

# 가변하는 채널 신호의 차성분을 이용한 스테레오 음향 반향 제거기

정일규\*, 김현태\*\*, 박장식\*\*\*, 손경식\*  
\*부산대학교 전자공학과, \*\*부산울산 중소기업청, \*\*\*동의공업대학 영상정보과

## Stereo Acoustic Echo Canceller Using Variable Difference Components of Channel Signals

\*Il-Kyoo Jung, \*\*Hyun-Tae Kim, \*\*\*Jang-Sik Park, \*Kyung-Sik Son  
\* Dept. of Electronics Engineering, Pusan National University.  
\*\* Regional Office of Pusan/Ulsan Small and Medium Business Admin.  
\*\*\* Dept. of Image Processing, Dong-Eui Institute of Technology.

### 요 약

스테레오 음향 반향 제거기는 채널신호간에 상호상관(cross-correlation) 때문에 적응필터 계수 가 수신실의 반향 경로를 정확하게 추정하지 못한다. 본 논문에서는 상호상관을 줄이기 위해서 채널신호간 차의 절대값과 시변 감쇠상수를 이용하는 새로운 전처리 필터를 제안한다. 시물레이션을 통해서 제안하는 전처리 필터가 기존의 방법에 비해서 우수함을 보인다.

### 1. 서론

생생하고 깨끗한 오디오, 비디오 매체를 전송하기 위해 통신시스템에서는 100 ~ 7000Hz 정도의 대역폭을 가져야 한다. 또한 현장감을 느끼기 위해서 최소한 두 채널 이상의 전이중방식(full duplex)으로 원격통신 시스템을 구성해야 한다. 일반적으로 원격통신 시스템에서는 스피커와 마이크 사이의 경로를 따라 발생하는 하울링(howling)이나 반향(echo)이 발생한다. 따라서 원격통신 시스템에는 반드시 음향 반향 제거기를 갖추어야 하고 두 채널의 음향 신호를 전송할 때 발생하는 음향 반향을 제거하는 것을 스테레오 음향 반향 제거라 한다. 단일채널 음향 반향 제거기와 달리 스테레오 음향 반향 제거기는 스테레오 신호가 단일 화자로부터 발생한 신호이므로 두 신호간의 상호상관이 크기 때문에 스테레오 음향 반향 제거기의 성능이 저하 되는 것이다[1][2]. 본 논문에서는 스테레오 신호간의 차성분의 절대값과 시변하는 감쇠상수를 이용하여 스테레오 신호 사이의 상호상관을 줄임으로써

음향 반향 제거 성능을 향상시키는 방법을 제안한다. 시물레이션을 통해서 기존의 스테레오 음향 반향 제거기 중에서 효율적으로 반향을 제거하는 반과정류필터를 이용한 방법[6]과 비교하여 스테레오 음향반향 제거 성능이 우수하다는 것을 확인하였다.

### 2. 스테레오 음향 반향 제거기

스테레오 음향 반향 제거기의 기본구조는 그림 1과 같다. 그림 1에서 전송실은 왼쪽, 수신실은 오른쪽에 위치한다.  $s(k)$ 는 화자의 음성신호이며,  $g_1(k)$ ,  $g_2(k)$ 는 전송실의 화자와 마이크 사이의 경로에 대한 임펄스응답이고,  $x_1(k)$ ,  $x_2(k)$ 는 각 채널신호이다.

$h_1(k)$ ,  $h_2(k)$ 는 수신실의 스피커와 마이크 사이의 실제 반향경로 임펄스응답이다. 스테레오 반향신호  $y(k)$ 는 수신실 마이크 입력신호이며,

$$y(k) = h_1(k) * x_1(k) + h_2(k) * x_2(k) \quad (1)$$

이다.

