

# Warped Common Acoustical Pole and Zero 방법을 이용한 효율적인 공간 등화 Effective Room Equalization Using Warped Common Acoustical Pole and Zero

이 준 호\*, 박 영 철\*\*, 윤 대 희\*, 이 석 필\*\*\*  
(Junho Lee\*, Young-cheol Park\*\*, Dae-hee Youn\*, Seokpil Lee\*\*\*)

\*연세대학교 전기전자공학과, \*\*연세대학교 원주캠퍼스 정보기술학부,  
\*\*\*전자부품연구원(KETI) 디지털미디어 연구센터  
(접수일자: 2008년 6월 30일; 수정일자: 2008년 11월 7일; 채택일자: 2008년 11월 27일)

본 논문에서는 warped common acoustical pole and zero (WCAPZ) 모델링 방법을 이용한 새로운 공간 등화 방법을 제안한다. 제안한 방법은 저주파 영역의 등화 성능을 감소시키지 않으면서 등화 필터의 차수를 줄일 수 있는 장점을 가진다. 따라서 제안된 공간 등화 시스템은 기존의 블록 변환 방법에 비해 연산량은 비슷하면서도 적은 입력 지연을 가지게 된다. 컴퓨터 모의실험을 통해 제안된 방법이 기존의 기법에 비해 저주파 영역에서 좋은 공간 등화 성능을 보임을 검증하였다.

**핵심용어:** 공간 등화, 워핑, 공동 음향 극·영점

**투고분야:** 음향신호처리 분야 (1,5) (1,3)

This paper presents a new method of designing room equalization filters using a warped common acoustical pole and zero (WCAPZ) modeling. The proposed method is capable of significantly reducing the order of the equalization filters without sacrificing the filter performance, especially, at low frequencies. Thus, the associated input-output delay is much smaller than the conventional block transform method while its computational complexity is comparable to it. The computational complexity also is still comparable to the conventional room equalization method, since the filter is implemented in the linear frequency domain after the pole-zero dewarping. Simulation results confirm that the use of the proposed algorithm significantly improves the room equalization over a range of low frequencies.

**Keywords:** Room equalization, Warping, Common acoustical pole and zero

**ASK subject classification:** Acoustic Signals Processing (1,5) (1,3)

## I. 서론

3차원 공간 역 필터링 (inverse filtering) 이란 디지털 신호 처리 기법을 사용하여 3차원 청취 공간에 의해 왜곡된 전달 함수를 등화시키는 기술이다 [1]. 멀티 채널 재생 시스템에서 3차원 공간 등화는 청취자의 귀주변에서 레코딩된 원 음상을 완벽하게 재생해 내는 것을 목적으로 한다. 이러한 목적을 달성하기 위해 스피커 입력 단에 역

필터 (inverse filter)를 연결하여 스피커에 의한 왜곡과 청취 공간의 음향적인 특성에 의한 변형을 보상한다. 이를 역 필터링 (inverse filtering) 과정이라고 한다.

일반적인 공간 응답은 수초간 지속되므로, 공간 응답을 Moving Average (MA) 모델로 모델링하기 위해서는 많은 수의 파라미터가 필요하다. 이것은 필연적으로 공간 등화 필터의 탭수를 증가시키고, 이와 같이 설계된 공간 등화 필터를 실시간 시스템에 사용하기 위해서는 시간 영역의 컨볼루션 (convolution) 방법보다 적은 연산량으로 구현 가능한 FFT 기반의 블록 컨볼루션 (overlap-add 또는 overlap-save) 방법을 이용하는 것이 효과적이다 [2]. 그러나 이 방법은 50% overlap-add를 가정하였을

책임저자: 이 준 호 (gladday@dsp.yonsei.ac.kr)  
120-749 서울시 서대문구 신촌동 연세대학교 공대 제2공학관 601호  
디지털 신호처리 연구실  
(전화: 02-2123-4534; 팩스: 02-312-4584)

경우, 출력 신호 한 프레임을 만들기 위해 두 개의 입력 신호 프레임이 필요하고, 입력 신호를 FFT 하고 연산하고 다시 IFFT 하는데 지연이 발생하게 된다. 즉, 블록 크기의 2배 이상의 입출력 지연을 발생시켜 비디오 신호와 오디오 신호의 비동기를 유발하는 문제점을 가진다. Gardner는 이와 같은 문제를 해결하기 위해 블록 변환 방법을 사용하면서도 지연을 없게 만드는 필터링 방법을 제안하였다 [3]. 이 방법은 시간 영역의 콘볼루션 보다는 적은 연산량을 가지지만, 주파수 영역의 블록 컨볼루션에 비해서는 여전히 많은 연산량을 가지는 단점을 지니고 있다. 따라서 공간 등화 시스템의 실시간 동작을 위해서는 필터링시 필요한 연산량을 줄이는 노력보다는 적은 필터 뎀수로 효율적인 공간 등화를 할 수 있는 필터를 설계하는 것이 중요하다고 볼 수 있다. 본 논문의 목적은 청취에 중요한 저주파 영역의 등화 성능을 유지하면서 기존의 방법보다 적은 뎀수를 가지는 공간 등화 필터를 설계하는 것이다.

Haneda et. al.은 공간 응답의 공통 음향 극점 (common acoustical pole) 을 이용한 공간 등화 방법을 제안하였다 [4]. 이 방법은 공간 응답을 직접 등화시키지 않고, 여러 개의 공간 응답으로부터 공통되는 음향 극점을 찾아 등화 시킴으로써 적은 뎀수로 공간 등화를 가능하게 한다. 그러나 이 방법은 공통 음향 극점의 차수가 이론 차수보다 적은 경우 성능 저하가 심하고, 공통 음향 극점을 구하는데 과도한 연산량이 드는 단점을 가지고 있다. 본 논문에서는 적은 차수의 등화 필터를 이용하여 청취에 중요한 역할을 하는 저주파 영역의 성능을 유지하는 등화 필터 설계 방법을 제안한다. 이를 위해 기존에 제안된 공통 음향 극점에 기반한 공간 등화 방법과 warped 필터의 장점을 결합하였다. 제안한 방법은 warped 주파수 영역에서 최적화된 등화 필터를 구한 후, 선형 주파수 영역으로 변환하여 구현되므로 기존의 등화 필터 설계 방법과 유사한 연산량을 가진다. 또한 등화 필터의 뎀수가 적으므로, 기존의 필터링 알고리즘보다 입출력 지연이 작은 장점을 가지게 된다.

본 논문의 구성은 다음과 같다. II 장에서는 기존의 공간 등화 방법에 대해 설명하고 이의 문제점을 도출한다. 이를 기반으로 III 장에서는 기존 방법의 문제점을 해결하기 위한 알고리즘과 이를 기반으로 한 공간 등화 시스템을 제안한다. IV장에서는 모의 실험을 통해 제안된 시스템의 성능을 검증하며, V장에서 결론을 맺는다.

