스테레오 원거리회의 시스템을 위한 음향 반향 제거기

Acoustic Echo Canceller for Stereophonic Tele-conferencing System

정 규 화*, 이 원 철**, 윤 대 희*, 차 일 환*
(Kyu-Hwa Jeong*, Won-Cheol Lee**, Dae-Hee Youn*, Il-Hwan Cha*)

※본 연구는 한국과학재단의 연구비 지원에 의한 것입니다.

요 약

본 논문은 스테레오 원거리 회의 시스템에서 원단 화자의 움직임이나 전송실 환경이 변하는 순간에 생기는 반향 제거 성능의 저하를 막기 위한 새로운 스테레오 음향 반향 제거기를 제안한다. 스테레오 음향 반향 제거기의 경우 반향 제거를 위한 적응 필터들의 계수들의 해가 유일하지 않게 되고 그들의 최적값으로 수렴하지 않게 된다. 따라서 원단의 전송실 환경 변화는 ERLE(Echo Return Loss Enhancement)값의 급격한 저하를 초래하게 된다. 또한 단일 채널과 달리 적응 필터 계수의 증가에 따라 계산량이 증가하고 수렴 속도가 느리게 되어 실용화에 어려움이 있다. 이와 같은 문제점을 해결하기 위하여 본 논문에서는 의사 스테레오 신호(pseudo stereophonic signal)를 만드는 적응 필터로 구성된 전처리기를 제안하여 원단 환경의 변화 순간에 성능을 개선하였고 반향 제거에 필요한 전체 계산량을 감소시켰다.

ABSTRACT

This paper proposes a new stereophonic acoustic echo canceller to prevent the performance degradation at the instant of far-end talker's movement or change in transmission room environment in stereophonic teleconferecing system. In stereophonic acoustic echo canceller, the filter coefficients of the adaptive filters for echo cancellation do not have unique solutions and not converge to their optimum values. Therefore, the change of the far-end transmission room environment leads to degradation of ERLE(Echo Return Loss Enhancement). Moreover, their computational complexity increases as the number of the filter coefficeint increases, and their adaptive filters converge very slowly. So the real-time implementation is very difficult. To overcome these problems we propose a pre-processor consisting of an adaptive filter for making pseudo stereophonic signal, and it results in improved performance at the instant of environment change at far-end side and reduction of the total complexity.

I. 서 론

정보화 사회의 발전과 더불어 빠르게 진행되고 있는 통신 수단의 발전은 기존의 유선 통신 개념으로부터 무선 통신 개념으로, 또한 편리한 통화 환경을 위해 핸드프리(hands-free)방식으로 발전해 나가고 있다. 이러한 핸드프리 방식이 응용될 수 있는 분야로는 스피커폰 시스템, 핸드프리 전화기, 원거리 회의 시스템(tele-conferencing system) 등이 있다[1]. 이러한 핸드프리 통신 방식에서는 필연적으로 음향 반향(acoustic echo)이 생기게 되는데, 이는 수신된 상대방의 신호(이하 원단화자 신호)가 스피커와 통화자가 위치한 실내공간을 거쳐 다시 마이크로폰으로 입력 되는 경우를 말하며 이로 인해 원단화자는 자기가 한 말을 통화 왕복 시간 뒤에 듣게 되어 통화에 지장을 주게 된다[2]. 이러한 문제점을 해결하기 위해서 적응 필터(adaptive filter)를 사용하여 반향 경로(echo path)를 매시간 추정, 새로 생성된 반향 성분만을 제거하여 통화의 부자연성 등을 줄여주는 방법이 가장 많이 연구되고 있다[2][3]. 현재 대부분의 원거리회의 시스템은 하나의 전-이중(full-duplex) 오디오 채널만을 이용하여 통신을 하기 때문에 회의 참석자들에게 공간감을 느끼게 해줄 수가 없다. 특히 화상 회의 시스템인 경우 화면으로 화자의 움직임 등을 전달 할 수 있으나 실제로 듣는 신호는 전혀 변화가 없는 신호가 들리기 때문에 통화의 부자연성이 발생하게

*연세대학교 전자공학과
**숭실대학교 정보통신공학과
접수일자 : 1997년 1월 18일

된다[4]. 따라서 회의 참석자들이 마치 한 공간에서 회의를 하는 듯한 느낌을 주기 위해서는 적어도 2개의 오디오 채널을 사용하는 스테레오 원거리회의 시스템이 필요하다[4][5]. 스테레오 원거리 회의 시스템에서도 마찬가지로 음향 반향이 생길 수 가 있는데 이런 경우 2개의 마이크로폰과 2개의 스피커를 사용하므로 4개의 반향 경로가 생기게 되고 이를 제거하기 위해서는 4개의 반향 제거기가 필요하게 된다. 따라서 단일 채널 시스템에 비해 반향 제거를 위해 더 많은 계산량이 필요하게 된다. 또한 각 채널의 입력 신호들간에 높은 상호상관(cross-correlation)에 의해 단일 채널인 경우에서 생기지 않는 여러 가지 문제점들이 발생하게 된다. 지금까지 스테레오 음향 반향 제거기에 관한 연구는 그 필요성에 비해 아직까지는 문제제기 단계에 머물고 있는 수준이다[5]. 본 논문에서는 스테레오 음향 반향 제거기의 문제점을 해결하기 위해 적응필터를 사용한 전처리기를 통해 의사 스테레오(pseudo stereophonic) 신호를 만든 후 이를 이용하는 새로운 구조를 제안하였다.

