

## 스테레오 음향 반향 제거를 위한 적응 알고리즘

정양원, 이지하, 박영철\*, 윤대희  
연세대학교 전자공학과, \*연세대학교 신호처리 연구센터

### A New Adaptive Algorithm for Stereophonic Acoustic Echo Cancellation

Yang Won Jung, Ji Ha Lee, Young Cheol Park\*, Dae Hee Youn  
Dept. of Electronic Eng, Yonsei Univ, \*Center for Signal Processing Research, Yonsei Univ.  
e-mail : ywjung@radar.yonsei.ac.kr

#### 요 약

현장감 있는 원거리 회의를 위해서는 두 채널 이상의 신호를 전송하는 스테레오 시스템이 필요하다. 이러한 원거리 회의 시스템의 경우 스피커와 마이크로폰간의 커플링에 의해 발생하는 음향 반향을 제거하기 위해 스테레오 음향 반향 제거기가 요구된다. 본 논문에서는 스테레오 음향 반향 제거기의 성능을 효과적으로 개선하기 위하여 인간의 청각 특성과 음성 신호 특성을 이용한 전처리단을 제안한다. 제안한 전처리단은 마스킹 현상을 이용하여 청각적으로 인지할 수 없는 크기의 부가 잡음을 음성 신호 분석을 통해 발생시켜 원 신호에 더해줌으로 두 채널 신호의 상호 상관 관계를 낮추어 준다. 또한 본 논문에서는 적응 반향 제거 알고리즘인 Affine Projection(AP) 알고리즘이 Gram-Schmidt(GS) 직교화 과정을 통한 벡터 연산으로 갱신될 수 있음을 이용하여, 별도의 음성 분석 과정 없이 AP 알고리즘의 GS 직교화 단으로 제안한 전처리단을 구현할 수 있음을 보이고, 최종적으로 전처리단과 적응 알고리즘이 결합된 스테레오 음향 반향 제거를 위한 적응 알고리즘을 제안한다.

#### 1. 서 론

현장감 있는 원거리 회의를 위해서는 두 채널 이상의 신호를 전송하는 스테레오 시스템이 필요하다. 이러한 원거리 회의 시스템의 경우 스피커와 마이크로폰간의 커플링에 의해 발생하는 음향 반향을 제거하기 위해 스테레오 음향 반향 제거기가 요구된다[1]. 스테레오 음향 반향 제거기는 일반적으로 단일 채널 음향 반향 제거기를 두 개의 채널로 확장하여 구성되며, 그 기본 구조는 그림 1과 같다. 수신실의 두 채널 마이크로폰 입력은 서로 대칭적으로 구성되므로 편위상 하나의 마이크로폰으로 입력되는 반향 성분들에 대하여 고려한 후 이를 두 채널로 확장하여 적용할 수 있다. 전송실에 위치한 화자의 신호  $S$ 는 화자와 두 마이크로폰 사이의 반향경로  $G_1(z)$ 과  $G_2(z)$ 를 거쳐 각각의 마이크로폰으로 입력된다. 마이크로폰의 출력  $x_1(k)$ 과  $x_2(k)$ 는 수신실의 두 스피커의 입력이 되고, 두 스피커로부터 하나의 마이크로폰까지의

반향 경로  $H_1(z)$ 과  $H_2(z)$ 를 통과하여 마이크로폰의 반향 입력

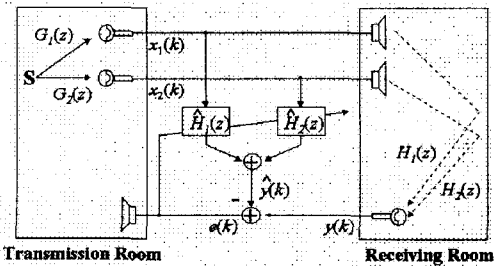


그림 1. 스테레오 음향 반향 제거기의 구조

신호  $y(k)$ 를 만들어 낸다. 따라서 한 채널에 대하여 반향 경로  $H_1(z)$ 과  $H_2(z)$ 를 추정하는 적응 필터  $\hat{H}_1(z)$ 과  $\hat{H}_2(z)$ 을 구성함으로써 음향 반향을 제거할 수 있다.