## II. 공통 음향 극·영점 방법을 이용한 3차원 공간 등화 시스템

멀티 채널 재생 시스템에서 3차원 공간 등화는 청취자의 귀주변에 레코딩된 원 음장을 완벽하게 재생해내기 위해 스피커 입력 단에 역 필터 (inverse filter)를 연결하여 스피커에 의한 왜곡과 청취 공간의 음향적인 특성에 의한 변형을 보상한다. 본 장에서는 멀티채널 재생 시스템에서의 공간 등화와 그것의 문제점을 기술한다. 그리고 이를 해결하기 위한 공통 음향 극·영점 (common acoustical pole and zero; CAPZ) 방법을 소개한다.

### 2.1. 멀티 채널 재생 시스템에서 3차원 공간 등화 문제

그림 1은 멀티채널 사운드 재생 시스템의 블록 다이어그램을 도시한 것이다. 그림에서 입력 신호, 스피커와 마이크로폰은 각각  $K$ 개,  $M$ 개,  $L$ 개가 있으며, 행렬  $C$ 는 스피커 입력 신호  $y_m(n)$ 과 마이크 위치에서의 에러 신호  $e_l(n)$ 간의 전달 함수, 행렬  $A$ 는 적당한 모델링 지연을 가지는 목적 임펄스 응답 행렬을 의미한다. 따라서 재생하고자 하는 복제신호  $d_l(n)$ 은 레코딩된 입력 신호  $x_k(n)$ 가 전달 함수  $A$ 를 통과한 신호이다. 이 때, 마이크 위치에서 관찰되는 신호  $e_l(n)$ 을 벡터-행렬 식 형태로 적으면 수식 (1)과 같다.

$$e(n) = d(n) - z(n) \tag{1}$$

여기서

$$e^T(n) = [e_1(n) \ e_2(n) \ \dots \ e_L(n)]$$

$$d^T(n) = [d_1(n) \ d_2(n) \ \dots \ d_L(n)]$$

$$z^T(n) = [z_1(n) \ z_2(n) \ \dots \ z_L(n)]$$

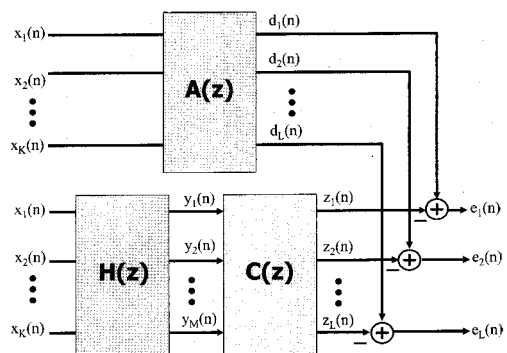


그림 1. 멀티 채널 음향재생 시스템에서의 역 필터링 문제  
Fig. 1. Inverse filtering for a multi-channel sound reproduction system.

이고,  $z(n)$ 은  $y_{ok}(n)$ 이 전달 함수  $C$ 를 통과한 신호를 의미한다. 보상 필터의 최적 계수 벡터는 마이크 위치에서 관찰되는 신호  $e(n)$ 의 전체 에너지를 최소화하는 방법으로 구할 수 있다. 수식 (2)는 보상 필터를 최적해를 구하기 위한 비용 함수를 나타낸다.

$$J(n) = E[e^T(n)e(n)] \tag{2}$$

여기서  $L=K$ 가 같고, 복색 신호가 레코딩된 원음이  $A$ , 만큼 단순 지연된 신호라고 가정하면 목적신호는 다음 수식과 같이 표현될 수 있다.

$$d_l(n) = x_l(n - \Delta_l), \quad l = 0, 1, \dots, L. \tag{3}$$

수식 (4)는 입력 신호가 백색잡음이라고 가정했을 때, 최소자승오차 방법을 이용하여 보상 필터  $H$ 의 계수를 구한 것이다 [1].

$$H_0 = C^{\#} \delta \tag{4}$$

여기서  $\delta$ 는  $A$ 의 위치의 값만 1이고 나머지는 0인 벡터이고,  $\#$ 은 pseudo inverse를 의미한다.

### 2.2 멀티 채널 공간 등화 시스템의 문제점 및 연구의 목표

일반적인 룸 임펄스 응답은 수초간 진행이 되며, 이로부터 구해지는 공간 등화 필터는 긴 차수를 가지게 된다. 일반적으로 홈 엔터테인먼트 시스템은 5.1 채널 환경으로 구축되며, 이 경우 공간 등화 필터는 25개가 필요하다. 이때 각 등화 필터의 차수가 512이고 샘플링 주파수가 44.1 kHz 인 경우, 공간 등화 시스템의 필터링을 위해 시간 영역의 콘볼루션을 사용한다면 약 500 MIPS (million instruction per second)의 연산량이 필요하다. 따라서 연산량을 고려한다면 주파수 영역에서 필터링을 수행하는 블록 콘볼루션을 사용하여야 한다. 하지만 블록 콘볼루션 방법은 필연적으로 블록 크기의 2배 이상의 입출력 지연을 유발하며, 연산량을 줄이기 위해 블록 크기를 크게 하면 입출력 지연이 길어져 실시간 필터링을 할 수 없게 된다 [2]. 하이브리드 콘볼루션 알고리즘은 이러한 문제점을 개선하기 위해 지연 시간과 연산량 사이의 트레이드오프를 조절하여 시간 영역의 콘볼루션 방법에 비해 적은 계산량을 가지면서도 블록 콘볼루션 방법에 비해 적은 지연시간을 가지도록 구현이 가능한 방법이다 [3]. 하지만 이 방법 역시 지연 시간을 작게 하기 위해서는 블록 콘볼루션보다 많은 연산량을 필요로 하는 단점을

가지고 있다. 그림 2는 블록 크기에 따른 출력 한 샘플당 필요한 실수 곱셈 수를 비교한 것이다 [3]. 시간 영역의 필터링 방법을 "direct form"으로 표기하였고, 블록 콘볼루션 방법은 하이브리드 방법과 비교하기 위해 "large block"으로 표기하였다. 하이브리드 방법의 경우 "direct form"과 같이 지연 시간을 0으로 만든 경우와 1.5 ms의 지연시간을 가지는 경우의 연산량을 기존의 방법들과 비교하였다. 필터의 차수를 512라고 할 때, 시간 영역 콘볼루션 방법은 한 샘플당 512번의 실수 곱셈 연산이 필요하고, 지연이 없으며 최소의 연산량이 필요한 하이브리드 콘볼루션 방법은 약 135번의 곱셈 연산이 필요하며, 블록 콘볼루션 방법은 약 35번의 곱셈 연산이 필요하다. 즉 하이브리드 콘볼루션 방법은 시간 영역의 콘볼루션보다 적은 연산량을 가지지만, 블록 콘볼루션에 비해 여전히 많은 연산량을 가짐을 확인할 수 있다. 하지만 블록 콘볼루션 방법은 샘플링 주파수를 44.1 kHz라고 가정했을 때, 블록 크기의 2배에 해당하는 23 msec 이상의 지연 시간을 유발한다. 실제로 동영상에서 오디오 신호에 이와 같은 지연 시간을 주고 동영상을 관찰하면 비디오와 오디오 신호사이의 동기화 문제가 발생한다.

따라서 저지연 공간 등화 시스템을 구축하기 위해서는 적은 탭수로 등화 필터를 설계해야 한다. 본 논문에서는 실시간 3차원 공간 등화 시스템을 위한 저연산, 저지연 등화 필터 설계 알고리즘을 제안하고자 한다.

