

MINT법을 이용한 실내 잔향곡선의 초기감쇠시간 단축에 관한 연구

A Study on the to Shorten of Early Decay Time in the Reverberation Curve Using MINT

차 경 환*
(Kyung-Hwan Cha*)

동서대학교 인터넷공학부*
(접수일자: 2001년 11월 5일; 채택일자: 2001년 11월 21일)

본 논문에서는 실내 잔향곡선의 초기감쇠시간을 복수 채널을 사용하여 단축시켰다. 실내에서 수음된 음성신호는 MINT (Multiple - input / output INverse Theorem)에 기초하여 풀밴드와 서브밴드로 역필터링하였으며, 복수 채널의 적응필터는 LMS (least Mean Square)와 NLMS (Normalized Least Mean Square) 알고리즘을 사용하였다. 실험결과 잔향곡선에서 -20 dB까지 감소하는데 걸리는 시간이 2개의 마이크로폰을 사용하였을 때 풀밴드 NLMS에서 약 1/3로 단축됨을 확인하였다. 실제 음장의 전달함수를 사용한 주관평가에서는 80%의 청취자가 MINT에 의해 초기감쇠시간이 단축된 음성의 명료도가 향상되었다고 평가하였다.

핵심용어: 잔향곡선, 초기감쇠시간, 역필터링, 음성의 명료도

투고분야: 전기음향 분야 (3, 4)

In this paper, we made shorter EDT (early decay time) of room reverberation curve using multiple-channel. The speech signal was processed inverse filtering with full-band and sub-band in the basis MINT, and then the multiple-channel adaptive filters were used LMS (Least Mean Square) and NLMS (Normalized Least Mean Square) algorithm. Experimental results, we could get 1/3 of time reduction at 20dB level in the reverberation curve using full-band NLMS when two microphones were used. Also, it is shown that the speech articulation was improved 80% from the test listeners with the speech, which was to shorten EDT by MINT in the subjective assessments using real room impulse response.

Keywords: Reverberation curve, Early decay time, Inverse filtering, Speech articulation

ASK subject classification: Electro-acoustic (3, 4)

I. 서론

실내에서 소리를 발생시키면 소리가 벽, 천정 그리고 바닥 등으로부터 반사음이 서로 중첩되어 아주 복잡한 음장이 형성된다. 이러한 반사음을 시간축 상에서 보면, 음원 발생 후에 청취자에게 제일 먼저 직접음이 도래하

고, 계속해서 천정과 벽면으로부터 1차 반사음, 2차 반사음 등이 도래한다. 직접음 도래 후 도달하는 반사음을 초기 반사음 (early reflections)이라고 하고, 초기 반사음 뒤에 도래하는 음을 잔향음 (reverberation)이라고 한다. 잔향음은 공간에서 발생된 공간 왜곡 (spatial distortion)이며, 이 왜곡의 효과가 적당하면 콘서트 홀과 같이 실내 음향상태는 울림이 적당하지만 음성의 명료도 (articulation)를 저하시키는 문제점을 가지고 있다.

실내에서 수음되는 음향신호에서 잔향을 제거하는 연

책임저자: 차경환 (khcha@dongseo.ac.kr)
617-716 부산광역시 사상구 주례동 산 69-1
동서대학교 인터넷공학부
(전화: 051-320-1759; 팩스: 051-327-8955)

구는 정보통신망에서 확성 통화 시스템의 통화품질 향상, 음성의 명료도 향상, 실감통신 등 다양한 분야에서 그 응용이 기대되고 있다. 특히 음성의 명료도는 발성자에서 수음자까지의 거리와 잔향시간보다는 잔향곡선의 감쇠 기울기에 따라서 크게 차이가 난다. 즉 음압 레벨이 -10 dB 떨어질 때까지의 시간인 초기감쇠시간 (early decay time)이 짧을수록 음성의 명료도가 향상된다고 한다 [1-3].