스테레오 음향 반향 제거에서는 적응 필터의 입력으로 사용되는 두 채널 신호간의 강한 상호 상관 관계에 의하여 단일 채널에서는 발생하지 않는 여러 가지 문제들이 발생한다. 먼저, 이론적으로 적응 필터의 최적해가 유일하게 존재하지 않으며, 또한 전송실에서의 반향 경로 변화를 수신실의 반향 경로 변화로 오인하여 적응 필터의 수렴이 잘못된 방향으로 이루어 질 수 있다. 이러한 문제를 극복하기 위해 두 채널 입력 신호 간의 상호 상관 관계를 낮추는 여러 가지 방법들이 제안되었다. 각 채널에 상호 상관 관계가 없는 부가 백색 잡음을 더하거나, 각 채널 입력을 비선형 함수에 통과시켜 두 채널 입력 간의 상호 상관 관계를 낮추는 등의 방법은 반향 제거기의 성능 향상에는 효과적이었으나, 전송 신호의 품질을 저해하는 단점을 가지고 있다[2].

본 논문에서는 스테레오 음향 반향 제거기의 성능을 효과적으로 개선하기 위하여 인간의 청각 특성과 음성 신호 특성을 이용한 전처리단을 제안한다. 제안한 전처리단은 마스킹 현상을 이용하여 청각적으로 인지할 수 없는 크기의 부가

잡음을 음성 신호 분석을 통해 발생시켜 원 신호에 더해줌으로 두 채널 신호의 상호 상관 관계를 효과적으로 낮추어 준다. 또한 본 논문에서는 적응 반향 제거 알고리즘인 Affine Projection(AP) 알고리즘이 Gram-Schmidt(GS) 직교화 과정을 통한 벡터 연산으로 갱신될 수 있음을 이용하여, 별도의 음성 분석 과정 없이 AP 알고리즘의 GS 직교화단으로 제안한 전처리단을 구현할 수 있음을 보이고, 최종적으로 전처리단과 적응 알고리즘이 결합된 스테레오 음향 반향 제거를 위한 적응 알고리즘을 제안한다.

## 2. 제안된 알고리즘

### 2.1 제안된 전처리단

스테레오 음향 반향 제거를 위한 기존의 방법들은 전송되는 두 채널 신호의 상호 상관 관계를 낮추기 위해 전처리단을 구성하여 전송 신호에 변형을 가한다. 그러나, 이러한 전처리단의 사용으로 인해 전송되는 신호의 정보가 변화되어서는 안되며, 수신실의 청취자가 신호의 왜곡을 인지할 수 없어야 한다. 기존의 방법들은 스테레오 음향 반향 제거기의 반향 제거 성능을 향상시키는 데에는 효과적이었지만, 재생되는 신호의 품질을 저하시키는 단점을 가지고 있다.

이러한 단점을 극복하기 위하여 심리 음향 현상인 마스킹 효과를 이용하여 전처리단을 구성하는 방법이 제안되었다 [3]. 이 방법에서는 부가 잡음을 마스킹 임계값보다 작은 크기로 각 채널에 더함으로, 전송 신호의 왜곡이 인지되지 않으면서도 신호의 상호 상관 관계를 낮출 수 있다. 그러나, 심리 음향 모델을 이용하기 위해서는 마스킹 임계값을 구하여야 하는데, 이 과정은 매우 복잡한 계산을 필요로 하기 때문에 실제 시스템에 적용하기에는 적합하지 않다.

본 논문에서는 기존의 방법들의 단점을 극복하고 스테레오 음향 반향 제거기의 성능을 향상시키기 위해 음성신호 특성을 이용한 전처리단을 제안한다. 일반적으로 핸드프리가 사용되는 환경은 주로 음성 신호의 전송을 목적으로 하므로, 그림 1에서의 화자 신호  $S$ 를 음성 신호라 가정할 수 있다. 음성 신호의 경우, 음성의 AR 스펙트럼인  $A(z)$ 는 음성의 파워 스펙트럼의 포락선을 나타내기 때문에, AR 스펙트럼을 이용하여 마스킹 패턴을 근사적으로 표현할 수 있다. 따라서 AR 분석을 통해 각 채널에 인지되지 않는 유사 음성 잡음을 더해줌으로써 재생 신호의 왜곡이 인지되지 않으면서 상호 상관 관계를 효과적으로 낮출 수 있다.

AR 분석은 Levinson-Durbin 알고리즘이나 격자 예측기, GS 직교화기 등의 어떠한 선형 예측 방법으로도 가능하다. 유사 음성 잡음을 생성하기 위해서는 AR 분석의 역과정인 AR 합성을 이용할 수 있다. 그림 2는 한 채널에 대하여 제안된 방법으로 유사 음성 잡음을 더해주는 과정을 나타낸다.