### 2.3 공통 음향 극·영점을 이용한 등화 시스템

공통 음향 극·영점 모델링 방법은 룸 임펄스 응답이나 머리전달함수 (head-related transfer function; HRTF)를 공통의 극점 (pole)을 이용하여 모델링하는 방법이다 [5]. Haneda et al.은 앞서 언급한 등화 시스템의 연산량과 입출력 지연 문제를 해결하기 위해 공간 응답의 공통

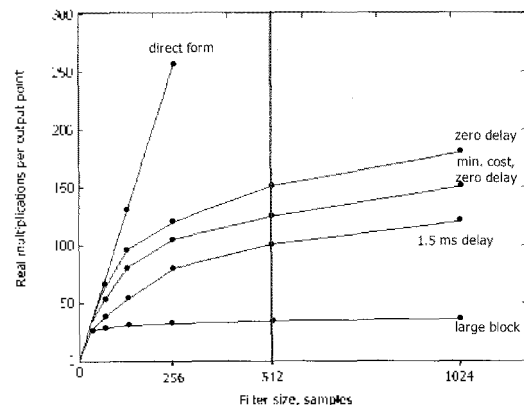


그림 2. 블록 크기에 따른 연산량 비교  
Fig. 2. Complexity comparison according to block size.

음향 극점을 구하여 등화시키는 방법을 제안하였다 [4]. 그림 3은 공통 음향 극점을 이용한 등화 시스템의 블록 다이어그램이다. 각각의 룸 전달 함수는 위치와 무관한 공통의 음향 극점 전달 함수  $A_{CAP}(z)$ 와 위치에 따라 다른 음향 영점 함수  $B_i(z)$ 로 모델링될 수 있다. 이와 같은 성질을 이용하여 공통의 음향 극점 전달 함수  $A_{CAP}(z)$ 만을 이용하여 공간 등화를 수행한다. 공통 음향 극·영점 방법은 기존의 all-zero 모델이나 pole/zero 모델에 비해 적은 파라미터의 개수를 필요로 하는 장점을 가진다. 다음은 공통 음향 극점과 영점을 구하는 방법을 설명한 것이다.

공통 음향 극·영점 모델의 임펄스 응답  $\hat{h}(\mathbf{r}_j, k)$ 는 Auto Regressive Moving Average (ARMA) 모델을 이용하여 다음과 같이 표현할 수 있다.

$$\hat{h}(\mathbf{r}_j, k) = -\sum_{i=1}^P a_i \hat{h}(\mathbf{r}_j, k-i) + \sum_{i=0}^Q b_i(\mathbf{r}_j) \delta(k-i) \quad (5)$$

여기서  $P$ 와  $Q$ 는 각각 극점과 영점의 차수를 나타내고,  $\mathbf{a}$ 와  $\mathbf{b}_i(\mathbf{r}_j)$ 는 추정된 공통 극점과 영점을 의미한다. 수식 (5)에서 볼 수 있듯이,  $\mathbf{a}$ 가 위치  $\mathbf{r}_j$ 에 따른 함수가 아님을 주목해야 한다. 원래의 임펄스 응답  $h(\mathbf{r}_j, k)$ 과 모델된 임펄스 응답  $\hat{h}(\mathbf{r}_j, k)$ 과의 오차  $\varepsilon(\mathbf{r}_j, k)$ 를 구하여 다음과 같은 비용 함수를 최소화하도록 하는 임펄스 응답을 추정할 수 있다.

$$J = \sum_{j=1}^M \sum_{k=0}^{\infty} \varepsilon^2(\mathbf{r}_j, k) \quad (6)$$

여기서,  $M$ 은 추정에 사용되는 임펄스 응답의 개수를 의미한다. 룸 임펄스 응답과 공통 음향 극·영점 방법에 의해 모델링된 임펄스 응답 사이의 오차인 "equation error"는 다음 수식과 같이 표현된다.

$$\varepsilon_{eq}(\mathbf{r}_j, k) = h(\mathbf{r}_j, k) + \sum_{i=1}^P a_i h(\mathbf{r}_j, k-i) - \sum_{i=0}^Q b_i(\mathbf{r}_j) \delta(k-i). \quad (7)$$

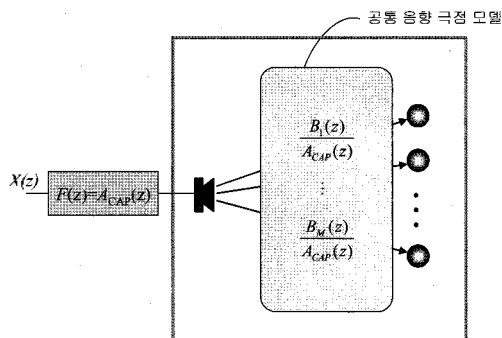


그림 3. 공통 음향 극점을 이용한 등화 시스템  
Fig. 3. Equalization system using common acoustical pole.

수식 (7)은 다음과 같은 벡터와 행렬을 정의함으로써 좀더 간단히 표현할 수 있다.

$$\mathbf{e}_j = [\varepsilon_{eq}(\mathbf{r}_j, 0) \ \varepsilon_{eq}(\mathbf{r}_j, 1) \ \dots \ \varepsilon_{eq}(\mathbf{r}_j, N+P)]^T, \quad (8)$$

$$\mathbf{h}_j = [h(\mathbf{r}_j, 0) \ h(\mathbf{r}_j, 1) \ \dots \ h(\mathbf{r}_j, N) \ 0 \ \dots \ 0]^T, \quad (9)$$

$$\mathbf{a} = [-a_1 \ -a_2 \ \dots \ -a_P]^T, \quad (10)$$

$$\mathbf{b}(\mathbf{r}_j) = [b_0(\mathbf{r}_j) \ b_1(\mathbf{r}_j) \ \dots \ b_Q(\mathbf{r}_j)]^T. \quad (11)$$

위의 벡터와 행렬을 이용하여 표현한 "equation error"는 다음과 같다.

$$\mathbf{e} = \mathbf{h} - \mathbf{A}\mathbf{x} \quad (12)$$

여기서  $\mathbf{e}$ ,  $\mathbf{h}$ , 그리고  $\mathbf{x}$ 는 다음과 같이 정의된다.

$$\mathbf{e} = [\mathbf{e}_1^T \ \mathbf{e}_2^T \ \dots \ \mathbf{e}_M^T]^T, \quad (13)$$

$$\mathbf{h} = [\mathbf{h}_1^T \ \mathbf{h}_2^T \ \dots \ \mathbf{h}_M^T]^T, \quad (14)$$

$$\mathbf{x} = [\mathbf{a}^T \ \mathbf{b}^T(\mathbf{r}_1) \ \mathbf{b}^T(\mathbf{r}_2) \ \dots \ \mathbf{b}^T(\mathbf{r}_M)]^T. \quad (15)$$

그리고  $\mathbf{A}$ 는  $\mathbf{A}_j$ 행렬과  $\mathbf{D}$  행렬을 원소로 갖는 행렬이다.