실내의 잔향시간을 변화시키는 방법은 크게 건축 구조체에 의한 방법과 전기 음향학적인 방법으로 나눌 수 있다. 특히 전기 음향에 의한 능동적인 방법으로 잔향에 의한 왜곡을 저감시키는 종래의 연구방법은 그림 1과 같이 실내 전달함수의 역함수를 이용하는 역필터링 방법이다. 그러나 실내 전달함수는 일반적으로 최소 위상 함수 (minimum phase function)로 고려되지 않기 때문에 단일채널에 의해 역함수를 정확히 구하기가 어려워서 시스템은 불안정하다[4,5].

본 논문에서는 실내에서 발생된 음성의 명료도를 향상시키기 위해서 실내 잔향곡선의 초기감쇠시간을 MINT (Multiple-input/output INverse Theorem)법을 이용하여 단축시켰다. MINT법을 이용한 잔향저감 시스템은 풀밴드 (full-band)와 서브밴드 (sub-band)로 모델링하고, LMS (Least Mean Square)와 NLMS (Normalized Least Mean Square)알고리즘으로 잔향곡선의 초기감쇠시간을 비교 검토하였다. 그리고 음성의 명료도 평가는 실제 음장의 전달함수를 사용하여 주관평가하였다.

II. MINT법에 의한 초기감쇠시간 단축

스피커에서 마이크로폰까지의 실내 전달함수가 인과적이면서 안정하고 그 역시스템도 인과적이고 안정할 필

요 충분조건은 $H(z)$ 의 극점과 영점이 단위원 내에 있을 때이다. 일반적인 유리 시스템 함수는 다음과 같이 표현될 수 있다.

$$H(z) = H_{\min}(z)H_{\text{ap}}(z) \tag{1}$$

윗식에서 $H_{\min}(z)$ 는 단위원 내에 있는 $H(z)$ 의 극점과 영점을 포함하고 단위원 밖에 있는 $H(z)$ 영점의 켈레 역수 영점 (conjugate reciprocal zero)도 포함한다. $H_{\text{ap}}(z)$ 는 $H_{\min}(z)$ 에서 반사 켈레 역수 영점을 상쇄하기 위한 극점과 함께 단위원 외부에 $H(z)$ 의 모든 영점을 포함한다. 식 (1)을 이용하여, 단위원 밖의 모든 영점을 단위원 내의 켈레 역수 위치로 반사시킴으로써 비최소 위상 시스템 (non-minimum phase system)을 최소 위상 시스템 (minimum phase system)으로 변형할 수 있다.

MINT법은 단일 채널에 의해 정확히 구할 수 없는 역함수를 복수개의 채널을 이용하여 최소 위상 함수로 변형시켜 안정한 역함수를 구할 수가 있다. 그림 2는 2개의 채널을 이용한 잔향 저감 시스템을 나타내었다. 음원 $S(n)$ 은 실내의 전달경로가 다른 $z^{-d}h_1(z)$ 와 $z^{-(d+d')}h_2(z)$ 에 의해 잔향이 포함된 상태로 Mic. 1과 Mic. 2에 입력된다. 2개의 마이크로폰 출력은 τ 만큼 지연된 후 역필터링을 위한 적응필터인 $g_1(z)$, $g_2(z)$ 필터에 의해 처리된다. 출력신호 $y(n)$ 은 식 (2)와 같이 표현된다.

$$y(n) = h_1(z)s(n-d) - z^{-\tau} \{h_1(z)g_1(z) + z^{-d'}h_2(z)g_2(z)\}s(n-d) \tag{2}$$

여기서 $h_1''(z)$ 를 식 (3)과 같이, $h_1(z)$ 를 식 (4)와 같이 두고, 식 (2)에 식 (3)과 식 (4)를 대입하면 식 (5)와 같이 된다.