먼저, AR 계수를 얻기 위해 선형 예측기로 전송된 음성 신호  $x(k)$ 를 AR 분석한다. 유사 음성 잡음  $\hat{n}(k)$ 는 추정된 AR 계수를 이용하여 AR 합성을 통해 발생된다. 최종적으로 입력 채널에 더해지는 부가 잡음  $\hat{n}(k)$ 는 시변하는 이득  $\alpha$ 에 의해 인 지되지 않는 크기로 유지된다. 시변 이득  $\alpha$ 는 가변 SNR에 의해 다음 식과 같이 조절된다.

$$\alpha = \sqrt{\frac{\sigma_x^2(k)}{\sigma_n^2(k)} \cdot 10^{-\frac{SNR}{10}}} \quad (1)$$

여기서  $\sigma_x^2(k)$ 는 입력 신호의 파워,  $\sigma_n^2(k)$ 는 유사 음성 잡음의 파워, 그리고 SNR은 실험을 통하여 정해진 잡음의 크기

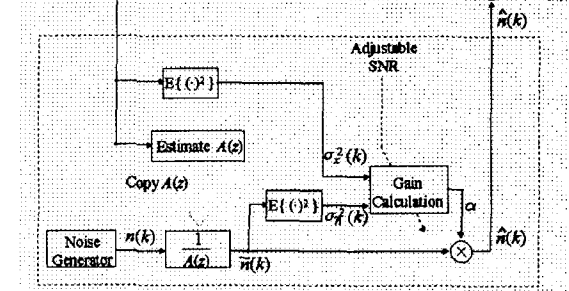


그림 2. 제안된 전처리단의 구조도

가 인지되지 않도록 하는 가변 값이다.

제안된 전처리단의 복잡도는 심리음향 모델을 적용하는 기존의 방법에 비해 월등히 적다. 그러나 제안된 방법을 사용함으로 음성의 품질 저하 없이 전송신호의 상호 상관관계를 크게 낮출 수 있다.

### 2.2 GS-AP 알고리즘

AP 알고리즘은 P차의 후행 오차를 0으로 만드는 적응 알고리즘으로, 다음과 같은 계수 갱신 식을 갖는다[4].

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \mathbf{X}(k) (\mathbf{X}(k)^T \mathbf{X}(k))^{-1} \mathbf{e}(k) \quad (2)$$

여기서  $\hat{\mathbf{h}}(k)$ 는 L개의 계수를 갖는 적응 필터의 계수 벡터이고,  $\mathbf{X}(k) = [\mathbf{x}(k-P+1) \ \mathbf{x}(k-P+2) \ \dots \ \mathbf{x}(k)]$ 는 입력 신호 벡터  $\mathbf{x}(k) = [x(k) \ x(k-1) \ \dots \ x(k-L+1)]^T$ 로 구성되는 입력 신호 행렬이고,  $\mathbf{e}(k) = [e(k-P+1) \ e(k-P+2) \ \dots \ e(k)]^T$ 는 오차 벡터이다.  $[\cdot]^T$ 는 벡터와 행렬의 전치를 나타낸다. 여기서 신호 벡터는 수식 유도상의 편의를 위해 시간 역순으로 구성되어 있다.

AP 알고리즘의 후행 오차 성질에 의하여 오차 벡터는 다음과 같이 나타난다.

$$\mathbf{e}(k) = [0 \ 0 \ \dots \ e(k)]^T \quad (3)$$

입력 행렬  $\mathbf{X}(k)$ 를 직교화 시키는 GS 직교화 행렬  $\mathbf{T}(k)$ 에 의해 다음과 같이 변환된 입력 신호 행렬  $\mathbf{Z}(k)$ 를 얻을 수 있다.

$$\mathbf{Z}(k) = \mathbf{X}(k) \mathbf{T}(k) \quad (4)$$

여기서  $\mathbf{Z}(k) = [\mathbf{z}_1(k) \ \mathbf{z}_{p-1}(k) \ \dots \ \mathbf{z}_1(k)]$  이고  $\mathbf{z}_i(k)$  ( $i=1, 2, \dots, P$ )은  $(L \times 1)$ 의 변환 입력 벡터이다.