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{D} & \mathbf{0} \\ \mathbf{A}_2 & & \mathbf{D} \\ \vdots & & \\ \mathbf{A}_M & & \mathbf{D} \end{bmatrix}, \quad (16)$$

$$\mathbf{D} = \begin{bmatrix} 1 & \mathbf{0} \\ & 1 \\ & & \ddots \\ \mathbf{0} & & & 1 \\ \vdots & & & \vdots \\ \mathbf{0} & & & \vdots \end{bmatrix}_{N+P+1}, \quad (17)$$

$$\mathbf{A}_j = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ h(\mathbf{r}_j, 0) & \mathbf{0} & \dots & \mathbf{0} \\ h(\mathbf{r}_j, 1) & h(\mathbf{r}_j, 0) & \dots & \vdots \\ \vdots & h(\mathbf{r}_j, 1) & \ddots & \mathbf{0} \\ h(\mathbf{r}_j, N) & \vdots & \ddots & -h(\mathbf{r}_j, 0) \\ \mathbf{0} & h(\mathbf{r}_j, N) & \ddots & h(\mathbf{r}_j, 0) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & h(\mathbf{r}_j, 0) \end{bmatrix}_{N+P+1}. \quad (18)$$

수식 (6)의 비용함수를 앞에서 정의한 벡터  $\mathbf{e}_k$ 를 이용하여 다음과 같이 다시 쓸 수 있다.

$$\mathbf{J} = \mathbf{e}^T \mathbf{e}. \quad (19)$$

이 때 수식 (19)의 비용함수를 최소화하는 벡터  $\mathbf{x}$ 는 다음과 같다.

$$\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{h}. \quad (20)$$

$M$ 개의 룬 임펄스 응답을 공통 음향 극점과 영점을 이용하여 모델링하는 경우, 일반적인 ARMA 모델링 방법이 가지는 차수는  $M \times (P+Q+1)$ 개인데 비해 공통 음향 극·영점 모델링 방법을 이용하는 경우에는 차수를  $P+M \times (Q+1)$ 개로 줄일 수 있는 장점이 있다.

공통 음향 극점의 차수를  $P$ , 영점의 차수를  $Q$ 라 하고 룬 임펄스 응답 한 개의 길이를  $N$ 샘플이라고 가정했을 때,  $\mathbf{A}$  행렬의 크기는  $\{M(M+P+1) \times (P+M(Q+1))\}$ 이 된다. 샘플링 주파수가 8 kHz인 경우, 0.25초의 진향을 -60 dB의 정확도로 표현하기 위해서는 룬 임펄스 응답은 2000 샘플의 길이를 가져야 한다. 이와 같은 룬 임펄스 응답 10개를 공통 음향 극·영점 방법을 이용하여  $P=256$ ,  $Q=256$ 으로 모델링하기 위해서는,  $2257(0 \times 2826)$ 의 크기를 가지는  $\mathbf{A}$  행렬의  $(\mathbf{A}^T \mathbf{A})^{-1}$ 을 구하는 연산이 필요하다. 따라서, [5]의 방법은 HRTF와 같은 비교적 짧은 길이의 임펄스 응답을 모델링하는데 적절한 방법임을 알 수 있다. 이와 같은 문제를 해결하고자 averaging 방법이 제안되었다 [5]. 이 방법은 각각의 임펄스 응답에서 구한 극점을 구한 후에 평균을 취하여 공통 음향 극점을 구한다.

$$a_i = \frac{1}{M} \sum_{j=1}^M a_i(r_j) \quad (i=1, 2, \dots, P). \quad (21)$$

이 방법은 연산량 측면에서는 이득이 있으나, 정확도가 떨어지는 단점이 있다. 또한 공통 음향 극·영점 모델링 방법은 필터의 차수가 충분한 경우에는 원래의 공간 응답을 잘 모델링하지만, 부족한 경우에는 저주파 영역에서의 모델링 성능이 떨어지는 문제점이 있다. 이것은 공통 음향 극·영점 모델링 방법의 특성상 차수가 부족한 경우에 Q-factor가 높은 주파수 대역을 먼저 모델링하기 때문이다.

### III. 제안한 공간 등화 알고리즘

본 장에서는 앞 장에서 언급한 공통 음향 극·영점 방법의

두 가지 문제점을 해결할 수 있는 적응 구조를 제안한다. 또한 저주파 영역의 공간 등화 성능을 높이기 위해 제안한 적응구조와 frequency-warping 기술을 결합한다.

### 3.1. Iterative algorithm

그림 4는 적응 IIR (Infinite Impulse Response) 필터를 이용하여 공통의 음향 극점과 영점을 구하기 위해 제안된 구조의 블록 다이어그램을 도시한 것이다. 블록 다이어그램에서 볼 수 있듯이 영점을 추정하는 부분은 임펄스 응답의 개수만큼 있으며 극점을 추정하는 부분은 공통적으로 한 개가 있다. 수식 (19)와 같은 비용함수를 최소화하는 적응 계수를 구하기 위해 gradient를 기반으로 한 steepest descent 방법을 이용하였다. 그림 4의 시스템에서 오차는 수식 (22)와 같으며, 비용함수는 수식 (23)과 같다.

$$\mathbf{e}_i = \mathbf{h}_i - [\mathbf{A}_i \mathbf{a}_i + \mathbf{D} \mathbf{b}_i]. \quad (22)$$

$$J = \mathbf{e}^T \mathbf{e} = \mathbf{e}_1^T \mathbf{e}_1 + \mathbf{e}_2^T \mathbf{e}_2 + \dots + \mathbf{e}_M^T \mathbf{e}_M. \quad (23)$$

이와 같은 비용함수를 최소화하는 공통 음향 극점과 영점의 계수는 수식 (23)의 비용함수를 각각의 계수로 미분함으로써 구할 수 있다.

$$\nabla_{\mathbf{a}_i} J = -2 \mathbf{D}^T \mathbf{e}_i, \quad i=1, 2, \dots, M. \quad (24)$$

$$\nabla_{\mathbf{b}_i} J = -2 \sum_{i=1}^M \mathbf{A}_i^T \mathbf{e}_i. \quad (25)$$

수식 (24)와 (25)의 미분치를 steepest descent 방법에 대입하면 다음과 같은 적응식을 얻게 된다.

$$\hat{\mathbf{b}}_i(m+1) = \hat{\mathbf{b}}_i(m) - \frac{\mu}{2} (\nabla_{\mathbf{b}_i} J) - \hat{\mathbf{b}}_i(m) + \mu \mathbf{D}^T \hat{\mathbf{e}}_i(m), \quad i=1, 2, \dots, M, \quad (26)$$

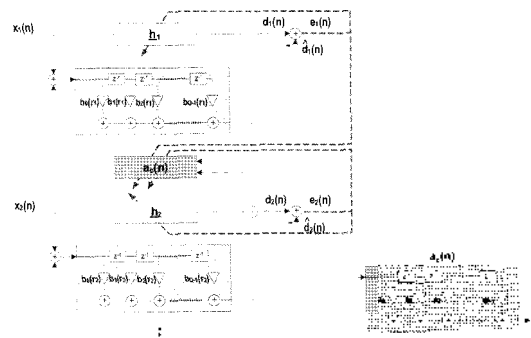


그림 4. 적응 CAPZ 추정 구조  
Fig. 4. Structure of the adaptive CAPZ estimator.

$$\hat{\mathbf{a}}_c(m+1) = \hat{\mathbf{a}}_c(m) - \frac{\mu}{2} (\nabla_{\mathbf{a}_c} J(m)) = \hat{\mathbf{a}}_c(m) + \mu \sum_{i=1}^M \mathbf{A}_i^T \hat{\mathbf{e}}_i(m), \quad (27)$$

여기서,  $\hat{\mathbf{b}}_i(m)$ 과  $\hat{\mathbf{a}}_c(m)$ 는 각각  $m$  번째 반복에서 추정된 공통 음향 극·영점 계수 벡터를 의미하고,  $\hat{\mathbf{e}}_i(m) = \mathbf{h}_i - [\mathbf{A}_i \hat{\mathbf{a}}_c(m) + \mathbf{D} \hat{\mathbf{b}}_i(m)]$ 은  $m$  번째 반복에서 추정된 오차들의 의미한다.