$$h_1''(z) = h_1(z)g_1(z) + z^{-d'}h_2(z)g_2(z) \tag{3}$$

$$h_1(z) = h_1'(z) + z^{-\tau}h_1''(z) \tag{4}$$

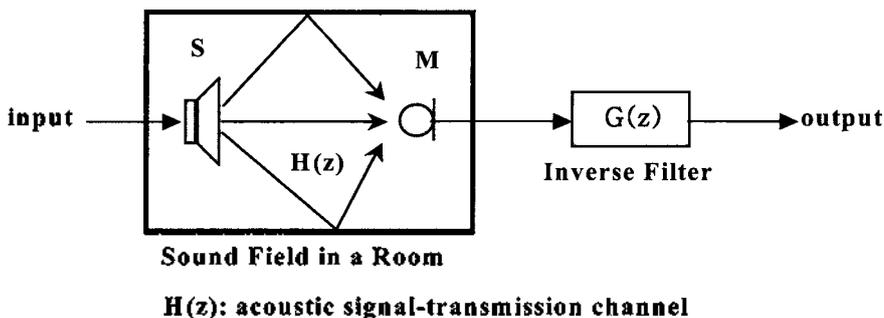


그림 1. 역필터링 방법에 의한 잔향 저감
 Fig. 1. Reduction of reverberation using inverse filtering.

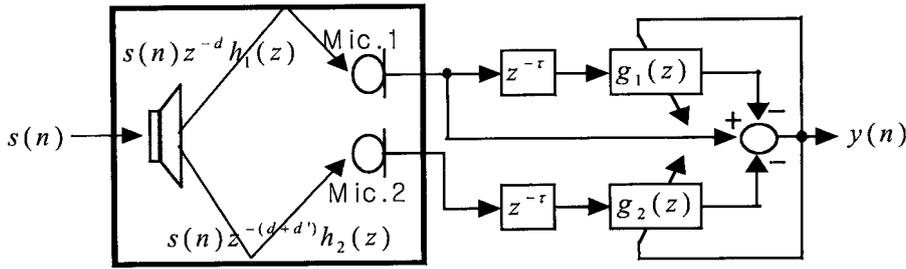


그림 2. MINT법에 의해 2개의 마이크로폰을 이용한 잔향 저감 시스템
 Fig. 2. Reverberation reduction system using 2 microphones by MINT.

$$y(n) = h_1'(z)s(n-d) \tag{5}$$

식 (2)를 만족하는 필터 $g_1(z), g_2(z)$ 를 구하는 것이 가능하면, 음원과 Mic.1사이의 실내 전달함수 $h_1(z)$ 의 τ 차 이하의 잔향 저감이 가능하다. 적응필터 $g_1(z), g_2(z)$ 는 식 (6)과 같이 2승 평균치를 최소화하는 적응처리에 의해서 표현이 가능하다.

$$G_{opt} = \frac{H^T E[W_{n-\tau} W_{n-\tau}^T] h'' + H^T E[W_{n-\tau} W_n^T] h'}{H^T E[W_{n-\tau} W_{n-\tau}^T] H + \sigma^2 I} \tag{6}$$

- G_{opt} : $g_1(z), g_2(z)$ 의 계수로 구성된 벡터
- H : $h_1(z), z^{-d'} h_2(z)$ 의 계수로 구성된 convolution행렬
- W_n : $s(n-d)$ 로 구성된 벡터
- h', h'' : $h_1'(z), h_1''(z)$ 의 계수로 구성된 벡터
- $E[\]$: 기대치

III. 시뮬레이션

시뮬레이션에 사용된 음원과 마이크로폰 사이의 실내 전달함수는 크기가 76 m³인 직육방체실에서 백색잡음에 의한 크로스 스펙트럼법을 이용하여 측정하였다. 그림 3은 스피커와 마이크로폰 사이의 거리를 2 m로 하였을 때 측정된 실내 전달함수를 나타내고 있다. 샘플링 주파수는 12 kHz이고 정밀도는 16비트이며 2048포인트까지 계산하였다. 음원신호는 백색잡음을 이용하여 $\sigma = 0, \tau = 1$ 로 해서 18,000 회 갱신하였다. 잔향 저감 시스템의 적응필터를 풀밴드와 서브밴드 그리고 적응알고리즘을 LMS와 NLMS로 하였을 때 잔향곡선의 초기감쇠시간 단축은 그림 2의 Mic.1 마이크로폰을 기준으로 해서 비교하였다.