본 논문의 구성은 2장에서 스테레오 음향 반향제거기가 갖는 문제점에 대해 설명하고, 3장에서는 이를 해결하기 위해 제안된 구조에 대한 설명을 한다. 그리고 4장에서 모의 실험 결과를 비교하고 5장에서 결론을 맺는다.

Ⅱ. 스테레오 음향 반향제거기의 문제점

스테레오 원거리회의 시스템에서 반향을 제거하기위한 기본 구조는 그림 1과 같다[5]. 이는 적응 필터를 사용하여 반향을 제거하는 기본 원리를 2개의 채널에 대해 적용한 것이다. 그림 1에서 G_1과 G_2는 원단에 위치한 전송실(transmission room)에서의 화자와 각 마이크로폰사이의 전달함수를 나타내고, 근단 화자가 위치한 수신실(receiving room)에서의 H_1과 H_2는 각 스피커로부터 하나의 마이크로폰간의 전달함수인 반향 경로를 나타낸다. 각 채널이 대칭성을 유지하므로 본 논문에서는 하나의 마이크로폰으로 입력되는 반향 성분들에 대해서만 살펴보도록 한다. 원단화자의 신호가 각 채널을 통해 상대방에게 전송될 때 수신룸의 스피커 출력 신호가 음향 경로 H_1과 H_2

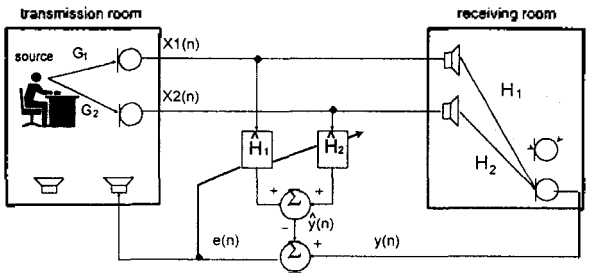


그림 1. 스테레오 음향 반향 제거기의 기본구조

를 통해 다시 마이크로폰으로 유입되는 경우 반향 신호 y(n)이 생기게 된다. 이 때 적응 필터 $\hat{H}_1$과 $\hat{H}_2$를 사용하여 반향 경로 H_1과 H_2를 추정함으로써 반향 신호를 추정할 수 있다. 이렇게 추정한 값을 반향 신호로부터 빼주어서 반향 성분을 제거할 수 있다.

일반적으로 반향 신호의 제거를 위한 적응필터는 계산이 용이한 LMS(Least-Mean-Square) 알고리듬[6]을 많이 사용한다. 스테레오 반향제거기에서의 반향 신호는 다음과 같다.

$$y(n) = h_1(n) * x_1(n) + h_2(n) * x_2(n) \tag{1}$$

여기서 $h_1(n)$과 $h_2(n)$는 반향 경로인 스피커와 마이크로폰 사이의 임펄스 응답을 나타내고 $x_1(n)$과 $x_2(n)$은 각 채널을 통해 수신실에 전달되는 신호이다. 적응필터를 사용해 추정한 반향신호는 다음과 같다.

$$\hat{y}(n) = \hat{h}_1^T(n) X_1(n) + \hat{h}_2^T(n) X_2(n) \tag{2}$$

여기서 $\hat{h}_1(n)$과 $\hat{h}_2(n)$는 탭수가 L인 적응 필터의 계수를 나타내는 벡터이고 $X_1(n)$과 $X_2(n)$는 적응 필터의 기준 입력 신호 벡터이며 T는 전치행렬을 나타낸다. 기준 입력 신호와 필터 계수로 주어지는 벡터 $X(n)$과 $\hat{h}(n)$을 아래와 같이 정의하면 반향 제거기의 필터계수 추정을 위한 LMS알고리듬은 다음과 같다.

$$X(n) \equiv [X_1^T(n)\ X_2^T(n)]^T,$$

$$\hat{h}(n) \equiv [\hat{h}_1^T(n)\ \hat{h}_2^T(n)]^T$$

$$\hat{h}(n+1) \equiv \hat{h}(n) + \mu e(n) X(n)$$

$$e(n) = y(n) - \hat{y}(n) \tag{3}$$

따라서, 적당한 수렴상수 μ에 의해 $e(n)$의 평균자승이 최소가 되도록 적응 필터의 계수를 갱신한다. 적응 필터가 완전히 수렴한 후 적응 필터 계수의 최적값은 이상적으로 각각의 반향 경로의 임펄스 응답과 일치하여야 한다. 그러나 평균 자승 오차를 최소화 하는각 적응 필터에 대한 최적 계수 벡터들을 구하면 다음과 같이 주어진다.