앞 절에서 제안한 적응 구조는 공통의 음향 극점과 영점을 적은 연산량으로 구하는 좋은 방법이 된다. 하지만, 기본적으로 공통 음향 극·영점 방법이 가지고 있는 문제점을 해결하지는 못한다. 기본적으로 공통 음향 극·영점 방법은 적은 차수로 공통의 음향 극점과 영점을 구하는 경우에 Q-factor가 높은 공통의 음향 극점과 영점을 찾는다. 이것은 비유합수를 자승 오차의 합으로 두었기 때문에 발생하는 현상이다. 효율적인 등화를 위해서는 청취에 중요한 저주파 영역에서 정확한 등화가 필요하므로 저주파 영역에서 기존의 방법보다 좀 더 정확한 공간 응답 모델링 방법이 필요하다.

### 3.2. Frequency-Warped Common Acoustical Pole and Zero (WCAPZ)

많은 음성/오디오 연구분야에서 특정 주파수 영역의 해상도를 높이기 위해 주파수 해상도를 원하는 형태로 조절할 수 있는 주파수 워핑 (frequency warping) 기술을 사용한다 [6]. 주파수 워핑은 유닛 지연 소자를 all-pass 소자로 대체함으로써 구현될 수 있다.

$$D(z) = \frac{z^{-1} - \lambda}{1 - \lambda z^{-1}} \quad (28)$$

여기서  $\lambda$ 는 -1과 1사이의 범위를 가지는 워핑 파라미터이다. 양수의  $\lambda$ 값은 저주파 영역의 해상도를 높여주는

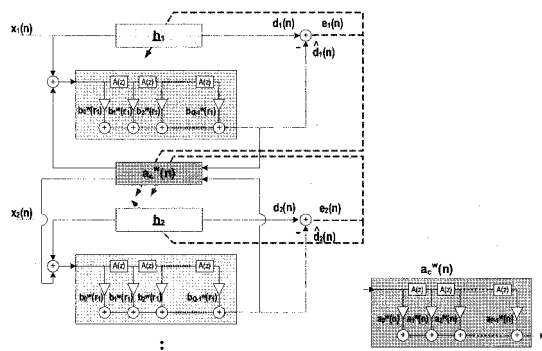


그림 5. 적응 WCAPZ 추정 구조  
Fig. 5. Adaptive estimator for the warped CAPZ.

반면, 음수의  $\lambda$ 값은 고주파 영역의 해상도를 높여주는 역할을 한다. 이와 같은 warped 필터와 공통 음향 극·영점 모델링 방법을 결합한 것을 warped common acoustical pole and zero (WCAPZ) 방법이라 하고, 그림 5에 블록 다이어그램을 도시하였다. 그림 5의 블록 다이어그램에서 구한 공통 음향 극점과 영점은 warping 영역에서 최소의 오차를 가지도록 하는 최적의 계수들을 구해낼 것이며, 이것은 앞서 언급한 적은 차수로 공통 음향 극점과 영점을 찾을 때 발생하는 문제점을 해결하여 저주파 영역에서 좋은 모델링 성능을 보여주게 된다.

일반적으로 warped 필터의 구현은 그것의 all-pass 소자 때문에 FIR 필터 구현보다 많은 연산량이 필요로 하게 된다. 따라서 warping 영역의 계수를 선형 영역으로 매핑시켜 IIR 필터 형태로 구현하게 된다. Warping 영역의 계수로부터 선형 영역의 계수로 변환하는 방법은 [6]에 기술되어 있다.

### 3.3. WCAPZ 방법을 이용한 공간 등화 시스템

그림 6은 WCAPZ를 이용한 공간 등화 시스템의 블록 다이어그램이다. 먼저 WCAPZ 모델링 방법을 이용하여 공간 응답 함수를 모델링하고 MA부분에 해당하는 역 필터를 설계한다. 그리고 Auto Regressive (AR) 부분에 해당하는 계수를 가지고 전체 입력 채널을 필터링한 후, 마지막으로 역 필터 계수로 입력 채널을 필터링한다. 여기서 등화 필터의 설계는 주파수 영역의 Fast deconvolution 설계 방법을 이용하였다 [7].

제안된 공간 등화 시스템은 WCAPZ 방법으로 모델링된 공간 전달 함수를 가지고 공간 등화 필터를 구하기 때문에 적은 차수의 등화 필터로 저주파 영역의 성능을 향상시킬 수 있는 장점이 있다. 또한 제안한 알고리즘은 CAPZ에 기반한 등화 방법과 유사한 연산량을 가지면서

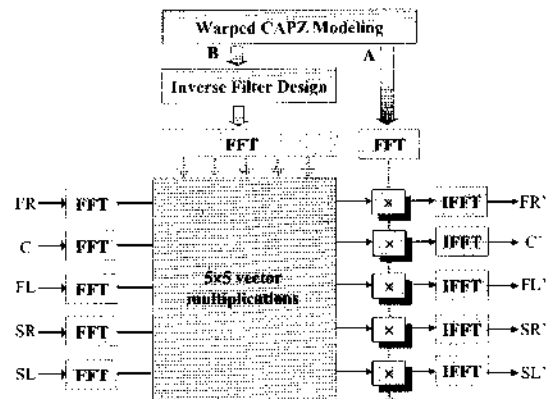


그림 6. WCAPZ를 이용한 공간 등화기  
Fig. 6. Room Equalizer Based on Warped CAPZ.

도 입출력 지연은 적은 장점을 가짐을 알 수 있다.

### IV. 모의 실험

본 논문에서 제안한 등화 시스템의 성능을 평가하기 위해 all-zero 모델을 이용한 등화 방법, CAPZ를 이용한 등화 방법, 그리고 WCAPZ를 이용한 등화 방법의 등화 성능을 비교하였다. 이 때 all-zero 모델은 512차로, CAPZ와 WCAPZ를 이용한 등화 방법의 극점과 영점의 차수는 128로 고정하였다. 제안된 방법에서 워핑 매개 변수  $\lambda$ 의 값이 클수록 저주파 영역의 해상도가 증가하여 저주파 영역의 성능이 향상된다. 하지만  $\lambda$ 값이 크면 많은 수의 극점과 영점들이 저주파 영역에 집중이 되고 이로 인해 warping 영역의 계수를 선형 영역의 계수로 변환하는 과정이 불안정해지게 된다. 이와 같은 성능과 안정성의 tradeoff 를 고려하여  $\lambda$ 값을 0.4로 결정하였다. 실험에 필요한 공간 전달함수는  $6 \times 4 \times 3$  m크기의 실험실에 ITU-T 권고안 [8]에 따라 스피커를 위치시킨 후, Image method [9]를 이용하여 512차의 all-zero 모델로 생성하였다. 복귀 응답은 MIT HRTF DB로부터 추출하여 사용하였다. 예를 들면 전방 좌채널로부터 왼쪽 마이크로로 도달하는 복귀 응답은 305의 HRTF를 이용하였다.