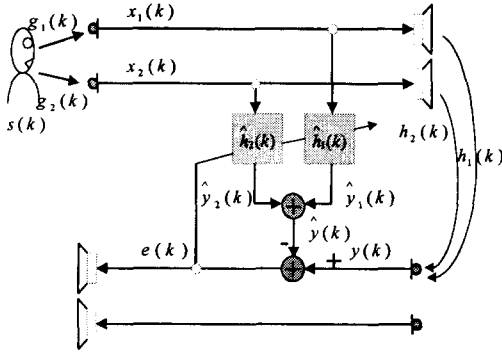


그림 1. 스테레오 음향반향 제거기의 기본구조  
Fig. 1. Basic scheme of stereo echo canceller

기호 \* 는 컨볼루션을 의미하고 k는 시간지수(time index)이다. 오차신호는

$$e(k) = y(k) - \hat{\mathbf{h}}_1^T(k) \mathbf{x}_1(k) - \hat{\mathbf{h}}_2^T(k) \mathbf{x}_2(k) \quad (2)$$

가 된다. 여기서  $\hat{\mathbf{h}}_1(k)$ ,  $\hat{\mathbf{h}}_2(k)$ 는 L개의 차수를 가지는 적응필터의 계수벡터이며, T는 전치를 의미한다.

$$\mathbf{x}_1(k) = [x_1(k), x_1(k-1), \dots, x_1(k-L+1)]^T \quad (3)$$

$$\mathbf{x}_2(k) = [x_2(k), x_2(k-1), \dots, x_2(k-L+1)]^T \quad (4)$$

이며 오차신호를 보다 간편하게 나타내면

$$e(k) = y(k) - \hat{\mathbf{h}}^T(k) \mathbf{x}(k) \quad (5)$$

여기서  $\hat{\mathbf{h}}(k)$ ,  $\hat{\mathbf{x}}(k)$ 는 아래와 같이 정의한다.

$$\hat{\mathbf{h}}(k) = [\hat{\mathbf{h}}_1^T(k) \mid \hat{\mathbf{h}}_2^T(k)]^T \quad (6)$$

$$\hat{\mathbf{x}}(k) = [\hat{\mathbf{x}}_1^T(k) \mid \hat{\mathbf{x}}_2^T(k)]^T \quad (7)$$

충분히 큰 L에 대해서 스테레오 반향신호 y(k)는

$$y(k) = \mathbf{h}_1^T(k) \mathbf{x}_1(k) + \mathbf{h}_2^T(k) \mathbf{x}_2(k) = \mathbf{h}^T(k) \mathbf{x}(k) \quad (8)$$

로 표현할 수 있다. 적응 알고리즘으로 LMS 알고리즘을 활용할 때 다음식과 같다.

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \mu e(k) \mathbf{x}(k) \quad (9)$$

식(9)에서 오차신호의 평균자승이 최소가 되도록 적응필터의 계수를 갱신한다. 각 채널에 대한 최적계수 벡터들을 구해보면 다음과 같다.

$$\mathbf{h}_{1opt} = \mathbf{R}_1^{-1} \mathbf{P}_1 = \mathbf{h}_1 + \mathbf{R}_1^{-1} \mathbf{R}_{12} \mathbf{h}_2 \quad (10)$$

$$\mathbf{h}_{2opt} = \mathbf{R}_2^{-1} \mathbf{P}_2 = \mathbf{h}_2 + \mathbf{R}_2^{-1} \mathbf{R}_{21} \mathbf{h}_1 \quad (11)$$

여기서  $\mathbf{R}_1$ ,  $\mathbf{R}_2$ 는 각 채널신호들의 자기상관이고  $\mathbf{R}_{12}$ ,  $\mathbf{R}_{21}$ 는 두 채널간의 상호상관이고,  $\mathbf{P}_1$ ,  $\mathbf{P}_2$ 는 각 채널의 입력신호와 목표신호와의 상호상관이다.