$$\hat{h}_{1opt}(n) = R_1^{-1}(n) P_1(n)$$
$$= h_1 + R_1^{-1}(n) R_{12}(n) h_2 \tag{4-1}$$

$$\hat{h}_{2opt}(n) = R_2^{-1}(n) P_2(n)$$
$$= h_2 + R_2^{-1}(n) R_{21}(n) h_1 \tag{4-2}$$

여기서 R은 적응 필터 입력 신호들의 상관 행렬(correlation matrix)이고 P는 입력 신호와 반향 신호간의 상호

상관 벡터를 나타내며 다음과 같다.

$$R_1(n) = E\{X_1(n)\,X_1^T(n)\},\quad R_2(n) = E\{X_2(n)\,X_2^T(n)\} \quad (4\text{-}3)$$

$$R_{12}(n) = E\{X_1(n)\,X_2^T(n)\},\quad R_{21}(n) = E\{X_2(n)\,X_1^T(n)\} \quad (4\text{-}4)$$

$$P_1(n) = E\{X_1(n)\,y(n)\},\quad P_2(n) = E\{X_2(n)\,y(n)\} \quad (4\text{-}5)$$

위에서 볼 수 있듯이 두 채널 신호간의 상호상관(cross-correlation)에 의해 적응 필터의 최적 계수 벡터들은 $\hat{h}_{1opt} \neq h_1$, $\hat{h}_{2opt} \neq h_2$인 결과가 된다. 즉, 각 적응 필터의 계수들과 각 반향 경로를 나타내는 임펄스 응답들과의 일치가 보장되지 않는다[4][5]. 이로 인해 스테레오 음향 반향 제거기의 문제점이 발생하게 되는데 그것은 각 적응 필터의 해가 유일하지 않다는 것이다[5][7]. 즉 각 적응 필터의 최적 계수 벡터가 경우에 따라서 바뀔 수 있음을 의미한다. 이를 위해 이것들의 Z-변환 함수들을 살펴 보면 다음과 같다.

$$X_1(z) = G_1(z)\,S(z),\qquad X_2(z) = G_2(z)\,S(z) \quad (5)$$

여기서 $G_1(z)$과 $G_2(z)$는 원단화자와 마이크로폰 사이의 전달함수들이고, $S(z)$는 원단화자 원신호의 Z-변환이다. 적응 필터가 수렴한 후 오차 신호 $e(n)$의 Z-변환은

$$E(z) = \{H_1(z)G_1(z) + H_2(z)G_2(z) - \hat{H}_1(z)G_1(z) - \hat{H}_2(z)G_2(z)\}\,S(z) \quad (6)$$

가 된다. 식 (6)에서 오차 신호가 완전히 제거되기 위한 조건은 다음과 같다.

$$[H_1(z) - \hat{H}_1(z)]\,G_1(z) + [H_2(z) - \hat{H}_2(Z)]\,G_2(z) = 0 \quad (7)$$

식 (7)을 만족시키는 이상적인 적응 필터의 해가 $\hat{H}_1(z) = H_1(z)$, $\hat{H}_2(z) = H_2(z)$이면 별 문제없이 반향 성분을 제거할 수 있다. 그러나 위에서 살펴본 것과 같이 각 적응 필터의 해는 각 반향 경로와 일치하지 않기 때문에 식(7)을 만족시키는 적응 필터의 해는 여러 가지의 조합으로 나타날 수 있다. 이로 인해 $H_1(z)$와 $H_2(z)$가 변하지 않더라도 $G_1(z)$이나 $G_2(z)$가 바뀌게 되면 적응 필터의 해인 $\hat{H}_1(z)$와 $\hat{H}_2(z)$의 값이 바뀌게 된다. 일반적으로 반향 경로가 변하게 되면 잔여 반향 신호가 커지게 되어 ERLE(Echo Return Loss Enhancement)가 떨어지게 된다. 이런 경우를 위해 수렴이 빠른 적응 필터가 필요하게 된다. 스테레오 음향 반향 제거기에서는 반향 경로가 변하지 않더라도 원단화자의 움직임 등으로 인해 원단화자와 마이크로폰 사이의 전달함수가 바뀌게 될 경우 마치 반향 경로가 변한 것과 같은 영향을 미칠뿐더러 두 채널 신호간의 상호상관이 매우 크기 때문에 적응 필터의 수렴 속도 또한 느리게 되는 문제가 발생한다. 그리고 단일 채널에 비해 4배 이상의 계산량을 필요로 한다. 따라서 본 논문에서는 이러한 문제를 해결하는 구조를 제안하였다.

Ⅲ. 제안된 스테레오 음향 반향제거기

본 논문에서는 전송룸에서의 환경변화에 둔감한 스테레오 음향 반향제거기를 위해 스테레오 신호에 별 영향을 주지 않는 의사 스테레오 신호(pseudo stereophonic signal)를 만드는 전처리단을 제안하였고 이를 이용하여 계산량이 줄어드는 구조를 제안하였다.