먼저 WCAPZ의 모델링 성능을 검증하기 위해 CAPZ 방법과 모델링 성능을 비교하였다. 그림 7은 생상한 룸 임펄스 응답 중 전방 좌채널 스피커로부터 왼쪽 마이크로

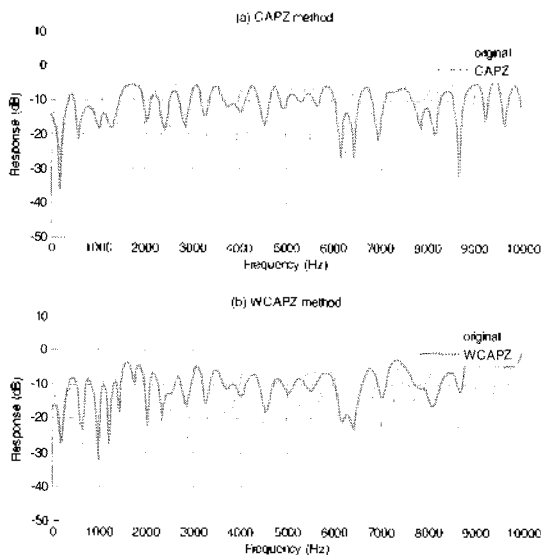


그림 7. 룸 임펄스 응답의 모델링 성능 (a) CAPZ method (b) WCAPZ method  
Fig. 7. Room transfer function modeling (a) CAPZ method (b) WCAPZ method.

폰으로 도달하는 룸 임펄스 응답을 모델링한 결과를 비교한 것이다. 그림에서 보듯이 WCAPZ 방법이 같은 차수를 사용한 CAPZ 방법보다 저주파 영역에서의 모델링 성능이 좋음을 확인할 수 있다. 이것은 CAPZ 방법의 경우 극점과 영점이 선 대역에 골고루 분포하는데 비해, WCAPZ 방법은 워핑 영역에서의 계수들을 선형 영역으로 변환하는 과정 중에 극점과 영점이 저주파 영역에 많이 분포하여 상대적으로 저주파 영역의 해상도가 증가하기 때문이다. 그림 8은 CAPZ와 WCAPZ 방법의 극·영점의 분포도를 나타낸 것이며, WCAPZ의 극·영점이 저주파 영역에 많이 분포하고 있음을 보여준다.

다음으로 전체 시스템의 등화 성능을 비교하였다. 그림 9는 all-zero 모델 방법에 기반한 등화 시스템과 CAPZ 방법에 기반한 등화 시스템의 전체 주파수 응답을 그린 것이고, 그림 10은 all-zero 모델 방법과 WCAPZ 방법에 기반한 등화 시스템의 전체 주파수 응답을 그린 것이다. 그림 9에서 CAPZ 모델과 all-zero 모델의 주파수 응답의 차이가 전방 좌채널 스피커에서 왼쪽 귀로 도달하는 경우에는 약 5dB, 전방 우채널 스피커에서 오른쪽 귀로 도달하는 경우에는 약 20dB의 차이를 보였다. 하지만 그림 10의 WCAPZ 모델은 전체적으로 고르게 all-zero 모델과 비슷한 성능을 보임을 확인할 수 있다. 이것은 동일한 차수를 사용하는 두 시스템을 비교할 때, WCAPZ 방법에 기반한 등화 시스템이 CAPZ 방법에 기반한 시스템보다 1kHz 이하의 저주파 영역에서 좋은 성능을 보여주는 것을 의미한다. 따라서 제안한 시스템은 512차를 이용하는 all-zero 모델 기반의 등화 시스템의 1/2에 해당하는 차수만으로 설계되었지만, 1kHz 이하의 저주파 대역의 등화 성능이 512차의 all-zero 기반의 등화 시스템과 비슷

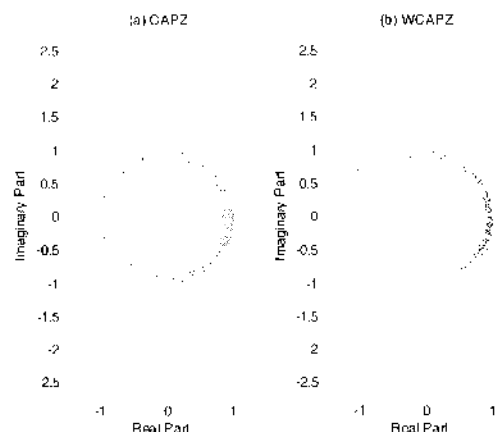


그림 8. 극·영점의 분포도 (a) CAPZ 방법 (b) 제안된 방법  
Fig. 8. Pole/zero diagram (a) CAPZ method (b) WCAPZ method.

