

# Hyper-plane Projection 알고리즘을 이용한 의사 스테레오 음향 반향 제거기

박 필구, 조 영민, 이 원철  
숭실대학교 정보통신공학과

## Pseudo Stereophonic Acoustic Echo Canceller using Hyper-plane Projection Algorithm

Phil Goo Park, Young Min Cho, Won Cheol Lee  
CSP Lab., Dept. of Information and Telecommunication Eng., Soongsil University  
e-mail : [wflee@saint.soongsil.ac.kr](mailto:wflee@saint.soongsil.ac.kr)

### 요약

스테레오 음향 반향 제거기에서 발생하는 주요한 문제점들은 모노 환경에서와는 다르게 반향 경로 시스템의 긴 임펄스 응답으로 인한 느린 수렴속도와 원단화자 주위의 환경변화에 의한 최적해의 변화등을 들 수 있다. 이러한 문제점들을 극복하기 위해 본 논문은 전송실에서 환경 변화에 의한 반향 제거 성능저하와 저속의 수렴속도 및 과도한 계산량의 문제점을 해결하기 위하여 본 논문에서는 전송실의 환경 변화에 강인하고 계산량을 줄일 수 있는 Hyper-plane projection 알고리즘을 이용한 의사 스테레오 음향 반향 제거기를 제안한다

### 1. 서론

현재 오디오 채널에 의해 전송되는 음성신호는 공간상의 변화를 제대로 전달할 수 없기 때문에 오히려 통화의 부자연스러움이 발생하게 된다. 따라서 마치 한 공간에서 회의를 하는 듯한 느낌을 주기 위해서는 적어도 2 개 이상의 오디오 채널을 사용하여 음성신호를 전달하는 다채널 원격회 회의 시스템이 미래의 원격회의 시스템으로 주목 받고 있다 [5]. 이러한 다채널 원격회의 시스템의 경우, 여러 개의 마이크로폰들과 스피커의 사용에 기인하는 상호 커플링 현상에 의하여 복잡한 음향 반향 신호들이 발생하게 되고 이 반향 신호는 그대로 전송실로 전달돼 통화에 불편을 주게 된다. 이러한 전송실로 유입되는 반향 신호를 제거하기 위해 음향 반

향제거기는 필수적으로 요구된다.

본 논문에서는 스테레오 원격회의 시스템 상에서 두 스피커 출력들간의 강한 상관도와 음향 반향 경로의 긴 임펄스 응답으로 인하여 적응 반향 제거시 느린 수렴속도와 원단화자 주위의 환경변화로 모노 시스템과는 달리 여러 심각한 성능상의 문제들을 내포한다[4]. 이러한 문제들을 극복하기 위해서 본 논문에서는 계산량을 줄이고 환경 변화에 강인하며 높은 ERLE (Echo Return Loss Enhancement) 를 유지하는 Hyper-plane Projection 알고리즘을 이용한 의사 스테레오 음향 반향 제거기를 제안한다.

### 2. 스테레오 음향 반향 제거기의 문제점

스테레오 원격회 회의 시스템에서 음향 반향을 제거하기 위한 기본 구조는 그림 1과 같다.

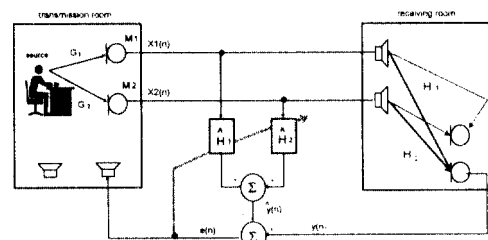


그림 1 스테레오 음향 반향 제거기의 기본 구조

그림 1에서  $G_1$  과  $G_2$  는 전송실에서 원단화자와 각 마이크로폰  $M_1$  과  $M_2$  사이의 음향경로에 대한 전달함

수를 나타내며  $H_1$  과  $H_2$  는 각 스피커로부터 특정 마이크로폰간의 반향 경로를 대표하는 전달함수를 나타낸다. 스테레오 음향 반향 제거기에서 근단화자의 신호가 존재하지 않을 때 마이크로폰으로 유입되는 반향 신호  $y(n)$  를 적응 필터를 사용해 추정한 반향 신호  $\hat{y}(n)$  은 다음과 같다.

$$\hat{y}(n) = \hat{\mathbf{H}}_1^T(n) \mathbf{X}_1(n) + \hat{\mathbf{H}}_2^T(n) \mathbf{X}_2(n) \quad (1)$$

