신호의 2차 통계적 특성을 이용하여 steepest descent method로 구한다. 한편 출력신호는 음원의 추정치 $y_1(k) = \tilde{s}_1(k)$, $y_2(k) = \tilde{s}_2(k)$이므로 식 (6), (7)로부터 출력신호의 분리모델은

$$y_1(k) = x_1(k) - \tilde{a}_{12}(k)y_2(k) \tag{8}$$
$$= \frac{x_1(k) - \tilde{a}_{12}(k)x_2(k)}{1 - \tilde{a}_{12}(k)\tilde{a}_{21}(k)}$$

$$y_2(k) = x_2(k) - \tilde{a}_{21}(k)y_1(k) \tag{9}$$
$$= \frac{-\tilde{a}_{21}(k)x_1(k) + x_2(k)}{1 - \tilde{a}_{12}(k)\tilde{a}_{21}(k)}$$

여기서 분리모델의 계수 행렬은

$$\mathbf{w}(k)=\frac{1}{1-\tilde{a}_{12}(k)\tilde{a}_{21}(k)}\begin{bmatrix}1 & -\tilde{a}_{12}(k)\\ -\tilde{a}_{21}(k) & 1\end{bmatrix} \quad (10)$$

분리계수를 구하는 대신 식 (4), (5)로부터 혼합계수를 직접 구할 수 있다. 오차신호를 비용함수로 사용하기 위하여 혼합모델을 다시 표현하면

$$x_1(k)-y_1(k)-\tilde{a}_{12}(k)y_2(k)=\varepsilon_1(k) \quad (11)$$

$$x_2(k)-\tilde{a}_{21}(k)y_1(k)-y_2(k)=\varepsilon_2(k) \quad (12)$$

따라서 LMS 알고리즘을 이용한 갱신값은

$$\nabla\rho_1(k)=\frac{\partial\{\delta_1(k)\}^2}{\partial a_{12}(k)}=-2[x_1(k)-y_1(k)-\tilde{a}_{12}(k)y_2(k)]y_2(k) \quad (13)$$

$$\nabla\rho_2(k)=\frac{\partial\{\delta_2(k)\}^2}{\partial a_{21}(k)}=-2[x_1(k)-y_1(k)-\tilde{a}_{12}(k)y_2(k)]y_1(k) \quad (14)$$

그리고 새로운 계수의 갱신식은

$$\tilde{a}_{12}(k+1)=\tilde{a}_{12}(k)-\mu_1\nabla\rho_1(k)=\tilde{a}_{12}(k)+2\mu[x_1(k)-\tilde{s}_1(k)-\tilde{a}_{12}(k)\tilde{s}_2(k)]\tilde{s}_2(k) \quad (15)$$

$$\tilde{a}_{21}(k+1)=\tilde{a}_{21}(k)-\mu_2\nabla\rho_2(k)=\tilde{a}_{21}(k)+2\mu[x_1(k)-\tilde{s}_1(k)-\tilde{a}_{12}(k)\tilde{s}_2(k)]\tilde{s}_1(k) \quad (16)$$

여기서 μ는 수렴속도를 결정하는 스텝 크기를 나타낸다. BSS 출력을 사용하여 식 (6), (7)을 다시 표현하면

$$y_1(k+1)=x_1(k)-\tilde{a}_{12}(k)y_2(k) \quad (17)$$

$$y_2(k+1)=x_2(k)-\tilde{a}_{21}(k)y_1(k) \quad (18)$$

따라서 식 (15), (16)과 식 (17), (18)을 각 샘플마다 반복적으로 계산하면 분리하고자 하는 음원을 구할 수 있다. 이러한 순환적 연산과정을 그림 3에서 도식적으로 보여주고 있다.