잔향곡선에서 초기감쇠시간의 단축에 따른 음성의 명료도 평가는 20대 남성 화자가 발성한 /안녕하십니까/

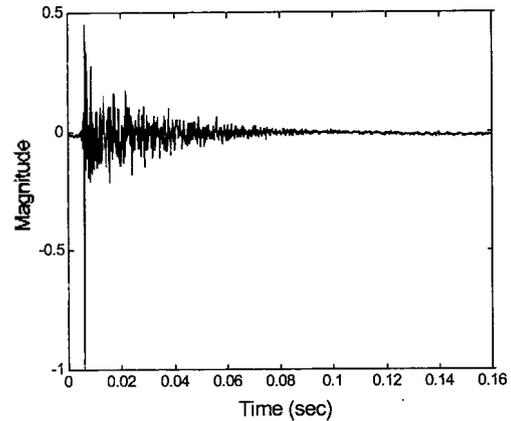


그림 3. 실음장에서 측정된 음원과 마이크로폰 사이의 실내전달 함수
 Fig. 3. Room impulse response from microphone to sound source in the real sound field.

라는 문장과 /아, 야, 어, 여, 오, 요, 우, 유, 으, 이/ 음성을 반무향실에서 녹음하여 사용하였다. 주관 평가방법은 평가 실험을 하는 실험자도 제시된 시료를 알지 못하는 “double blind법”을 이용하였다[5]. 청취자는 정상적인 청각을 갖고 있는 20대 대학생 남성 5명과 여성 5명이며 헤드폰으로 평가하였다. 실험에 사용된 세가지 음성 A,B,C는 다음과 같다. A는 반무향실에서 녹음한 잔향이 없는 source음성이며, B는 그림 3의 실내 전달함수에 의해 잔향이 포함된 음성이고, C는 풀밴드 NLMS에 의해 초기감쇠시간을 단축시켜 처리한 음성이다. 청취자는 불규칙하게 배열된 A,B,C 음성 중에서 가장 명료도가 좋은 차례로 1, 2, 3등급을 부여하도록 하였다.

그림 4는 2개의 마이크로폰을 사용하여 풀밴드의 LMS와 NLMS 그리고 8 서브밴드의 NLMS에 의해서 잔향이 저감된 실내 전달함수의 잔향곡선을 나타내었다. 잔향에너지가 -20 dB까지 감소하는데 걸리는 시간을 비교하면 측정된 실음장에서는 약 160 msec였지만, 풀밴드 LMS에서는 약 60 msec, 풀밴드 NLMS에서는 약 50 msec로 단축되었다. 8 서브밴드 NLMS에서는 풀밴드에 비해서 약 80

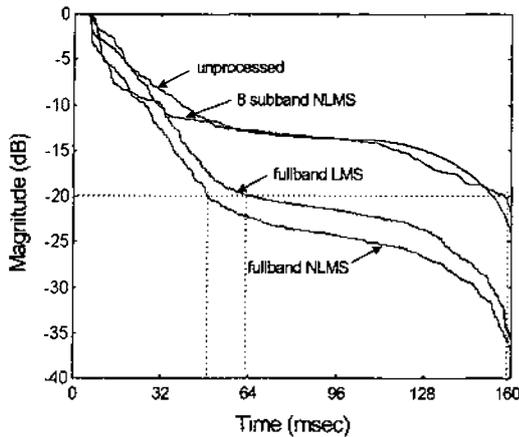


그림 4. 2개의 마이크론을 사용한 잔향곡선
Fig. 4. Reverberation curve using two microphones.