여기서,  $\mathbf{X}_1(n)$  과  $\mathbf{X}_2(n)$  은 각각의 채널 입력신호 벡터들이며  $\hat{\mathbf{H}}_1(n)$  과  $\hat{\mathbf{H}}_2(n)$  은 반향 제거 필터의 계수 벡터들이다. 이때 잔여 반향 오차 신호의 평균 자승값을 최소화 하기 위한 적응 필터  $\hat{\mathbf{H}}_1(n)$  과  $\hat{\mathbf{H}}_2(n)$  의 최적 계수 벡터는 다음과 같다.

$$\mathbf{H}_{opt} = \mathbf{R}^{-1} \mathbf{p} \quad (2)$$

여기서,  $\mathbf{H}_{opt} = [\mathbf{H}_{1opt}^T \mathbf{H}_{2opt}^T]^T$  이고,  $\mathbf{R}$  과  $\mathbf{p}$  는 각각 스테레오 입력 신호들의 상관 행렬과 반향 신호와 입력 신호간의 상호 상관 벡터이다.

식 (2) 에서 최적해를 구하기 위해서는 입력 상관 행렬의 역행렬을 구해야 하는데,  $x_1(n)$  과  $x_2(n)$  는 서로 상관도가 크기 때문에 상관 행렬  $\mathbf{R}$  은 고유치 확산비가 크게 되고 이를 적응 알고리즘에 적용 할 경우 매우 느린 수렴속도를 성립하게 된다. 또한 일반적으로 불 전달함수의 임펄스 응답은 긴 잔여 신호 성분을 가지며 이를 유한 임펄스 응답 형태의 적응 필터로 추정할 경우 많은 계산량을 요구하게 된다. 또 다른 문제점은 전송선에서의 환경변화에 의한 성능 저하이다. 이러한 영향을 살펴 보기 위하여 잔여 반향 신호  $e(n)$  의 Z-변환을 정리하면 다음과 같다.

$$E(z) = [H_1(z) - \hat{H}_1(z)]G_1(z) + [H_2(z) - \hat{H}_2(z)]G_2(z) = 0 \quad (3)$$

여기서, 잔여 반향 신호가 완전히 제거되기 위한 조건, 즉  $E(z)=0$  을 만족시키는  $\hat{H}_1(z)$  과  $\hat{H}_2(z)$  가 존재한다 하더라도  $G_1(z)$  과  $G_2(z)$  가 바뀌게 되면 반향 제거를 위한  $\hat{H}_1(z)$  과  $\hat{H}_2(z)$  의 값이 변하게 된다. 따라서, 전송선의 환경이 변하게 되면 잔여 반향 오차 신호가 커져서 ERLE 값이 저하 된다.

### 3. 의사 스테레오 음향 반향 제거기

의사 스테레오 반향 제거기는 원단화자 측에서 전송된

두 채널 신호들의 합성신호  $x_c(n) = x_1(n) + x_2(n)$  을 가지고 적응 필터를 이용한 전처리단에 의해 의사 스테레오 신호들을 재구성하여 스피커로 출력하고 이에 대한 반향 신호를 제거하는 구조를 지닌다. 이때  $L+M$  길이의 탭수를 갖는 적응 반향 제거 필터를 사용하여 얻을 수 있는 추정 반향 신호  $\hat{y}(n)$  은 다음과 같이 행렬형태로 표현할 수 있다.

$$\hat{y}(n) = \hat{\mathbf{h}}_1^T(n) \hat{\mathbf{X}}_1(n) + \hat{\mathbf{h}}_2^T(n) \hat{\mathbf{X}}_2(n) \quad (4)$$

여기서, 의사 스테레오 신호 벡터  $\hat{\mathbf{X}}_1(n)$  과  $\hat{\mathbf{X}}_2(n)$  는 다음과 같다. [3]

$$\hat{\mathbf{X}}_1(n) = \mathbf{F}(n) \tilde{\mathbf{X}}_c(n) \quad (5)$$

$$\hat{\mathbf{X}}_2(n) = (\mathbf{D} - \mathbf{F}(n)) \tilde{\mathbf{X}}_c(n) \quad (6)$$

여기서,  $\tilde{\mathbf{X}}_c(n) = [x_c(n), x_c(n-1), \dots, x_c(n-L-M+1)]^T$  는  $(L+M)$  개의 이전 합성 신호 샘플로 구성된  $(L+M) \times 1$  벡터이고,  $\mathbf{F}(n)$  은 의사 스테레오 신호를 발생하기 위한 전처리단 적응 필터계수들로 구성된  $L \times (L+M)$  행렬이며,  $\mathbf{D}$  는 첫 번째 구성요소가 1 이고 나머지는 0 인  $(L+M) \times 1$  벡터이다. 식 (5) 와 (6) 을 식 (4) 에 대입하고 행렬의 형태로 정리하면 식 (4) 를 다음과 같이 나타낼 수 있다.