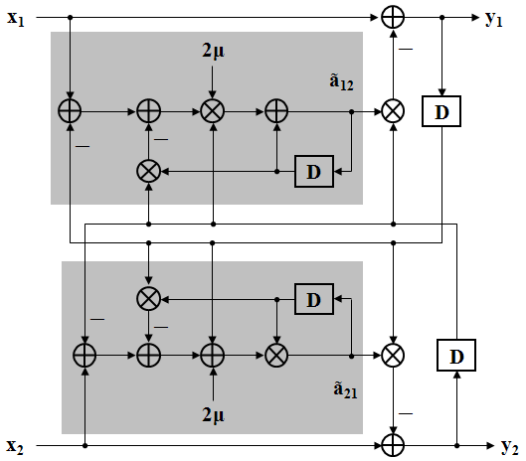

그림 3. 신호분리과정
Fig. 3 Signal demixing structure

Ⅲ. 음향반향제거기의 암묵신호분리

스피커-폰 시스템에서 발생하는 동시통화 문제를 해결하기 위하여 BSS 방법을 사용하고자 한다. 제안하는 BSS를 사용한 음향반향제거기를 그림 4에서 보여주고 있다.

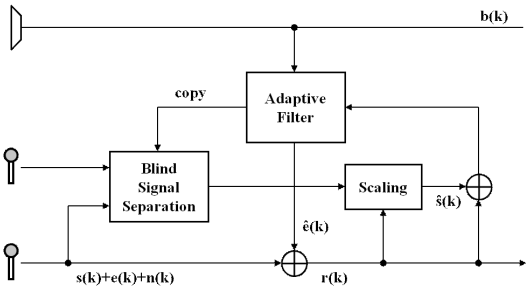

그림 4. BSS를 사용한 음향반향제거기
Fig. 4 Acoustic echo canceller with BSS

동시통화 구간에서 음향반향제거기의 계수는 잔차신호에 포함된 근단신호로 인하여 수렴상태에서 이탈하여 발산하고 수렴상태로 복귀하지 못할 수도 있다. 따라서 BSS를 사용하여 근단화자신호를 추출하고 잔차신호에서 근단화자신호를 제거함으로써 반향제거기의 계수가 수렴상태를 유지하도록 한다. 이때 BSS 출력의 크기의 모호성이 문제가 된다. 이를 위해 근단화자신호의 잔차신호의 평균값을 구하여 그 크기에 따라 분리된 근단화자신호의 크기를 조정한다. 앞에서 설명한 비정상상태 검출방법이 안고 있는 문제점을 해결하기 위하여 다음과 같은 새로운 검출방법을 제안하고자 한다.

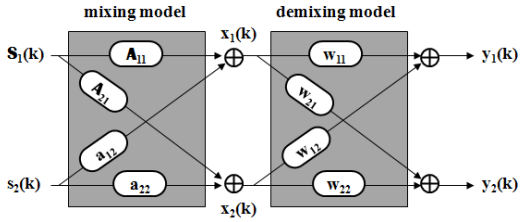

그림 5. 다채널 반향에 대한 암묵신호분리의 순방향모델
Fig. 5 Feedforward structure of blind signal separation with multi-channel echo

스피커에서 발생한 음향반향신호는 다채널의 반향경로를 거쳐 마이크로 유입된다. 스피커신호의 행렬을 $\mathbf{S}_1(k)$, 근단화자신호를 $s_2(k)$, 그리고 반향신호에 대한 혼합계수의 행렬을 $\mathbf{A}_{11}(k)$ 및 $\mathbf{A}_{21}(k)$라 하면 혼합모델은 식 (19)와 (20)으로 표현된다.

$$x_1(k)=\mathbf{A}_{11}^{\mathrm{T}}(k)\mathbf{S}_1(k)+a_{12}(k)s_2(k) \tag{19}$$

$$x_2(k)=\mathbf{A}_{21}^{\mathrm{T}}(k)\mathbf{S}_1(k)+a_{22}(k)s_2(k) \tag{20}$$