GS 변환 행렬  $\mathbf{T}(k)$ 는 다음 식과 같이 하위 단계로 분해하여 나타낼 수 있다[5].

$$\mathbf{T}(k) = \mathbf{T}_1(k) \mathbf{T}_2(k) \dots \mathbf{T}_{p-1}(k) \quad (5)$$

여기서  $\mathbf{T}_i(k)$ 는 i번째 하위 단계의 변환 행렬을 나타내며, 아래 식과 같이 i번째 행과 대각에만 0이 아닌 값을 가지고 있다.

$$\mathbf{T}_i(k) = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 1 & a_{i,p-i} & \cdots & a_{i1} \\ 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \end{bmatrix}, \quad i=1,2,\dots,P-1. \quad (6)$$

변환된 입력 행렬  $\mathbf{Z}(k)$ 는 다음과 같은 특징을 갖는다.

$$\mathbf{Z}(k)^T \mathbf{Z}(k) = \text{diag}[\mathbf{z}_p(k)^T \mathbf{z}_p(k) \quad \mathbf{z}_{p-1}(k)^T \mathbf{z}_{p-1}(k) \quad \cdots \quad \mathbf{z}_1(k)^T \mathbf{z}_1(k)] \quad (7)$$

$\mathbf{T}(k)$ 가 unit upper triangular (UUT) 행렬이고, 행렬식이 1이므로,  $\mathbf{T}(k)^{-1}$ 이 언제나 존재하며 다음과 같은 관계를 항상 만족시킨다.

$$\mathbf{X}(k) = \mathbf{Z}(k) \mathbf{T}(k)^{-1}. \quad (8)$$

그러므로, AP 알고리즘의 계수 갱신항  $\Delta \hat{\mathbf{h}}(k)$ 는 다음과 같이 변형된다.

$$\begin{aligned} \Delta \hat{\mathbf{h}}(k) &= \mathbf{X}(k) (\mathbf{X}(k)^T \mathbf{X}(k))^{-1} \mathbf{e}(k) \\ &= \mathbf{Z}(k) (\mathbf{Z}(k)^T \mathbf{Z}(k))^{-1} \mathbf{T}(k)^T \mathbf{e}(k). \end{aligned} \quad (9)$$

$\mathbf{T}(k)$ 가 UUT 이고,  $\mathbf{e}(k)$ 가 식 (3)과 같이 나타나므로, 식 (9)의 행렬-벡터 곱셈은 다음과 같이 간단하게 표현된다.

$$\mathbf{T}(k)^T \mathbf{e}(k) = \mathbf{e}(k). \quad (10)$$

식 (9)와 (10)에 의하여  $\Delta \hat{\mathbf{h}}(k)$ 는 역행렬의 연산 대신에 1차원의 벡터 연산으로 구해질 수 있다.

$$\begin{aligned} \Delta \hat{\mathbf{h}}(k) &= \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \mathbf{e}(k) \end{bmatrix} \\ &= \frac{\mathbf{z}_1(k)}{\mathbf{z}_1(k)^T \mathbf{z}_1(k)} \cdot \mathbf{e}(k). \end{aligned} \quad (11)$$

최종적으로, GS-AP 알고리즘의 계수 갱신식은 다음과 같이 얻어진다.

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \frac{\mathbf{z}_1(k)}{\mathbf{z}_1(k)^T \mathbf{z}_1(k)} \cdot \mathbf{e}(k). \quad (12)$$

식 (12)에서부터 GS-AP 알고리즘은 NLMS 알고리즘과 매우 유사함을 볼 수 있다.

## 2.3 스테레오 음향 반향 제거를 위한 적응 알고리즘

이 절에서는 스테레오 음향 반향 제거를 위한 새로운 적응 알고리즘을 제안한다. 제안된 알고리즘은 GS-AP 알고리즘을 앞서 제안된 전처리단과 결합되어 구성된다. GS 직교화는 선형 예측의 한 방법으로 볼 수 있고, 따라서 주어진 입력 신호에 대해 AR 분석을 수행할 수 있다. GS-AP 알고리즘을 적응 알고리즘으로 이용하여 스테레오 음향 반향 제거를 구성할 경우, GS-AP 알고리즘을 구성하는 GS 직교화 행렬을 AR 분석과 입력 신호 변환의 수행에 동시에 이용할 수 있다.