식 (10)과 (11)에서 알 수 있듯이 스테레오 신호간의 상호상관에 의해 각 적응필터의 계수들은 반향경로 임펄스응답과 일치하지 않는다. 또한 스테레오 반향 제거기는 유일 해를 갖지 못한다[3]. 그리고 전송실의 갑작스런 경로변화가 생기면 적응필터의 계수는 전송실의 경로변화를 추정하려고 하여 수신실의 반향경로 추정이 어려워진다는 점이다. 그러므로 바람직한 스테레오 반향 제거기는 이러한 전송실의 경로변화에 민감하지 않아야 한다.

이러한 문제점들을 해결하고자 여러 가지 방법이 제안되었다. 기존의 제안된 방법들을 크게 두 가지 형태로 분류할 수 있는데 첫째로 전처리 필터를 통해서 스테레오 신호간의 상호상관을 줄여 성능을 개선하는 경우이다[1][2]. 둘째는 전처리 필터를 사용하지 않고, 적응 알고리즘 자체에서 적응필터의 계수를 갱신할 때 스테레오 신호간의 상호상관을 줄인 입력신호를 사용하여 스테레오 신호의 상호상관에 의한 영향을 줄이려고 하였다[5]. 전처리 필터 구조를 가지는 방법 중 구조가 간단하고, 효율적으로 채널 신호의 상관도를 줄인 방법은 반과정류필터를 사용한 경우다[6]. 단순히 각 채널신호의 양의 구간에 해당하는 값들에 대해서만 일정비를 감쇠시켜 본래의 채널신호에 더해 스테레오 신호간의 상호상관을 줄이는 구조다. 즉 반과정류필터의 비선형 특성을 이용하여 상호상관을 줄이는 구조이다. 반과정류필터를 이용한 스테레오 음향 반향 제거기는 구조가 간단하고 더하는 성분이 본래의 채널신호의 값과 많이 닮아 있어 음질(quality) 및

공간감(spatial perception)을 거의 손상시키지 않는다는 장점이 있지만 단순히 양의 구간동안 자기 자신의 신호를 더하는 구조이므로 스테레오 신호간의 상호상관을 크게 줄이지 못하는 단점이 있다.

### 3. 제안하는 스테레오 반향 제거기

본 논문에서는 스테레오 신호간의 차성분의 절대값을 이용한 새로운 전처리 필터를 제안한다.

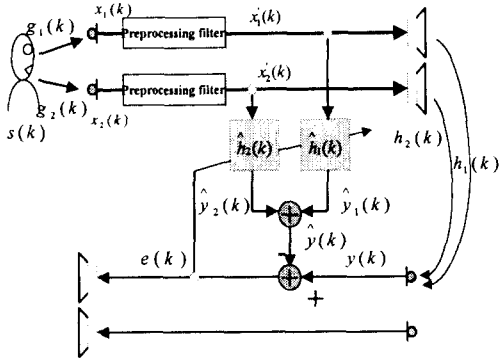


그림 2 전처리필터를 가진 스테레오 음향반향 제거기  
Fig. 2. Basic scheme of stereo echo canceller with preprocessing filter

스테레오 신호는 매우 넓어 있으므로 차성분은 작은 크기를 가진다. 그림 2에서  $x_1'(k)$ 와  $x_2'(k)$ 가 원래 스테레오 신호의 음질과 공간감을 그대로 유지하기 위해서  $x_1(k)$ 와  $x_2(k)$  신호의 차성분에 절대값을 취한 후 시변하는 감쇠상수  $\rho(m)$ 을 곱한다. 그리고 이 신호를 스테레오 신호  $x_1(k)$ 와  $x_2(k)$ 에 더해서  $x_1'(k)$ 와  $x_2'(k)$ 를 구한다. 절대값은 비선형 함수이기 때문에  $x_1'(k)$ 와  $x_2'(k)$ 사이에 선형성을 줄이고, 차성분에 시변하는 감쇠상수  $\rho(m)$ 를 곱하면 더해지는 성분의 크기가 작아진다. 제안하는 전처리 필터를 수식으로 전개하면 다음과 같다.