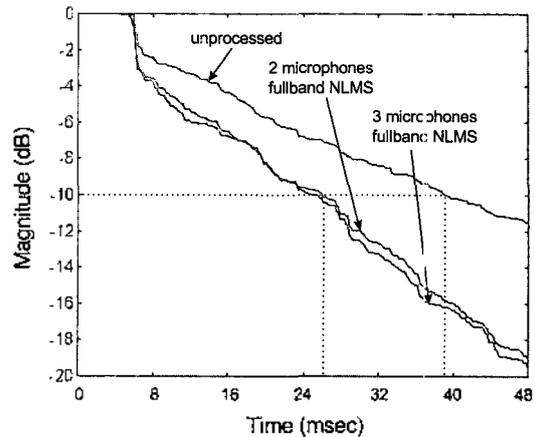


그림 5. 풀밴드 NLMS에서 마이크론의 개수에 따른 초기감쇠 시간
Fig. 5. Early Decay Time according to microphone number in the full-band NLMS.

%정도의 연산량을 감소시킬 수는 있었으나 잔향 저감 효과가 크지 않았다. 따라서 풀밴드 NLMS방법이 잔향 저감 효과가 우수한 것으로 확인되었다.

그림 5는 풀밴드 NLMS인 경우에 마이크론의 개수에 따른 잔향 저감을 나타내었다. Mic.1과 Mic.2사이의 거리는 0.2m이고, Mic.1과 Mic.3, Mic.2와 Mic.3사이의 거리는 0.14 m로 하였다. 실험결과에서는 마이크론을 2개와 3개 사용하였을 경우 큰 차이를 나타내지는 못하였다. 잔향에너지가 -10 dB까지 감소하는데 걸리는 초기감쇠 시간 (EDT)을 비교하면 실음장에서는 약 39 msec였으며 풀밴드 NLMS에서는 약 26 msec로 단축되었다.

그림 6은 잔향의 유/무와 잔향곡선에서 초기감쇠시간의 단축에 따른 음성의 명료도 차이를 확인하기 위하여 실시한 주관평가의 결과이다. 10명의 청취자 모두가 잔향이 없는 source 음성이 가장 명료도가 좋은 것으로 평가하였다. 7번과 8번 청취자를 제외하고는 풀밴드 NLMS에 의해 잔향이 저감된 음성이 실음장의 잔향이 포함되어 있는 음성보다 명료도가 좋은 것으로 평가하였다. 이러한 결과는 잔향곡선의 초기경사의 기울기인 초기감쇠시간에 따라서 음성의 명료도가 크게 영향을 받는다는 것을 확인할 수 있었다.

표 1은 10개의 음성에 대해서 주관 평가한 결과이다. 각 음성신호에 대해서 1등급/2등급/3등급의 형태로 표시하였다. 주관평가 결과 /아/ 음성의 경우 source는 9명이 1등급을 1명이 3등급으로 평가하였다. 잔향이 포함된 실제 음장에서는 3명이 2등급을 6명이 3등급으로 평가하였다. 초기감쇠시간이 단축된 풀밴드에서는 1명이 1등급, 7명이 2등급, 3명이 3등급으로 평가하였다. 10개의 음성

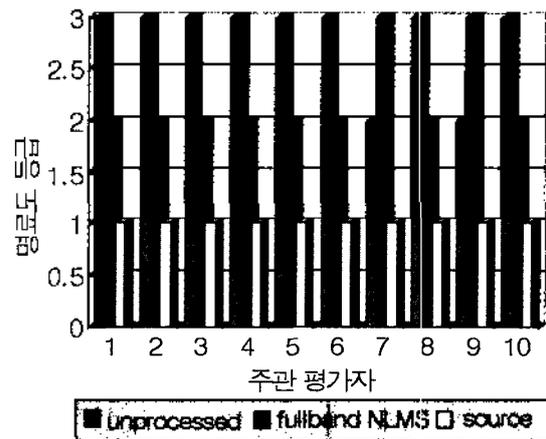


그림 6. 초기감쇠시간의 변화에 따른 문장의 주관평가 결과
Fig. 6. Result of subjective assessment one sentence according to changing EDT.