3.1 전처리기의 구조

그림 2는 전송룸의 환경 변화로 인한 성능 저하를 줄이기 위해 제안된 반향 제거기의 구조이다. 이 구조는 기존의 구조에다 전처리기를 첨가한 것이다. 즉, 원신호를 전송하는 것이 아니고 전처리기를 통해 재생된 신호를 수신룸의 스피커를 통해 출력한다. 기본 개념은 전송룸의 환경이 갑자기 변하게 될 경우 각 적응 필터의 해의 급격한 변화를 막아줌으로써 전송룸의 환경 변화에 둔감하도록 해주는 것이다. 전처리기에서는 원래의 신호 $x_1(n)$과 $x_2(n)$를 왜곡없이 재생하기 위해 적응 필터를 사용하였다. $x_c(n) = x_1(n) + x_2(n)$를 기준 신호(reference signal)로 하고 $x_1(n)$을 추정하는 적응 필터를 $F_1(n)$이라고 하고 마찬가지로 $x_c(n)$를 사용하여 $x_2(n)$를 만들기 위한 적응필터를 $F_2(n)$라고 하면 각 필터가 수렴한 후 필터계수의 Z-변환을 보면 다음과 같다.

$$F_1(z) = \frac{G_1(z)}{G_1(z) + G_2(z)} \quad (8)$$

$$F_2(z) = \frac{G_2(z)}{G_1(z) + G_2(z)} \quad (9)$$

그런데 $F_1(z)$과 $F_2(z)$ 사이에는 $F_2(z) = 1 - F_1(z)$의 관계가 성립되므로 그림 3과 같이 하나의 적응 필터 $F(z)$만으로 두 신호를 재생해 낼 수 있다. 적응 필터 $F(z)$의 계수 적응 식은 다음과 같다.

$$F(n+1) = F(n) + \eta e_p(n) X_c(n) \quad (10)$$

여기서 $X_c(n) = [x_c(n)\,x_c(n-2)\,\cdots\,x_c(n-M+1)]^T$이고

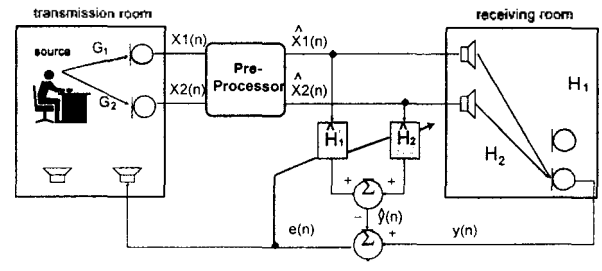


그림 2. 제안한 스테레오 음향 반향 제거기의 구조

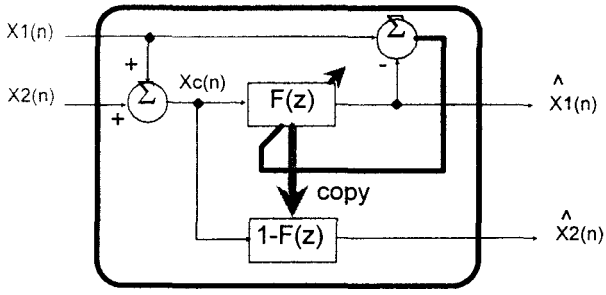


그림 3. 제안한 전처리기의 구조

$e_p(n)=x_1(n)-F(n)^T X_c(n)$, η은 전처리단 적응 필터의 수렴상수이며, M은 필터 탭수이다. 이때 재생한 스테레오 신호들은 다음과 같다.

$$\hat{x}_1(n)=F(n)^T X_c(n), \qquad \hat{x}_2(n)=(d-F(n))^T X_c(n) \tag{11}$$

여기서 $d=[1\ 0\ 0\ \cdots\ 0]^T$인 M×1 벡터를 나타낸다.

위의 식들에 의해 두 신호의 합으로부터 다시 스테레오 신호를 재생함으로써 스테레오 신호에는 별 영향을 주지 않는다. 이렇게 해서 재생한 신호 $\hat{x}_1(n)$과 $\hat{x}_2(n)$를 그림 2처럼 상대방의 스피커로 내보내 주고 동시에 반향 제거기의 기준 신호로 사용하여 LMS 알고리듬을 적용한 적응필터 $\hat{H}_1$과 $\hat{H}_2$를 이용해 반향을 제거한다. 위의 결과를 이용하여 반향 제거를 위한 적응 필터 계수 갱신식을 살펴보면 다음과 같다.

$$\hat{h}(n+1)=\hat{h}(n)+\mu e(n)\hat{X}(n) \tag{12}$$

$$\hat{e}(n)=y(n)-\hat{h}_1^T(n)\hat{X}_1(n)-\hat{h}_2^T(n)\hat{X}_2(n) \tag{13}$$

여기서 $\hat{X}(n)\equiv[\hat{X}_1^T(n)\ \hat{X}_2^T(n)]^T$,
$\hat{h}(n)\equiv[\hat{h}_1^T(n)\ \hat{h}_2^T(n)]^T$,
$\hat{X}_1(n)=[\hat{x}_1(n)\ \hat{x}_1(n-1)\ \cdots\ \hat{x}_1(n-L+1)]^T$,
$\hat{X}_2(n)=[\hat{x}_2(n)\ \hat{x}_2(n-1)\ \cdots\ \hat{x}_2(n-L+1)]^T$,

L은 반향 제거 적응필터의 탭수이다.