위 식에서 반향신호의 혼합계수는 음향반향제거기의 계수와 동일하다. 그리고 만일 두 마이크가 근접거리에 위치해 있다면 혼합계수의 행렬은 다음과 같이 근사화될 수 있다. 즉,

$$\mathbf{A}_{21}(k)\cong\eta\mathbf{A}_{11}(k) \tag{21}$$

여기서 η는 거의 1에 가까운 상수이다. 따라서 혼합모델을 다시 쓰면

$$x_1(k)=\mathbf{C}^{\mathrm{T}}(k)\mathbf{S}_1(k)+a_{12}(k)s_2(k) \tag{22}$$

$$x_2(k)=\eta\mathbf{C}^{\mathrm{T}}(k)\mathbf{S}_1(k)+a_{22}(k)s_2(k) \tag{23}$$

BSS의 출력에서 분리된 신호의 크기는 음원의 크기와 동일하지 않기 때문에 음원을 $\mathbf{S}_1'(k)=\mathbf{S}_1(k)/c_0(k)$, $s_2'(k)=s_2(k)/a_{22}(k)$과 같이 정규화함으로써 수식을 간략화한다. 그래서

$\mathbf{C}^{0\mathrm{T}}(k)\mathbf{S}_1'^{0}(k)=\sum_{n=1}^{N-1}c_n(k)s_1(k-n)/c_0(k)$라 하면

$$x_1(k)=s_1'(k)+\mathbf{C}^{0\mathrm{T}}(k)\mathbf{S}_1'^{0}(k)+a_{12}(k)s_2'(k) \tag{24}$$

$$x_2(k)=\eta\mathbf{C}^{\mathrm{T}}(k)\mathbf{S}_1'(k)+s_2'(k) \tag{25}$$

위 식에서 두 입력신호를 구하는 형태로 정리하면

$$\tilde{s}_1(k)=x_1(k)-\mathbf{C}^{0\mathrm{T}}(k)\tilde{\mathbf{S}}_1^{0}(k)-\tilde{a}_{12}(k)\tilde{s}_2(k) \tag{26}$$

$$\tilde{s}_2(k)=x_2(k)-\eta\mathbf{C}^{\mathrm{T}}(k)\tilde{\mathbf{S}}_1(k) \tag{27}$$

출력신호는 $y_1(k)=\tilde{s}_1(k)$, $y_2(k)=\tilde{s}_2(k)$, 그리고 $\mathbf{C}^{\mathrm{T}}(k)\mathbf{Y}_1(k)=x_1(k)-\tilde{a}_{12}(k)y_2(k)$이므로

$$\begin{aligned}y_1(k)&=x_1(k)-\mathbf{C}^{0\mathrm{T}}(k)\mathbf{Y}_1^{0}(k)-\widetilde{a_{12}}(k)y_2(k)\\&=-\mathbf{C}^{0\mathrm{T}}(k)\mathbf{Y}_1^{0}(k)+\eta\widetilde{a_{12}}(k)\mathbf{C}^{\mathrm{T}}(k)\mathbf{Y}_1(k)\\&\quad+x_1(k)-\widetilde{a_{12}}(k)x_2(k)\end{aligned} \tag{28}$$

$$\begin{aligned}y_2(k)&=x_2(k)-\eta\mathbf{C}^{\mathrm{T}}(k)\mathbf{Y}_1(k)\\&=\eta\tilde{a}_{12}(k)y_2(k)-\eta x_1(k)+x_2(k)\end{aligned} \tag{29}$$

그러므로 식 (28), (29)를 정리하여 역방향 형태로 분리모델을 표현할 수 있다.

$$\begin{aligned}y_2(k)=\ &-\mathbf{C}^{0\mathrm{T}}(k)\mathbf{Y}_2^{0}(k)+\eta\tilde{a}_{12}(k)\mathbf{C}^{\mathrm{T}}(k)\mathbf{Y}_2(k)\\&-\eta\mathbf{C}^{\mathrm{T}}(k)\mathbf{X}_1(k)+\mathbf{C}^{\mathrm{T}}(k)\mathbf{X}_2(k)\end{aligned} \tag{30}$$