유사 음성 잡음을 생성하기 위한 AR 합성에는 IGS(inverse GS) 변환을 이용할 수 있다. IGS 변환 행렬은 GS 행렬의

역행렬로 주어진다. GS 변환 행렬  $\mathbf{T}(k)$ 가 식 (5)와 같이 나타나므로, IGS 변환 행렬  $\mathbf{T}(k)^{-1}$ 는 다음과 같이 얻어진다.

$$\mathbf{T}(k)^{-1} = \mathbf{T}_{p-1}(k)^{-1} \mathbf{T}_{p-2}(k)^{-1} \cdots \mathbf{T}_1(k)^{-1}. \quad (13)$$

또한,  $\mathbf{T}_i(k)$ 가 식 (6)과 같이 표현되므로,  $\mathbf{T}_i(k)^{-1}$ 는 다음과 같이 구할 수 있다.

$$\mathbf{T}_i(k)^{-1} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 1 & -a_{i,p-i} & \cdots & -a_{i1} \\ 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \end{bmatrix}, \quad i=1,2,\dots,P-1. \quad (14)$$

여기서  $\mathbf{T}_i(k)^{-1}$ 는  $\mathbf{T}_i(k)$ 의 대각 위치에 있지 않은 0이 아닌 원소의 부호를 바꾸어줌으로 간단히 얻을 수 있다. 따라서, IGS는 식 (13)과 같이 GS 직교화의 모듈 구조를 그대로 이어받기 때문에 매우 간단하게 구성할 수 있다.

제안된 잡음 생성기는 GS 직교화 행렬의 원소를 이용하여 구성되며 다음과 같이 유사 음성 잡음을 합성한다.

$$\tilde{\mathbf{N}}(k) = \mathbf{N}(k) \mathbf{T}(k)^{-1} \quad (15)$$

여기서  $\mathbf{N}(k)$ 는 독립 백색 잡음 행렬이고,  $\tilde{\mathbf{N}}(k)$ 는 유사 음성 잡음 행렬이다.

잡음의 이득은 식 (1)과 동일하게 조절되어  $\hat{\mathbf{N}}(k)$ 를 얻어낸다. 제안된 알고리즘의 종합적인 과정을 그림 3에 나타내었다. 여기서 AR 계수를 추정하기 위한 입력 신호는  $\mathbf{x}_i(k)$ 가 아닌 유사 음성 잡음이 더해진  $\hat{\mathbf{x}}_i(k)$ 임을 그림에서 볼 수 있다. 그러나  $\hat{\mathbf{x}}_i(k)$ 를 사용하여 AR 분석을 수행하는 경우와  $\mathbf{x}_i(k)$ 를 사용하는 경우 주관적 음질 평가에서는 아무런 차이가 없음을 확인하였는데, 이것은 유사 음성 잡음과 음성 신호가 서로 같은 AR 특징을 공유하기 때문이다.

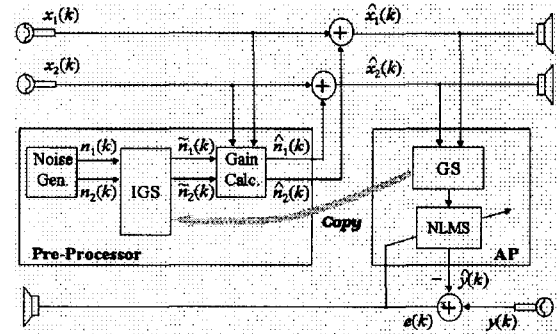


그림 3. 제안된 스테레오 음향 반향 제거 알고리즘

## 3. 모의 실험

### 3.1 주관적 음질 평가

제안된 전처리단의 성능을 평가하기 위하여 주관적 음질 평가를 수행하였다. 본 논문에서는 mean opinion score (MOS) test를 25명의 평범한 청취자를 대상으로 하였고, 백색 잡음

을 각 채널에 더하는 방법과 함께 음질 평가를 수행하였다. 주관적 음질 평가 결과를 그림 4에서 보였다.

제안된 방법은 백색 잡음을 이용하는 방법에 비하여 SNR에서 20dB 이상의 이득을 보였다. 이것은 바람직한 음질을 유지하면서 (MOS 4 점 이상) 백색 잡음 방법에 비해 20dB 나 많은 잡음을 더할 수 있음을 의미한다. 따라서, 제안된 방법을 이용하여 두 채널 신호간의 상호 상관 관계를 더욱 제거할 수 있으며, 이로 인해 음향 반향 제거기의 성능을 향상시킬 수 있다.