제안한 구조에서 잔여오차 신호 $e(n)$의 분산(variance) σ_e^2는

$$\sigma_e^2=\int_{-\infty}^{\infty} S_e(e^{jw})\,dw$$
$$=\int_{-\infty}^{\infty} |\Delta\tilde{H}(e^{jw})F(e^{jw})+\tilde{H}_2(e^{jw})|^2 S_{xc}(e^{jw})\,dw \tag{14}$$

이 된다.

여기서 $S_e(e^{jw})$는 오차 신호의 전력 스펙트럼 밀도 함수이고 $S_{xc}(e^{jw})$는 $x_c(n)$의 전력 스펙트럼 밀도 함수를 나타낸다.

또한 $\Delta\tilde{H}(e^{jw})=\tilde{H}_1(e^{jw})-\tilde{H}_2(e^{jw})$,
$\tilde{H}_1(e^{jw})=H_1(e^{jw})-\hat{H}_1(e^{jw})$,
$\tilde{H}_2(e^{jw})=H_2(e^{jw})-\hat{H}_2(e^{jw})$ 이다.

반향 신호의 분산은

$$\sigma_y^2=\int_{-\infty}^{\infty} S_y(e^{jw})\,dw$$
$$=\int_{-\infty}^{\infty} |\Delta H(e^{jw})F(e^{jw})+H_2(e^{jw})|^2 S_{xc}(e^{jw})\,dw \tag{15}$$

이고, 여기서 $\Delta H(e^{jw})=H_1(e^{jw})-H_2(e^{jw})$이다. 반향 제거의 성능을 평가하는 ERLE(Echo Return Loss Enhancement)는 다음과 같이 나타낸다.

$$\mathrm{ERLE}=10\log_{10}\frac{E[y^2(n)]}{E[e^2(n)]}\ dB \tag{16}$$

전송룸이 변하는 순간의 순간 ERLE를 $ERLE_{inst}$라고 하면 식 (14)와 식 (15)에 의해 $ERLE_{inst}$는 다음과 같다.

$$ERLE_{inst}\equiv 10\log_{10}\frac{\sigma_y^2}{\sigma_e^2}$$
$$\equiv 10\log_{10}\frac{\sigma_y^2}{\left[\left\{\int_{-\infty}^{\infty}|\Delta\tilde{H}(e^{jw})F(e^{jw})+\tilde{H}_2(e^{jw})|^2 S_{xc}(e^{jw})\,dw\right\}^{1/2}\right]^2} \tag{17}$$

여기서 전송룸의 환경 변화 등으로 인해 $G_1(z)$와 $G_2(z)$가 변하는 순간의 전처리기 적응 필터를 $\hat{F}(z)$라고 하면 $\hat{F}(z)=F(z)+\Delta F(z)$로 나타낼 수 있다. 식(17)을 Minkowski's 부등식[9]을 이용해 변하는 순간의 ERLE 값을 정리해 보면 다음과 같다.

$$ERLE_{inst}\geq 10\log_{10}\frac{\sigma_y^2}{\sigma_y+A(\hat{H}_1,\hat{H}_2,F,S_{xc})+B(\hat{H}_1,\hat{H}_2,\Delta F,S_{xc})]^2} \tag{18}$$

여기서 $A=\left\{\int_{-\infty}^{\infty}|\Delta\hat{H}(e^{jw})F(e^{jw})+\hat{H}_2(e^{jw})|^2 S_{xc}(e^{jw})\,dw\right\}^{1/2}$

$B=\left\{\int_{-\infty}^{\infty}|\Delta\hat{H}(e^{jw})|^2\,|\Delta F(e^{jw})|^2 S_{xc}(e^{jw})\,dw\right\}^{1/2}$

이다. 식 (18)의 우변을 $ERLE_{\min}$라 할 때, $G_1(z)$와 $G_2(z)$가 변하는 순간의 제안된 방법에서는 전처리기 적응 필터의 수렴 상수 η를 작게하면 $\Delta F(z)\approx 0$이 되고 기존의 방법에서는 $|\Delta F(z)|=\left|\frac{\Delta G_1(z)}{\Delta G_1(z)+\Delta G_2(z)}\right|>0$이 된다. 따라서 제안한 방법에서의 ERLE값의 최소한계 $ERLE_{\min}^{(p)}$와 기존의 방법에서의 최소한계 $ERLE_{\min}^{(c)}$는 다음의 관계를 갖는다.

$$ERLE_{\min}^{(p)}\geq ERLE_{\min}^{(c)} \tag{19}$$

식(19)로부터 제안한 방법에서 전송룸의 환경 변화에 영향을 덜 받는 것을 알 수 있다.