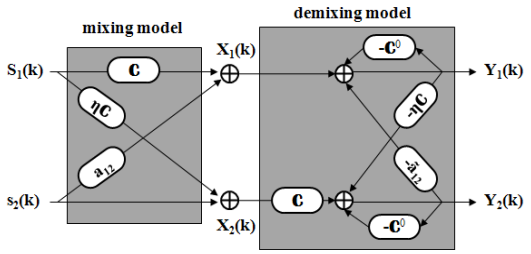

그림 6. 다채널 반향에 대한 암묵신호분리의 역방향모델

Fig. 6 Feedback structure of blind signal separation with multi-channel echo

혼합계수를 구하기 위한 비용함수로서 식 (28)로부터의 오차신호를 사용하면

$$x_1(k)-\mathbf{C}^{\mathrm{T}}(k)\mathbf{Y}_1(k)-\tilde{a}_{12}(k)y_2(k)=\varepsilon(k) \tag{31}$$

따라서 LMS 알고리즘의 갱신값은

$$\begin{aligned}\nabla\rho(k)&=\frac{\partial\{\varepsilon(k)\}^2}{\partial\tilde{a}_{12}(k)}\\&=-2[x_1(k)-\mathbf{C}^{\mathrm{T}}(k)\mathbf{Y}_1(k)\\&\qquad-\tilde{a}_{12}(k)y_2(k)]y_2(k)\end{aligned} \tag{32}$$

그리고 혼합계수를 업데이트하는 식은

$$\begin{aligned}\tilde{a}_{12}(k+1)&=\tilde{a}_{12}(k)-\mu\nabla\rho(k)\\&=\tilde{a}_{12}(k)+2\mu_1[x_1(k)-\mathbf{C}^{\mathrm{T}}(\mathrm{k})\mathbf{Y}_1(\mathrm{k})\\&\qquad-\tilde{a}_{12}(k)y_2(k)]y_2(k)\\&=\tilde{a}_{12}(k)[1-2\mu_1y_2^2(k)]+2\mu_1[x_1(k)\\&\qquad-\mathbf{C}^{\mathrm{T}}(\mathrm{k})\mathbf{Y}_1(\mathrm{k})]\mathrm{y}_2(\mathrm{k})\end{aligned} \tag{33}$$

또한, LMS 알고리즘을 이용하는 반향제거기와 관련된 수식은

$$\hat{e}(k)=\sum_{n=0}^{N-1}c_n(k)b(k-n) \tag{34}$$

$$d(k)=r(k)-\tilde{s}_2(k)=x_1(k)-\hat{e}(k)-\tilde{s}_2(k) \tag{35}$$

$$c_n(k+1)=c_n(k)+2\mu d(k)b(k-n) \tag{36}$$

여기서 $d(k)$는 잔차신호에서 BSS 추정 근단화자신호가 차감된 오차신호이다. 이 반향제거기는 몇가지 해결해야 할 문제들을 갖고 있다. 즉, BSS 출력신호에서 순서의 permutation 문제와 신호 크기의 scaling 모호성 문제가 있다. BSS 출력신호의 크기는 균일하지 않으며 입력신호와 같은 scale이 아니다. 추정 반향신호가 차감된 잔차신호로부터 분리된 근단화자신호의 scale 정보를 얻을 수 있다. BSS 출력신호의 scale을 정하기 위하여 분리된 근단화자신호 및 잔차신호의 이동평균을 계산한다. 그리고 이 평균값의 scale에 따라 분리된 근단화자신호의 크기를 조정한다.