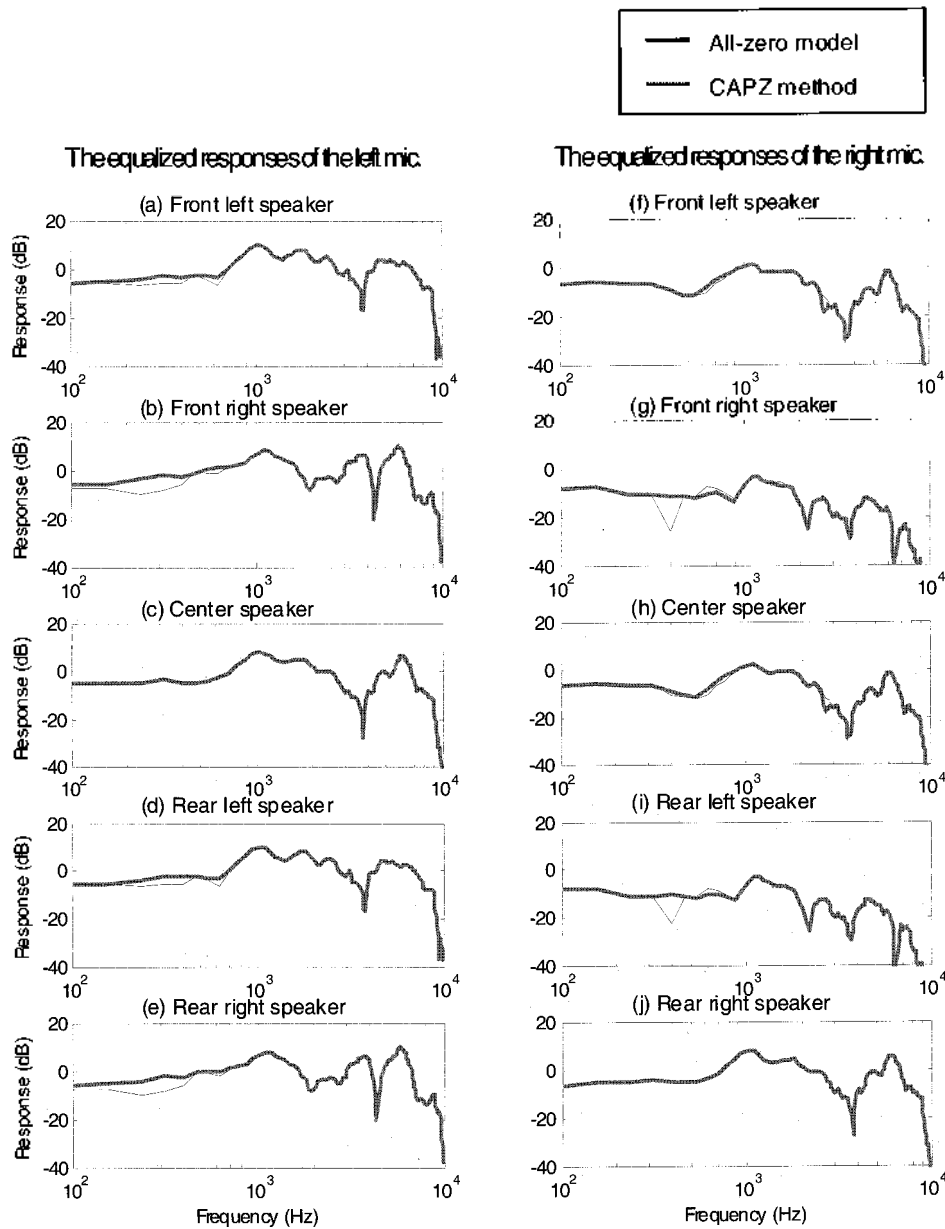


그림 9. All-zero 모델 방법과 CAPZ 방법에 기반한 공간 등화 시스템의 주파수 응답:

(a) 전방 좌채널 스피커 (b) 전방 우채널 스피커 (c) 센터 채널 스피커 (d) 후방 좌채널 스피커 (e) 후방 우채널 스피커 와 왼쪽 마이크로폰 사이의 주파수 응답, (f) 전방 좌채널 스피커 (g) 전방 우채널 스피커 (h) 센터 채널 스피커 (i) 후방 좌채널 스피커 (j) 후방 우채널 스피커 와 오른쪽 마이크로폰 사이의 주파수 응답.

Fig. 9. Frequency response of room equalization system based on all-zero modeling method and CAPZ method:

The frequency response of between (a) Front left speaker (b) Front right speaker (c) Center speaker (d) Rear left speaker (e) Rear right speaker and left microphone, The frequency response of between (a) Front left speaker (b) Front right speaker (c) Center speaker (d) Rear left speaker (e) Rear right speaker and right microphone.

한 성능을 가지며, warping을 이용하지 않은 선형 주파수 영역의 CAPZ 방법에 기반한 등화 방법보다 좋은 성능을 가진다. 비록 제한한 방법이 고주파 대역의 성능 감소를 야기하지만, 오디오 신호의 대부분의 에너지가 저주파 대역에 분포 하는 것을 감안한다면 고주파 대역의 성능 감소가 전체 등화 성능에 미치는 영향은 작다.

마지막으로 10명의 오디오 전문가를 대상으로 주관적

음질 평가를 수행하였다. 각 피실험자들은 all-zero 모델 방법, CAPZ 방법, 그리고 WCAPZ에 기반한 등화 시스템의 출력 신호를 듣고 음질 평가를 하였다. 이 때, 512차의 등화 필터를 사용한 기존 방법을 레퍼런스 시스템으로 놓고 128차의 극점과 영점을 사용하는 CAPZ, WCAPZ 방법과의 등화 성능을 비교하여 5점부터 1점까지로 평가하도록 하였다. 테스트 음원은 5.1채널 테스트 DVD인 "The



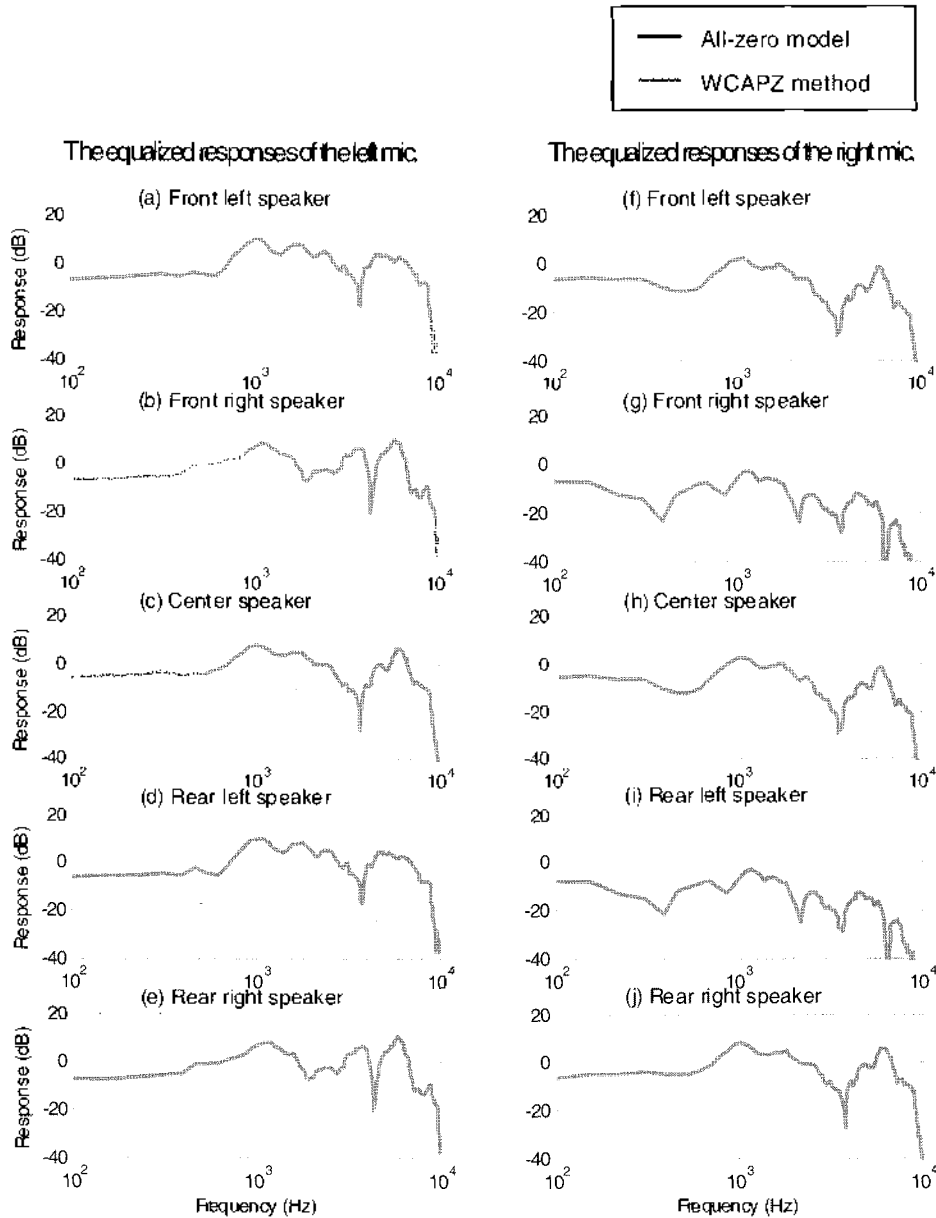


그림 10. All-zero 모델 방법과 WCAPZ 방법에 기반한 공간 등화 시스템의 주파수 응답

(a) 전방 좌채널 스피커 (b) 전방 우채널 스피커 (c) 센터 채널 스피커 (d) 후방 좌채널 스피커 (e) 후방 우채널 스피커 와 왼쪽 마이크로폰 사이의 주파수 응답, (f) 전방 좌채널 스피커 (g) 전방 우채널 스피커 (h) 센터 채널 스피커 (i) 후방 좌채널 스피커 (j) 후방 우채널 스피커 와 오른쪽 마이크로폰 사이의 주파수 응답.