3.2 계산량을 감소시킨 스테레오 음향 반향 제거기

제안된 구조에서 추정된 반향신호 $\hat{y}(n)$은

$$\hat{y}(n) = \hat{\boldsymbol{h}}_1^T(n)\,\hat{\boldsymbol{X}}_1(n) + \hat{\boldsymbol{h}}_2^T(n)\,\hat{\boldsymbol{X}}_2(n) = [\hat{\boldsymbol{h}}_1^T(n)\ \hat{\boldsymbol{h}}_2^T(n)]\,[\hat{\boldsymbol{X}}_1^T(n)\ \hat{\boldsymbol{X}}_2^T(n)]^T \tag{20}$$

으로 나타낼 수 있다. 그런데 $\hat{\boldsymbol{X}}_1(n)$과 $\hat{\boldsymbol{X}}_2(n)$는

$$\hat{\boldsymbol{X}}_1(n) = \boldsymbol{F}_M(n)\,\tilde{\boldsymbol{X}}_c(n), \qquad \hat{\boldsymbol{X}}_2(n) = (\boldsymbol{D} - \boldsymbol{F}_M(n))\,\tilde{\boldsymbol{X}}_c(n) \tag{21}$$

가 되고,

여기서 $\tilde{\boldsymbol{X}}c(n) = [x_c(n)\ x_c(n-1)\ \cdots\ x_c(n-L-M+1)]^T$,

$$\boldsymbol{F}_M(n) = \begin{bmatrix} f_1(n) & f_2(n) & \cdots & \cdots & f_M(n) & 0 & \cdots & \cdots & 0 \\ 0 & f_1(n) & f_2(n) & \cdots & \cdots & f_M(n) & 0 & \cdots & 0 \\ 0 & 0 & f_1(n) & f_2(n) & \cdots & \cdots & f_M(n) & \cdots & 0 \\ \vdots & & & & & & & & \\ 0 & \cdots & \cdots & \cdots & 0 & f_1(n) & \cdots & \cdots & f_M(n) \end{bmatrix}$$

$$\boldsymbol{D} = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & \cdots & 0 & \cdots & 0 \\ \vdots & & & & & & & & \\ 0 & \cdots & \cdots & \cdots & 0 & 1 & \cdots & \cdots & 0 \end{bmatrix}$$

를 나타낸다. $\tilde{\boldsymbol{X}}_c(n)$은 (L+M)×1 벡터를 나타내고, $\boldsymbol{F}_M(n)$은 전처리기 적응 필터의 계수들로 구성된 L×(L+M) 행렬이다. 또한 $\boldsymbol{D}$는 L×(L+M) 행렬이다. 여기서 L은 반향 제거 적응 필터의 탭수이고 M은 전처리기 적응필터의 탭수이다. 식 (21)을 사용해 식 (20)를 표시하면,

$$\hat{y}(n) = [\hat{\boldsymbol{h}}_1^T(n)\ \hat{\boldsymbol{h}}_2^T(n)] \begin{bmatrix} \boldsymbol{F}_M(n) \\ \boldsymbol{D} - \boldsymbol{F}_M(n) \end{bmatrix} \tilde{\boldsymbol{X}}_c(n) = \boldsymbol{M}^T(n) \tag{22}$$

가 되며, (L+M)×1 벡터 $\boldsymbol{M}(n)$은

$$\boldsymbol{M}^T(n) \equiv [\hat{\boldsymbol{h}}_1^T(n)\ \hat{\boldsymbol{h}}_2^T(n)] \begin{bmatrix} \boldsymbol{F}_M(n) \\ \boldsymbol{D} - \boldsymbol{F}_M(n) \end{bmatrix} \tag{23}$$

이다. 따라서 오차 신호는

$$e(n) = y(n) - \boldsymbol{M}^T(n)\,\tilde{\boldsymbol{X}}c(n) \tag{24}$$

이 된다. 식 (24)로부터 전처리기을 이용한 제한된 방법은 그림 4와 같이 $x_c(n)$를 기준 신호(reference signal)로 사용하고 $y(n)$를 desired 신호로 사용하는 적응 필터 $\boldsymbol{M}(n)$을 사용하는 구조로 간단히 할 수 있다. 이렇게 함으로서 계산량을 줄일 수가 있다. 기존의 방법인 경우 반향 제거를 위한 적응 필터의 탭수를 L이라고 하면 전체 계산량은 4*L이 필요하지만 제안한 방법에서는 2*(L+M)의 계산량이 필요하게 된다. 여기서 M은 전처리단 적응 필터의 탭수이며 M<L이다. 전체 곱셈량을 비교한 것이 〈표 1〉에 나타내었다.