$$y_m(k)=(1-\lambda)y_m(k)+\lambda|y_1(k)| \tag{37}$$

$$r_m(k)=(1-\lambda)r_m(k)+\lambda|r(k)| \tag{38}$$

$$y_1(k)=\frac{r_m(k)}{y_m(k)}y_1(k) \tag{39}$$

여기서 λ는 작은 상수이다. 또한, BSS 출력신호의 순서가 항상 일정하게 정해지지 않는 permutation 문제가 존재한다. 그러나 식 (28)과 (30)에서 추정 혼합계수 $\tilde{a}_{12}(k)$와 η의 크기에 따라 BSS 출력신호의 순서를 예측할 수 있다. 즉, $\tilde{a}_{12}(k)$와 η의 크기가 1 이하인 경우 $y_1(k)$에는 $s_1(k)$, $y_2(k)$에는 $s_2(k)$가 출력되고, 반대로 $\tilde{a}_{12}(k)$와 η의 크기가 1 이상인 경우 $y_1(k)$에는 $s_2(k)$, $y_2(k)$에는 $s_1(k)$가 출력된다. 이것은 계수행렬 $\mathbf{C}(k)$를 직접 계산하지 않고 반향제거기로부터 복사하여 사용하기 때문에 가능하다.

Ⅳ. 모의실험 결과

본 논문에서 제안한 복수의 마이크를 사용한 음원분리 및 반향제거 성능을 검증하기 위해 윈도우 XP에서 C 언어를 사용하여 시스템의 동작을 표현하는 시뮬레이션 프로그램을 작성하고, 이 알고리즘을 적용한

경우 반복된 심볼수에 대한 ERLE(Echo Return Loss Enhancement)의 수렴특성곡선을 Matlab을 이용하여 관찰하였다.

입력은 근단화자의 음성신호와 원단의 백색잡음 등 두 음원으로부터 각 마이크에 유입된 혼합신호를 사용하였다. 원단으로부터의 수신신호는 400Hz~4kHz의 주파수 사이에서 무색 성분을 갖는 잡음신호이다. 두 마이크는 최소 4cm 이상 간격을 두어야 하며, 입력신호는 8kHz로 샘플링된다. 그리고 계수 $\tilde{a}_{12}(k)$를 적응적으로 산출할 때 갱신 스텝의 크기는 $\mu = 2^{-13}$으로 설정하였다. 실험에 사용한 반향제거기는 128-tap FIR(Finite Impulse Response) 구조로서 NLMS(Normalized Least Mean Square) 적응 알고리듬을 사용하였으며 적응스텝의 크기는 μ=0.03으로 설정하였다. 이 실험에서 반향경로는 다음과 같은 수학적 모델을 사용하였다. 반향경로의 응답특성은 시간에 지남에 따라 크기가 지수함수적으로 감소하고 끝부분에서는 상당히 긴 꼬리가 지속되어 진다.

$$h(i) = 0.0914\,\mathrm{rand}(i)*(-1.0346)^{-i} \quad \text{for } i = 0,1,\cdots N-1 \tag{40}$$

여기서 $\mathrm{rand}(i)$는 시간에 따라 불규칙하게 (+1) 또는 (-1)을 갖는 함수이다. 본 기술은 두 음향신호가 혼합되어 유입될 때 둘 중 어느 하나의 신호를 분리해낸다. 먼저 음성신호와 잡음신호가 혼합되어 유입될 때 각 신호의 분리 실험을 수행하였다. 사용된 음성은 「실제 윈도우즈 프로그램 내에서」라는 여성의 발음이다. 다음 그림은 두 음원 및 음성신호를 분리한 경우와 잡음신호를 분리한 경우의 각 출력신호 파형을 보여주고 있다.