$$\hat{y}(n) = [\hat{\mathbf{h}}_1^T(n) \hat{\mathbf{h}}_2^T(n)] \begin{bmatrix} \mathbf{F}(n) \\ \mathbf{D} - \mathbf{F}(n) \end{bmatrix} \tilde{\mathbf{X}}_c(n) = \mathbf{W}_p^T(n) \tilde{\mathbf{X}}_c(n) \quad (7)$$

여기서,  $\mathbf{W}_p(n) = \hat{\mathbf{h}}_1^T(n) \mathbf{F}(n) + \hat{\mathbf{h}}_2^T(n) (\mathbf{D} - \mathbf{F}(n))$  은 크기가  $(L+M) \times 1$  인 변형된 계수 벡터이다. 식 (7) 에서 보듯이 의사 스테레오 신호를 스피커 출력으로 내보낼 경우 반향 신호는 단일 필터  $\mathbf{W}_p(n)$  을 사용하여 표현할 수 있다. 그리고, 오차 신호  $e(n)$  은 다음과 같다.

$$e(n) = y(n) - \hat{y}(n) = y(n) - \mathbf{W}_p^T(n) \tilde{\mathbf{X}}_c(n) \quad (8)$$

### 4. Hyper-plane Projection 이용한 의사 스테레오 음향 반향 제거기

일반적으로 스테레오 음향 반향 제거를 위한 적응 필터로는 AP (Affine Projection) 알고리즘이 어느 정도 적합하나 이 장에서는 계산량에 대한 부담을 해결하고 전송

실에서의 환경 변화에 강인한 Hyper-plane Projection 알고리즘을 이용한 의사 스테레오 음향 반향 제거기를 제안한다. 그림 2는 제안한 스테레오 음향 반향 제거기 구조를 나타내며 전처리단과 적교화 과정 및 반향 제거 필터로 구성된다.

의사 스테레오 신호의 합성 신호로 구성된  $\tilde{X}_c(n)$  벡터에 대해 크기가  $(L+M) \times (m-1)$  인  $A = [\tilde{X}_c(n-1) \tilde{X}_c(n-2) \dots \tilde{X}_c(n-m+1)]$ 의 행렬에서 각각 벡터들에 공통으로 직교하는 벡터  $u(n)$ 을 Hyper-plane Projection 알고리즘을 이용해 반복적으로 구하는 방법을 제안한다. 일반적인 Affine Projection 알고리즘은  $m$ 개의 a posteriori 반향 제거 오차들을 0으로 만족시키는 반향 제거 가중치들을 계산하는 동작원리를 갖고 있으며 이에 대한 성능은 이미 여러 논문을 통하여 입증되어 왔다. 특히 논문 [1]에서는 projection 차수가  $m=2$ 인 경우 과거 입력 벡터에 대하여 직교성질을 만족하는 새로운 벡터를 구하여 이를 반향 제거기의 입력으로 사용하였으며 이에 대한 결과는 일반적인 Affine projection 알고리즘과 동일함을 입증하였다. 본 논문에서는 일반화된 projection 차수에 대하여 과거 벡터들과 직교성질을 만족하는 벡터를 Hyper-plane projection 알고리즘을 통하여 반복적으로 계산하는 방법을 소개하였고 이에 대한 성능을 고찰하였다.

전처리단에서는 적응 필터로서 적은 계산량의 장점을 갖고 있는 NLMS 알고리즘을 사용하였으며 전처리단의 출력 신호들인  $\hat{X}_1(n)$ 과  $\hat{X}_2(n)$ 의 합성 신호인  $X_c(n)$ 을 Hyper-plane Projection 알고리즘을 이용하여 반향 제거기의 입력으로 사용하였다.