Ⅳ. 모의 실험

본 장에서는 유색잡음(colored noise)과 실제 음성신호를 사용하여 전송룸의 전달함수가 변할 때 기존의 방법과 제안된 방법과의 성능을 비교하였다. 전송룸의 임펄스 응답과 반향 경로를 모두 256개의 데이터를 가진 임펄스 응답을 사용하였고, 전처리기 적응필터의 탭수는 32개, 반향 제거를 위한 적응 필터의 탭수는 256개를 사용하였다. 모의 실험에 사용한 적응 필터로는 NLMS(normalized LMS) 알고리듬을 사용하였다. 전송룸의 환경이 바뀌는 상태를 나타내기 위해 50000 샘플 중 25000 샘플에서 G_1, G_2를 다른 전달함수로 바꿔 주었다. 그림 5는 모의 실험에 사용된 전송룸이 변한 상태에서의 유색 잡음 입력 신호들이다. 실험에 사용한 유색 잡음 신호는 일반적인 음성의 스펙트럼을 가지도록 AR(Auto Regressive)모델링한 신호를 사용하였으며 음성 신호로는 8kHz로 샘플링한 실제음성 신호를 사용하였다. 그림 6은 실험에 사용된 반향 경로인 임펄스 응답을 나타낸다. 그림 7는 유색 잡음을 원신호(source signal)로 사용했을 때의 ERLE 곡선을 비교한 것으로, 20번의 앙상블 평균(ensemble average)을 구한것이다. 그림에서 알수 있듯이 G_1, G_2가 변하는 순간에 제안한 방법의 성능이 우수함을 알 수 있다. 그림 8에서는 실제 음성 신호를 사용했을 경우 잔여 반향 신호를 비교하였다. 기존의 방법인 그림 8(a)는 G_1, G_2가 변하는 순간 잔여 반향이 갑자기 증가하는 반면 제안된 방법인 그림 8(b)에서는 이를 해결하였다. 마찬가지로 계산량을 줄인 구조에서도 그림 8(c)와 같이 성능이 우수함을 볼 수 있다.

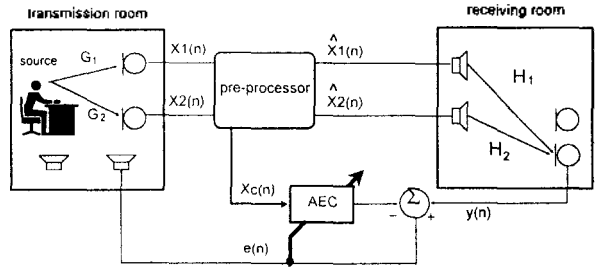


그림 4. 계산량을 감소시킨 스테레오 음향 반향 제거기

표 1. 전체 곱셈량의 비교(반향 제거 적응필터 탭수 L = 256, 전처리 적응필터 탭수 M = 32)

기존의 방법	제안한 방법
각 LMS 당 곱셈량 : 2L 전체 곱셈량 : 4*(2L) = 2048	전처리단-LMS 곱셈량 : 2M 필터링 : M 반향 제거기-각 LMS곱셈량 : 2(L +M) 전체 곱셈량 : 3M +4(L +M) = 672

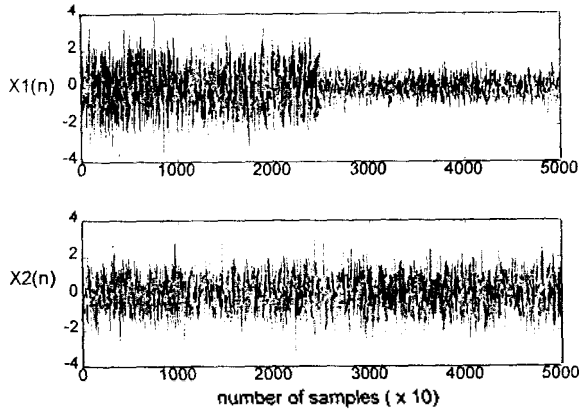


그림 5. 모의 실험에 사용한 유색 잡음 신호

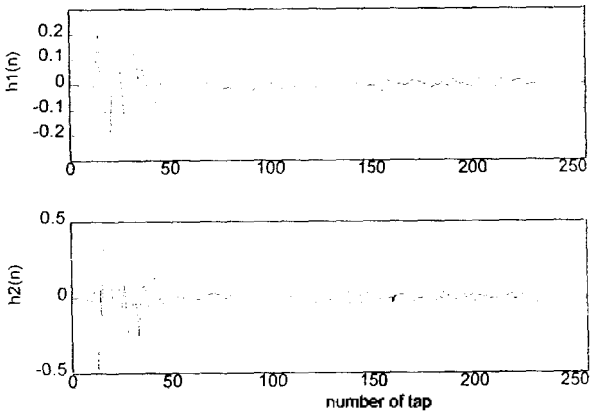
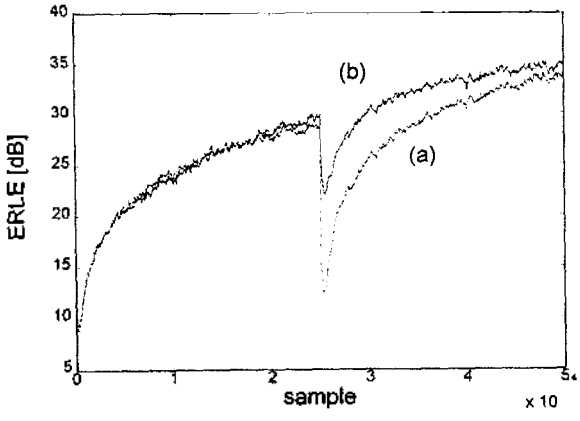
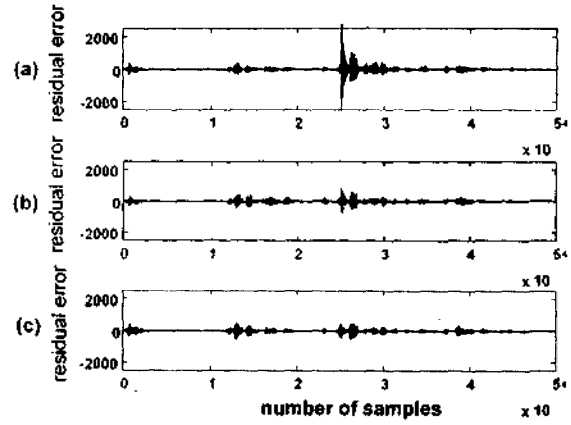