$$x_i'(mL) = x_i(mL) + \rho(m) \cdot |x_1(mL) - x_2(mL)| \quad (12)$$

$$\rho(m) = \frac{\|x_1\|^2 + \|x_2\|^2}{2x_1^T x_2} \sqrt{\frac{(\|x_1\|^2 + \|x_2\|^2)^2 - 4(x_1^T x_2)^2}{2x_1^T x_2}} \quad (13)$$

$$i=1,2 \quad m=1,\dots,M/L$$

여기서  $i=1,2$ 는 스테레오 신호를 의미하고  $L$ 은 적응 필터 차수,  $M$ 은 채널신호의 길이이다.

하지만 식 (13)에서  $\rho(m)$ 의 절대값이 0.5 보다 크면 음질과 공간감이 손상되기 때문에  $\rho(m)$  범위를 -0.5에서 0.5 사이로 제한한다. 적응 알고리즘은 NLMS[4]와 APA2(2nd order affine projection algorithm)[2]를 사용하였다. 제안하는 전처리 필터는 구조가 간단하고 반과정류필터와 달리 전구간에 걸쳐 더하는 구조이므로 감쇠상수가 작아도 스테레오 신호간의 상호상관을 보다 더 줄일 수 있고,  $\rho(m)$ 을 블록단위로 시변시킨다. 또한 감쇠상수가 작으면 본래의 스테레오 신호의 음질과 공간감을 해치지 않으며, 전송실의 경로변화에 의한 적응 알고리즘의 계수 오조정을 보다 줄일 수 있다. 그 이유는 상호상관이 줄어들면 식(10)과 식(11)에서 두 번째 성분이 줄어들어 최적값에 보다 더 가까이 접근함으로써 전송실의 경로변화가 생겨도 적응 필터 계수의 변화는 최적 값을 중심으로 가까운 곳에서는 적은 변화를 보다 먼 곳에서는 보다 큰 변화를 보이기 때문이다.

### 4. 시뮬레이션 결과 및 검토

시뮬레이션에 사용한 음성신호는 '택시는 곧장 앞으로 빠져 나왔다' 라는 여성음성이다. 이 음성신호는 무향실에서 녹음한 16 비트 양자화, 16kHz 표본화 된 음성신호이다. 전송실의 임펄스응답  $g_1(k), g_2(k)$ 와 수신실의 임펄스응답  $h_1(k), h_2(k)$ 는 2채널 DAT 레코더로 각각 실제 회의실에서 측정하였으며 필터 길이는 1024탭으로 했다. 스테레오 신호는 녹음한 음성데이터와 전송실의 임펄스응답을 각각 콘볼루션해서 얻었고, 수신실의 스테레오 반향신호는 수신실의 임펄스응답과 스테레오 신호를 각각 콘볼루션하고 더하여 얻었다. 적응필터의 길이는 1024탭으로 임펄스응답의 길이와 같이 했고, 시뮬레이션에 사용한 적응 알고리즘 NLMS와 APA2의 전력은 running power estimation을 이용해서 추정했으며, 망각지수는 0.998, 적응상수는 0.2로 동일하게 두었다. 그리고 본 시뮬레이션에서는 농시통화와 배경잡음을 고려하지 않았다. 그림 3,4에 시뮬레이션 결과를 나타내었다. 전처리 필터를 사용하지 않은 NLMS와 APA2 알고리즘보다 5~10dB 정도 개선된 것을 알 수 있다. 또한 반과정류필터를 이용한 경우보다 5dB 정도 개선된 것을 알 수 있다.

그리고 NLMS 알고리즘을 사용하는 것보다 APA2 알고리즘을 사용했을 때 전체적으로 5dB 정도 성능이 우수하다는 것을 알 수 있었다. APA2 알고리즘은 채널 신호내에서 상호상관을 줄이기 때문에 유색신호 입력에 대해서 NLMS 알고리즘 보다 적응속도가 빠르다.