논문 [1]의 정의에 의해  $\beta_1(n)$ 은 다음과 같은 관계가 성립한다.

$$\beta_1(n) = \frac{e(n)}{u^T(n)u(n)} \quad (9)$$

따라서, Projection 차수가  $m$ 인 Projection 알고리즘에 대한 벡터  $u(n)$ 은 다음과 같이 전개할 수 있다.

$$u(n) = x(n) + \frac{\beta_2(n)}{\beta_1(n)}x(n-1) + \frac{\beta_3(n)}{\beta_1(n)}x(n-2) + \dots + \frac{\beta_m(n)}{\beta_1(n)}x(n-m+1) \quad (10)$$

결국, 크기가  $(L+M) \times (m-1)$ 인 행렬  $A$ 는 식 (9)와 (10)에 의해서 다음 관계를 만족한다.

$$u(n) = (I - P_{m-1})\tilde{X}(n) \quad (11)$$

여기서,  $P_{m-1}$ 은  $A(A^*A)^{-1}A^*$ 이고  $u(n)$ 은 벡터  $\tilde{X}_c(n)$ 에 대해  $A$  행렬을 구성하는 각각의 벡터들이 공통으로 직교하는 벡터이다. 그리고, 식 (11)에서  $u(n)$ 을 구하기 위해 Hyper-plane Projection 알고리즘[2]을 이용하고  $K$ 번 반복하여 구한 벡터  $\hat{u}^k(n)$ 의 갱신식은 다음과 같다.

$$\hat{u}_i^k(n) = \hat{u}_i^{k-1}(n) + \frac{e^{i-1}}{\sum_{j=1}^L \|a_j\|^2} a_i(n), \quad \begin{matrix} i = 1 \rightarrow m \\ k = 1 \rightarrow K \end{matrix} \quad (12)$$

여기서  $\hat{u}^k(n)$ 은  $k$ 번째 projection 과정에서  $i$ 번째 과거 입력 벡터와 직교 성질을 만족시키는 벡터이며  $a_j$ 는 행렬  $A = [\tilde{X}_c(n-1) \tilde{X}_c(n-2) \dots \tilde{X}_c(n-m+1)]$ 의  $i$ 번째 행과  $j$ 번째 열 성분이다.

여기서, 식 (12)를 이용하여  $k \times m$ 번의 반복과정을 통하여 얻어진 벡터를  $\hat{u}(n) = \hat{u}_m^k(n)$ 라고 놓으면,  $m$ 개의 Affine Projection 알고리즘의 계수 갱신식은 다음과 같이 간소화된다.

$$\hat{W}(n+1) = \hat{W}(n) + \mu \frac{e(n)}{u^T(n)u(n)} \hat{u}(n) \quad (13)$$

그림 2에서 알 수 있듯이 각 채널에 대한 반향 신호를 제거하는데 두개의 적응 필터를 사용한 기존의 방법에 비해 단일 적응 필터의 사용만으로 가능하게 된다. 그리고 합성 신호  $X_c(n)$ 은  $(L+M)$ 의 길이를 갖는 단일 적응 필터  $W(n)$ 의 기준 입력으로 사용된다. 즉 반향 제거기에서 필요한 계산량은 일반적인 Affine Projection을 이용한 방법보다 줄어든다.

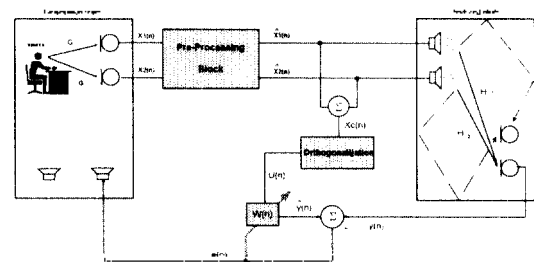


그림 2 제안한 의사 스테레오 음향 반향 제거기의 구조

## 5. 모의 실험

모의 실험에서는 두 개의 서로 다른 소스 신호를 이용하였다. 하나는 음성의 스펙트럼 특성과 유사한 AR 필

터를 통과시켜 만든 유색잡음 신호이며, 또 하나는 8kHz로 샘플링된 실제의 음성 신호이다. 본 논문에서 제안한 방법의 우수성을 보이기 위해 전송실에서의 환경을 순간적으로 변화시켰다. 변화된 환경을 만들어내기 위해 서로 다른 이득을 갖는 전송실에서의 임펄스 응답들을 사용하였고, 그림 3, 5는 고정된 환경이고, 그림 4, 6은 각각 샘플 10000과 5000에서 환경 변화를 주었을 때의 결과이다.

본 논문에서, NLMS(Normalized-LMS) 알고리즘을 전처리단과 음향 반향 제거기의 적응 필터로 사용하였으며, 적응 필터 탭수는 각각  $M=64$ ,  $L=256$ 이다. 그림 3, 4는 유색잡음 신호를 소스 신호로 사용했을 경우 NLMS 방법과 제안한 Hyper-plane 방법에 대한 ERLE 곡선을 나타내고 있다. 이때 수렴상수  $\mu$ 는 1로 고정시켰다.