Fig. 10. Frequency response of room equalization system based on all-zero modeling method and WCAPZ method:

The frequency response of between (a) Front left speaker (b) Front right speaker (c) Center speaker (d) Rear left speaker (e) Rear right speaker and left microphone, The frequency response of between (a) Front left speaker (b) Front right speaker (c) Center speaker (d) Rear left speaker (e) Rear right speaker and right microphone.

표 1. 테스트 음원

Table 1. Test materials.

Abbreviation	Artist	Title
SRC1	Ana Caram	Telephone
SRC2	Richard Wagner	Tannhauser
SRC3	The Conga Kings	Tumbao de Tamborito

ultimate DVD<sup>®</sup>에서 5.1채널 음원 3곡을 선택하였다. 표 1은 선택한 음원을 정리한 것이다. 그림 11에서 볼 수 있듯이 저주파 대역의 에너지가 부족한 SRC3 음원을 제외하고는 WCAPZ에 기반한 등화 시스템의 등화 성능이 512차의 all-zero 모델에 기반한 등화 방법과 유사한 성능을 보임을 확인할 수 있다.

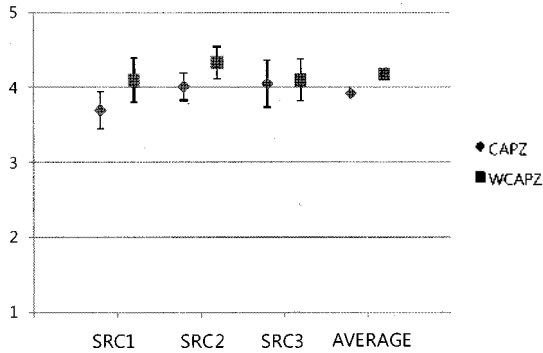


그림 11. 주관적 음질 평가 결과  
Fig. 11. Subjective quality test results.

### V. 결론

본 논문에서는 실시간 등화 시스템을 위한 저연산, 저지연 공간 등화 시스템을 제안하였다. 이를 위해 warped 필터와 기존의 공통 음향 극·영점 모델 방법을 결합한 WCAPZ 방법을 제안하였다. 제안된 WCAPZ 방법을 기반으로 하는 공간 등화 시스템은 CAPZ를 기반으로 하는 공간 등화 시스템에 비해 청취에 중요한 저주파 영역의 등화 성능이 뛰어나고, 기존의 all-zero 모델을 기반으로 하는 공간 등화 시스템의 1/2에 해당하는 차수로만으로도 기존의 등화 시스템과 유사한 성능을 가진다. 실험을 통해 제안된 방법이 적은 차수로 저주파 영역의 공간 등화 성능을 유지하는 효율적인 알고리즘임을 검증하였다.

### 참고문헌

1. P. A. Nelson, F. O. Bustamante, and H. Hamada, "Inverse filter design and equalization zones in multichannel sound reproduction," *IEEE Trans. On Speech and Audio Proc.*, 3(3), 185-192, May 1995.
2. A. V. Oppenheim and R. W. Schaler, *Digital Signal Processing*, (Prentice-Hall, Englewood Cliffs, NJ, 1975), Chap. 8, pp.576-588.
3. W. G. Gardner, "Efficient convolution without input-output delay," *J. Audio Eng. Soc.*, 43(3), 127-136, Mar. 1995.
4. Y. Haneda, S. Makino, and Y. Kaneda, "Multiple-Point equalization of room transfer functions by using common acoustical poles," *IEEE Trans. On Speech and Audio Proc.*, 5(4), 325-333, Jul. 1997.
5. Y. Haneda, S. Makino, and Y. Kaneda, "Common acoustical pole and zero modeling of room transfer functions," *IEEE Trans. On Speech and Audio Proc.*, 2(2), 320-328, Apr. 1994.
6. A. Härmä, M. Karjalainen, L. Savioja, V. Välimäki, U. K. Laine, and J. Huopaniemi, "Frequency-warped signal processing for audio applications," *J. Audio Eng. Soc.*, 48(11),

- 1011-1031, Nov. 2000.
7. O. Kirkeby, P. A. Nelson, H. Hamada, and F. Orduna-Bustamante, "Fast deconvolution of multichannel systems using regularization," *IEEE Trans. On Speech and Audio Proc.*, 6(2), 189-195, Mar. 1998.
8. ITU-R BS 775-1 3/2 Surround Format.
9. J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small room acoustics," *J. Acoust. Soc. Am.*, 65(4), 943-950, Apr. 1979.

### 저자 약력

•이 준 호 (Junho Lee)

2001년 2월: 숭실대학교 정보통신전자공학과 공학사  
2003년 2월: 연세대학교 전기전자공학부 공학 석사  
2003년 3월~ 현재: 연세대학교 전기전자공학부 박사과정  
\*주요연구분야: 디지털 신호처리, 오디오 신호처리, 적응 신호처리, 마이크로폰 어레이

•박 영 철 (Young-cheol Park)

1986년 2월: 연세대학교 전자공학과 (공학사)  
1988년 2월: 연세대학교 전자공학과 (공학 석사)  
1993년 2월: 연세대학교 전자공학과 (공학 박사)  
현재: 연세대학교 컴퓨터정보통신공학부 교수  
\*주요연구분야: 디지털 신호처리, 오디오 신호처리, 음성 신호처리, 적응 신호처리

•윤 대 희 (Dae-hee Yoon)

1977년 2월: 연세대학교 전자공학과 (공학사)  
1979년 8월: Kansas State Univ. (공학 석사)  
1982년 2월: Kansas State Univ. (공학 박사)  
현재: 연세대학교 전기전자공학부 교수  
\*주요연구분야: 디지털 신호처리, 적응 신호처리, 음성 신호처리, 오디오 신호처리

•이 석 필 (Seokpil Lee)

1990년 2월: 연세대학교 전기공학과 (공학사)  
1992년 8월: 연세대학교 전기공학과 (공학 석사)  
1997년 8월: 연세대학교 전기공학과 (공학 박사)  
현재: 전자부품연구원 방송통신융합연구센터 센터장  
\*주요연구분야: 차세대디지털방송, 디지털 신호처리, 메타데이터 처리