그림 6. 모의 실험에 사용한 반향 경로



(a) 기존의 방법 (b) 제안한 방법

그림 7. 유색 잡음을 사용했을 경우 ERLE의 비교



(a)기존의 방법 (b)제안한 방법 (c)계산량을 줄인 방법

그림 8. 음성 신호를 사용했을 경우 잔여 반향 신호의 비교

V. 결 론

현장감 있는 원거리 회의를 위해서는 스테레오 원거리 회의 시스템이 필요하게 되고 이때 생기는 반향을 제거하기 위해서는 스테레오 음향 반향 제거기가 필요하게 된다. 그러나 스테레오 음향 반향 제거기는 수신룸의 환경 변화뿐만 아니라 전송룸의 환경 변화에 민감한 영향을 받게 된다. 본 논문에서는 스테레오 원거리회의 시스템에서 발생하는 반향 제거의 문제점들 중 전송룸의 환경이 변할 때 반향 제거기의 성능이 저하되는 문제를 해결하기 위해 전처리기을 사용하는 새로운 구조를 제안하였다. 또한 이를 바탕으로 계산량을 줄일 수 있는 방법도 제안하였다. 제안한 방법에서는 현장감에 지장이 없는 의사 스테레오 신호를 사용하였다. 그리고 모의 실험을 통해 성능의 우수함을 보였다.

추후 연구과제로는 의사 스테레오 신호를 사용하지 않고 전송룸의 환경 변화에 영향 받지 않는 알고리듬에 대한 연구와 실시간 구현을 위해 수렴을 향상시킨 적응 필터 알고리듬[8]에 대한 연구가 필요하며 현재 이들에 대한 연구를 진행 중에 있다.

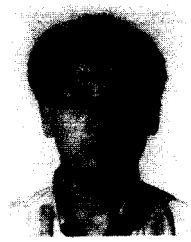
참 고 문 헌

1. S. H. Jensen, "Acoustic Echo Canceller for Hands-Free Mobile Radiotelephony," *Proc. EUSIPCO*, Belgium, pp. 1629-1632, 1992
2. K. Murano, S. Unagami and F. Amano, "Echo Cancellation and Applications," *IEEE Comm. Magzine*, pp. 49-55, Jan. 1990.
3. 유재하, 조성호, 윤대희, "음향 반향 제거기," 대한 전자공학회 논문지, 1995년 5월.
4. A. Hirano and A. Sugiyama, "A Compact Multichannel Echo Canceller with A Single Adaptive Filter per Channel," *Proc. Int. Conf. Acoust. Speech, Signal Processing*, pp. 1922-

1925, 1992.

5. M. M. Sondhi, D. R. Morgan and J. L. Hall, "Stereophonic Acoustic Echo Cancellation-An Overview of the Fundamental Problem," *IEEE Signal Prosessing Letters*, vol. 2, no. 8, pp. 148-151, Aug. 1995
6. S. Haykin, *Adaptive Filter Theory*, Prentice Hall. 1991.
7. J. Bebesty, F. Amand, A. Gilloire and Y. Grenier, "Adaptive Filtering Algorithms for Stereophonic Acoustic Echo Cancelation," *Proc. Int. Conf. Acoust. Speech. Signal Processing*. pp. 3099-3102. 1995
8. F. Amand, J. Benesty, A. Gilloire and Y. Grenier, "A Fast Two-Channel Projection Algorithm," *Proc. Int. Conf. Acoust. Speech. Signal Processing*, pp. 949-952, 1996
9. D. Zwillinger, *Standard Mathematical Tables and Formulae*, CRC Press.

▲정 규 화(Kyu-Hwa Jeong) 1972년 6월 12일생



1995년 2월 : 중앙대학교 공과대학 전기공학과 졸업(공학사)

1995년 9월~현재 : 연세대학교 대학원 전자공학과 석사과정

※주관심분야 : 디지탈 신호처리, 적응디지탈 필터

▲이 원 철(Won Cheol Lee)

한국음향학회지 15권 1E호 참조

▲윤 대 희(Dae Hee Youn)

한국음향학회지 14권 4호 참조

▲차 일 환(Il Hwan Cha)

한국음향학회지 12권 6호 참조

▲노 용 래(Yong-Rae Roh)

한국음향학회지 15권 5호 참조