그림 4, 6에서 알 수 있듯이 전송실의 환경이 변할 경우 반향 제거기의 성능은 나빠지게 된다. 변화된 순간에서의 ERLE 값은 NLMS 방법을 이용한 경우와 비교했을 때 Hyper-plane 방법이 더 높음을 알 수 있다. 게다가 변화 후의 적응 필터는 최적해에 보다 빨리 수렴한다. 여기서, 탭수  $L+M=320$ 을 갖는 단일 적응 필터가 사용되었으며 제안한 방법은 기존의 방법보다 적은 계산량과 보다 향상된 성능을 보임을 알 수 있다.

본 모의 실험에서 추정된 의사 스테레오 신호는 스피커 출력으로 사용된다. 따라서, 제안한 방법의 우수성은 얼마나 의사 스테레오 신호가 실제 스테레오 신호와 유사한지 그리고 송,수신실의 환경이 일정한지에 달려 있다. 따라서, 그림 7, 8에서 실제 신호들과 의사 신호들과의 음질 차이에 대한 객관적인 평가를 프레임 (160 샘플) 별로 나타낸다.

## 6. 결론

본 논문은 적은 계산량과 전송실에서의 환경 변화에 강인하고, 향상된 성능을 갖는 Hyper-plane Projection 알고리즘을 이용한 의사 스테레오 음향 반향 제거기를 제안한다. 전처리단의 적응 알고리즘으로 NLMS를 사용하고 제안한 반향 제거기의 성능을 향상시켰다. 다양한 모의 실험을 통하여 제안한 방법인 의사 스테레오 신호를 사용한 Hyper-plane Projection 알고리즘의 이용으로 일반적인 Affine Projection 알고리즘 보다 적은 계산량을 갖고 환경 변화에 강인함을 확인하였다.

## 참고 문헌

- [1] S.Makino and Y.Kaneda, "Exponentially Weighted Step-size Projection Algorithm for Acoustic Echo Cancellers" IEICE Trans. Fundamentals, vol. E75-A, no.11, pp. 1500-1508, Nov. 1992.
- [2] J. Von Neumann, Functial Operator, vol.II, Annals of Mathematics Studies, No.22, Princeton, New Jersey, 1950
- [3] 정 규화, 이 원철, 차 윤철, 강 현우, 윤 대희, 차 일환, "환경변화에 강인한 스테레오 음향 반향 제거기," 신호처리 합동 학술대회, 제 9권, Part1, pp.23-26, 1996.
- [4] M. M. Sondhi, D. R. Morgan and J. L. Hall, "Stereophonic Acoustic Echo Cancellation-An Overview of the Fundamental Problem," *IEEE Signal Processing Letters*, vol.2, no.8, pp.148-151, Aug. 1995.
- [5] A. Hirano and A. Sugiyama, "A Compact Multichannel Echo Canceller with A Single Adaptive Filter per Channel," *Proc. ICASSP*, pp.1922-1925, 1992.

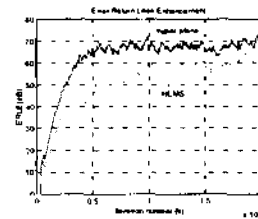


그림 3. AR(6), 고정된 환경

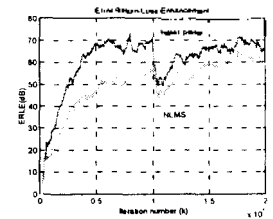


그림 4. AR(6), 변화된 환경

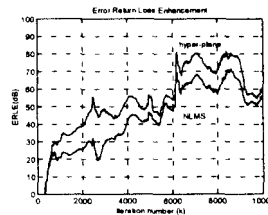


그림 5. 음성, 고정된 환경

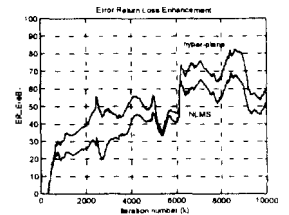


그림 6. 음성, 변화된 환경

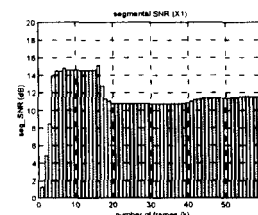


그림 7. 음성, Seg-SNR (X1)

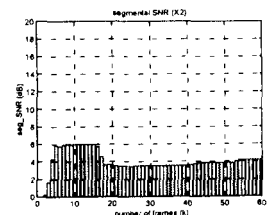


그림 8. 음성, Seg-SNR (X2)