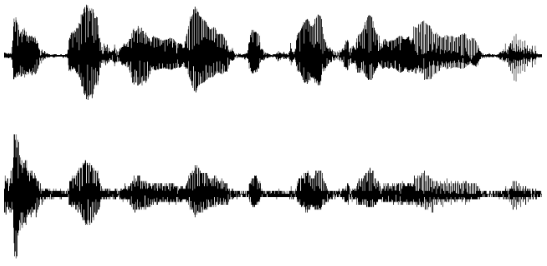

그림 7. 음성신호의 음원 및 분리된 출력신호
Fig. 7 Source and separated signal of speech

그림 8. 잡음신호의 음원 및 분리된 출력신호
Fig. 8 Source and separated signal of noise

이번에는 두 음성신호가 동시에 혼합되어 유입될 때 각 신호의 분리 결과에 대하여 살펴본다. 사용된 음성은 「실제 윈도우즈 프로그램 내에서」라는 여성의 발음과 「안녕하세요 황광일 입니다」라는 남성의 발음이다. 두 음원 A와 B 및 음성신호를 분리한 경우의 각 출력신호 파형을 다음 그림에서 보여주고 있다. 초기 계수가 수렴하지 않았을 때는 이상한 파형이 나타났으나 곧 수렴하여 분리된 음성신호를 보여준다.

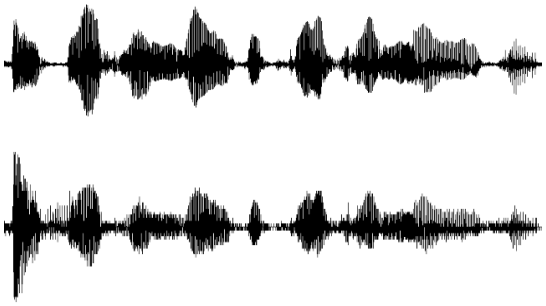

그림 9. 음성신호A의 음원 및 분리된 출력신호
Fig. 9 Source and separated signal of speech A

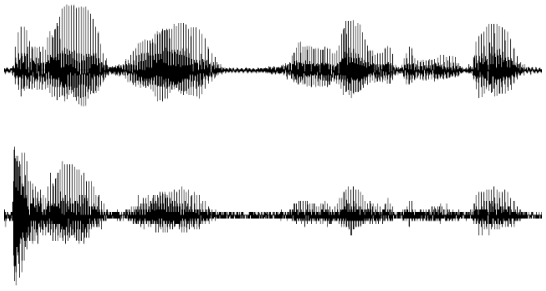

그림 10. 음성신호B의 음원 및 분리된 출력신호
Fig. 10 Source and separated signal of speech B

그림 11은 암묵신호분리에서 평균 오차신호 대 평균 근단화자신호의 비를 백분율로 나타낸다. 오차신호는 근단화자신호 입력과 신호분리된 근단화자신호 간의 차이다. 이 그림으로부터 BSS가 수렴되었을 때 평균 오차신호는 평균 근단화자신호의 5% 이하가 되는 것을 볼 수 있다.

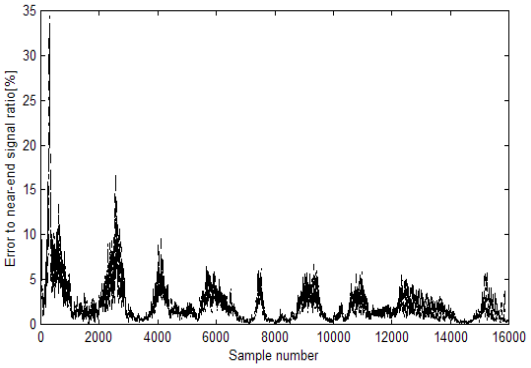

그림 11. 암묵신호분리에서 오차신호 대 근단화자신호의 비율
Fig. 11 Error to near-end signal ratio in BSS

그림 12는 음향반향제거기의 성능을 나타내는 ERLE 곡선을 보여준다. 점선으로 도시된 곡선은 일반적인 LMS 알고리즘을 사용했을 때의 ERLE를 표시하고, 실선으로 도시된 곡선은 암묵신호분리 알고리즘을 적용했을 때의 ERLE를 표시한다. 이 그림에서 제안한 알고리즘의 ERLE가 일반 LMS 알고리즘보다 평균 20dB 정도 향상되는 것을 볼 수 있다.