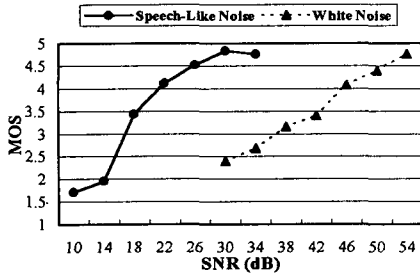


그림 4. 제안한 방법의 MOS

### 3.2 반향 제거 성능

제안된 알고리즘의 성능을 평가하기 위하여 모의 실험을 수행하였다. 성능 비교를 위하여 백색 부가 잡음을 이용하는 방법과 비선형 변환을 이용하는 방법을 비교 대상으로 선택하였다. 비선형 변환 방법에 대해서는 아주 적은 왜곡만을 느낄 수 있는 값인 0.5로 비선형 인수를 결정하였고[2], 백색 잡음을 이용하는 방법과 제안한 방법의 경우 SNR을 MOS가 4점이 될 때의 SNR, 즉 45dB와 20dB로 정하였다. 전송실의 음성 입력 신호  $S$ 는 8kHz로 표본화 되었으며, 두 채널 전송 신호  $x_1(k), x_2(k)$ 는 실측된 1024 탭의 음향 반향 경로를 이용하여 합성되었다. 전송실의 반향 경로 변화를 나타내기 위해 전송실의 반향 경로를 2.5초에서 3.5초 사이에 서서히 변화시켰다. 제안된 알고리즘은 각 채널당 256 탭을 이용하였고, AP 알고리즘의 투영 차수는 11차로 하였다. 음향 반향 제거기의 성능을 ERLE(Echo Return Loss Enhancement) 측면에서 평가하였으며, 스테레오 환경에서 적응 필터가 실제 반향 경로로 수렴하는 정확도를 평가하기 위하여 적응 필터의 misalignment를 평가하였다.

그림 5와 그림 6은 ERLE 곡선과 misalignment 곡선을 나타낸다. 그림 5로부터 제안된 알고리즘이 다른 방법에 비하여 20dB 이상 성능 향상을 보임을 확인할 수 있다. 또한 그림 6에서 제안된 알고리즘이 전송실의 경로 변화에 거의 영향을 받지 않음을 확인할 수 있다.

### 4. 결론

본 논문에서는 스테레오 음향 반향 제거를 위한 새로운 적응 알고리즘을 제안하였다. 제안된 알고리즘은 인간의 청각 특성과 음성 신호 분석을 이용하여 적은 연산량으로 인지되지 않는 부가 잡음을 생성하였다. 제안된 알고리즘은 전송

실의 환경 변화에 크게 영향을 받지 않음을 실험적으로 입증하였다.

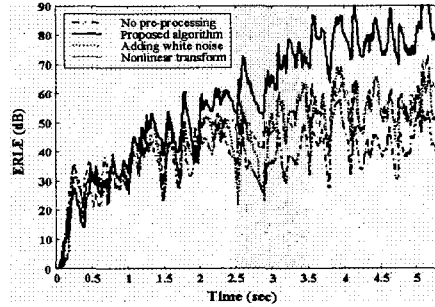


그림 5. 알고리즘간의 ERLE 비교

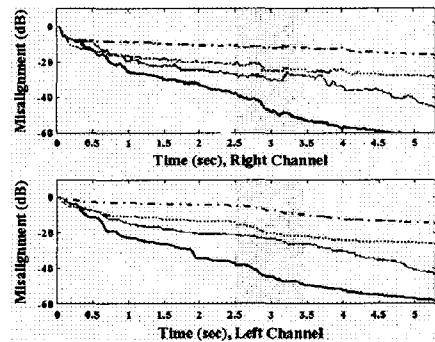


그림 6. 알고리즘간의 Misalignment 비교

### 5. 참고문헌

- [1] M. M. Sondhi, D. R. Morgan and 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
- [2] J. Benesty, D. R. Morgan, M. M. Sondhi, "A Better Understanding and an Improved Solution to the Specific Problems of Stereophonic Acoustic Echo Cancellation," *IEEE Trans. on Speech and Audio Processing*, Vol. 6, No. 2, pp 156-165, Mar., 1998
- [3] A. Gilloire and V. Turbin, "Using Auditory Properties to Improve the Behaviour of Stereophonic Acoustic Echo Cancellers," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing*, pp.3681-3684, 1998.
- [4] K. Ozeki and T. Umeda, "An Adaptive Filtering Algorithm Using an Orthogonal Projection to an Affine Subspace and Its Property," *Electronics and Communication in Japan*, vol.67-A, pp.19-27, May 1984.
- [5] N. Ahmed and D. H. Youn, "On a Realization and Related Algorithm for Adaptive Prediction," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-28, no.5, pp.493-497, Oct. 1980.