표 1. 10개 음성의 명료도 주관평가 결과
Table 1. Result of subjective assessment 10 words.

	A(source) 1/2/3등급	B(unprocessed) 1/2/3 등급	C(fullband) 1/2/3 등급
아	9 / 0 / 1	0 / 3 / 6	1 / 7 / 3
야	8 / 1 / 1	1 / 3 / 7	1 / 6 / 3
어	7 / 3 / 0	1 / 1 / 8	2 / 6 / 2
여	10 / 0 / 0	0 / 5 / 5	0 / 5 / 5
오	10 / 0 / 0	0 / 1 / 9	0 / 9 / 1
요	10 / 0 / 0	0 / 5 / 5	0 / 5 / 5
우	9 / 0 / 1	0 / 4 / 6	1 / 6 / 3
유	10 / 0 / 0	0 / 2 / 8	0 / 8 / 2
으	7 / 2 / 1	1 / 2 / 7	2 / 7 / 1
이	8 / 1 / 1	0 / 3 / 7	2 / 6 / 2

이 전체적으로 발음시간이 문장에 비해서 매우 짧으므로 청취자가 잔향의 저감 정도를 평가하기가 문장보다는 어려워 하였다. 하지만 잔향곡선에서 초기감쇠시간이 단축된 음성이 처리되지 않은 음성에 비해서 명료도가 높게 나타난 것을 실험을 통해서 확인할 수 있었다.

IV. 결론

본 논문에서는 실내 전달함수의 잔향곡선에서 초기감쇠시간을 단축시키기 위해서 복수채널에 의한 MINT에 기초하여 잔향을 저감시켰다. 잔향저감 시스템에서 마이크로폰 개수는 2개와 3개를 사용하였으며, 적응처리는 풀밴드 LMS, 풀밴드 NLMS, 8 서브밴드 NLMS로 모델링하여 잔향시간을 비교 검토하였다. 또한 초기감쇠시간의 단축에 따른 음성의 명료도 향상을 평가하기 위하여 하나의 문장과 10개의 음성에 대해서 주관평가를 실시하였다.

76 m³의 실제 음장에서 잔향에너지가 -20 dB까지 감소하는데 걸리는 시간은 2개의 마이크로폰을 사용한 풀밴드 NLMS에서 1/3로 단축할 수 있었다. 또한 초기감쇠시간도 39 msec에서 26 msec로 단축할 수 있었다. 10명의 청취자를 대상으로 주관평가한 결과에서는 80%의 청취자가 음성의 명료도가 향상된 것으로 평가하였다. 실험 결과 실내의 잔향곡선에서 MINT 법을 이용한 초기감쇠시간의 단축이 명료도 향상에 기여한다는 것을 확인하였다.

참고 문헌

1. B. Yegnanarayana and B. S. Ramakrishna, "Intelligibility of speech under nonexponential decay conditions," *J. Acoust. Soc. Am.*, 58(4), pp. 853-857, 1975.
2. 김성훈, "방송음향 총론", 기전연구소, 2001.
3. B. Atal, M. R. Schroeder and G. M. Sessler, "Subjective Reverberation time and its Relation to Sound Decay," *Proc. 5th Congr. Acoustica*, pp. 344-347, 1965.
4. J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small room acoustics," *J. Acoust. Soc. Am.*, vol. 65, No. 4, pp. 454-459, Apr., 1979.
5. M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. On Acoustics, Speech, and Processing*, vol. 36, No. 2, pp. 145-152, 1988.
6. ITU-R, "Method for subjective assessment of small impairments in audio system including multichannel sound systems," ITU-R Draft New Recommendation, Document 10/65, November 1993.
7. S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, New Jersey 1991.

저자 약력

- 차 경 환 (Kyung-Hwan Cha)
한국음향학회지 제19권 제3호 참조