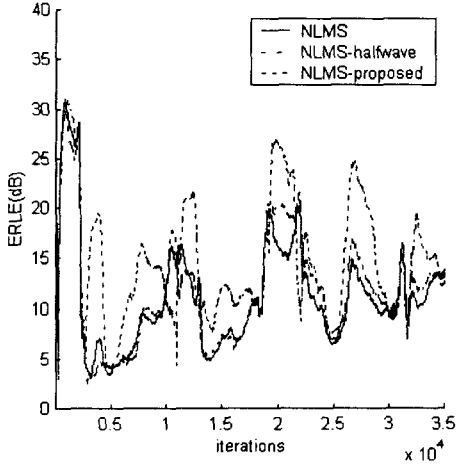


그림 3 NLMS 알고리즘을 사용한 스테레오 음향반향 제거기의 ERLE  
Fig. 3. ERLE curve of stereo echo cancellation using NLMS algorithm

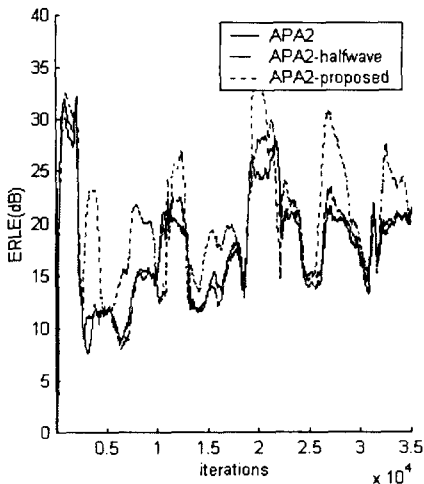


그림 4 APA2 알고리즘을 사용한 스테레오 음향반향 제거기의 ERLE  
Fig. 4. ERLE curve of stereo echo cancellation using APA2 algorithm

## 5. 결론

스테레오 음향 반향 제거기는 두 채널 신호의 높은 상호상관에 의해서 반향을 충분히 제거하지 못한다. 본 논문에서는 스테레오 음향 반향 제거기의 성능을 향상시키기 위해서 스테레오 신호간의 차성분과 시변 감쇠상수를 이용하여 상호상관을 줄이는 전처리 필터를 제안하였다. 컴퓨터 시뮬레이션을 통해서 제안하는 방법이 기존의 방법에 비해서 약 5~10dB 정도 성능이 개선됨을 보였다. 제안하는 방법의 계산량을 줄이기 위한 연구를 계속하고자 한다.

## [참고문헌]

- [1] M. M Sondhi and D. R. Morgan, "Acoustic Echo Cancellation for Stereophonic Teleconferencing", 1991 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 1991
- [2] S. Shimauchi, S. Makino, "Stereo Projection Echo Canceller with True Echo Path Estimation", International Conference on Acoustics, Speech and Signal Processing 94, pp.245-248, 1994
- [3] M. M Sondhi and D. R. Morgan, J. L. Hall, "Stereo Acoustic Echo Cancellation - An Overview of the Fundamental Problem", IEEE Signal Processing Letters, Vol. 2, No 8, pp. 148-151, Aug. 1995
- [4] J. Benesty, F. Amand, A. Gilloire, Y. Grenier, "Adaptive Filtering Algorithms for Stereophonic Acoustic Echo Cancellation", International Conference on Acoustics, Speech and Signal Processing 95, pp.3099-3101, 1995
- [5] J. Benesty, P. Duhamel, Y. Grenier, "A Multichannel Affine Projection Algorithm with Applications to Multichannel Acoustic Echo Cancellation", IEEE Signal Processing Letters, Vol. 3, No 3, pp. 35-37, Feb. 1995
- [6] J. Benesty, D. R. Morgan, M. M. Sondhi, "A Better Understanding and an Improved Solution to the Problems of Stereophonic Acoustic Echo Cancellation", International Conference on Acoustics, Speech and Signal Processing 97, pp.303-306, 1997