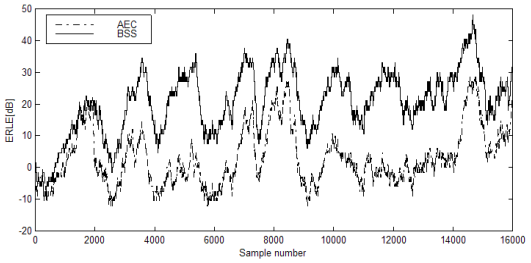

그림 12. 음향반향제거기에서 일반 LMS와 제안한 방법의 ERLE 곡선
Fig. 12 ERLE curves of the normal LMS and the proposed method in the AEC

V. 결 론

본 논문은 음향반향제거기에서 복수의 마이크를 사용하여 혼합된 신호로부터 반향신호와 근단화자신호를 분리하고 이를 반향제거기의 계수 업데이트에 반영하여 동시통화 중에도 반향제거기가 정상 동작하는 방법에 관한 것이다. 폐쇄된 반향환경에서 암묵신호분리의 혼합모델은 다채널이기 때문에 분리계수를 직접 계산하지 않고 반향제거기의 계수를 복사하여 그대로 사용할 수 있다. 모의실험 결과, 제안된 알고리즘은 비교적 간단한 과정을 통해 두 음원을 분리하고 동시통화 시에도 안정적으로 동작함을 보였다. 그리고 일반 LMS 알고리즘에 비해 ERLE가 평균 20dB 향상되는 것으로 나타났다.

참고문헌

[1] S. Minami, T. Kawasaki, "A Double Talk Detection Method for an Echo Canceller," *ICC*'85, pp.1492-1497, 1985.

[2] Hua Ye, Bo-Xiu Wu, "A New Double-Talk Detection Algorithm Based on the Orthogonality Theorem," *IEEE Trans. on Comm*., vol.39, no.11, pp.1542-1545, Nov. 1991.

[3] Wenbin Hsu, Frank Chui, David A. Hodges, "An Acoustic Echo Canceler," *IEEE J. of Solid-state Circuits*, vol.24, no.6, pp.1639-1646, Dec. 1989.

[4] P. Comon, "Independent component analysis, A new concept," *Signal Processing*, vol.36, pp.287-314, 1994.

[5] C. Jutten and J. Herault, "Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture," *Signal Processing*, vol.24, no.1, pp.1-10, Jul. 1991.

[6] Bell A. J., Sejnowski T. J., "An Information Maximization Approach to Blind Separation and Blind Deconvolution," *Neural Computation*, vol.7, pp.1129-1159, 1995.

[7] Kari Torkkola, "Blind Separation of Convolved Sources Based on Information Maximization," *Neural Networks for Signal Processing*, *IEEE Signal Processing Society Workshop*, 1996.

[8] D.W.E. Schobben, P. "Sommen, A frequency domain blind signal separation method based on decorrelation," *IEEE Trans. Signal Process.*, vol.50, no.8, pp.1855-1865, 2002.

[9] M. Kawamoto, K. Matsuoka, et al., "A method of blind separation for convolved non-stationary signals," *Neuro-computing*, vol.22, pp.157-171, 1998.

[10] L. Parra, C. Spence, "Convolutive blind separation of non-stationary sources," *IEEE Trans. Speech Audio Process.*, vol.8, no.3, pp.320-327, 2000.

[11] E. Weinstein, M. Feder, A.V. Oppenheim, "Multi-channel signal separation by decorrelation," *IEEE Trans. Speech Audio Process.*, vol.1, no.4, pp.405-413, 1993.

[12] D. Yellin, E. Weinstein, "Multichannel signal separation: methods and analysis," *IEEE Trans. Signal Process.*, vol.44, no.1, pp.106-118, 1996.

저자소개

이행우(Haeng-Woo Lee)

2001년 전북대학교 공학박사
2001년~ 남서울대학교 부교수

※관심분야 : VLSI 설계, 음향 신호처리, 적응 알고리즘

윤현민(Hyun-Min Yun)

2006년 인하대학교 공학박사
1991년~ 재능대학교 교수

※관심분야 : 반도체 공정, VLSI 개발, 디지털통신