

# 음향 피드백 제거를 위한 전대역, 협대역 적응 필터의 비교

## A comparative study of full-band and sub-band approaches to acoustic echo cancellation

° 신 민 철\*·김 상 명

Min-Cheol Shin and Sang-Myeong Kim

**Key Words :** AEC (Acoustic Echo Canceller), Fullband adaptive filter (전대역 적응 필터)  
Subband adaptive filter (협대역 적응 필터)

### ABSTRACT

The system in which a microphone and a loudspeaker are simultaneously used can cause an echo. The echo is caused by feedback between the output of the loudspeaker and the input of the microphone. The acoustic echo canceller is a device to cancel the echo in a communication system. Its general procedure for cancellation is first estimating the plant response of the feedback path and then eliminating the feedback signal from the input signal. In this paper, full-band and sub-band approaches are compared by using some simulation examples.

#### 기호설명

- AEC : Acoustic Echo Canceller
- SAEC : Subband Acoustic Echo Canceller
- LMS : Least Mean Square
- NLMS : Normalized Least Mean Square
- $J$  : Modeling error
- SEN : System Error Norm

### 1. 서 론

음향 피드백이란 스피커와 마이크를 동시에 사용하는 음향 시스템에서 스피커의 출력신호가 마이크로 되먹임되는 현상을 말한다. 여기서 되먹임되는 스피커의 출력은 그림 1에 나타낸 것과 같이 스피커에서 직접 마이크로 입력이 되는 음향 신호와 시스템이 구현되어있는 환경에 의해서 결정되는 메아리 성분으로 나뉜다[8].

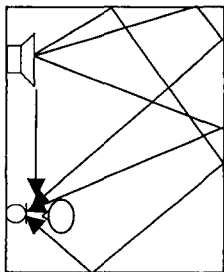


그림 1. 음향 피드백 신호

잘 알려진 바와 같이 음향 피드백 현상은 아주 오랫동안 음향 시스템의 큰 문제로 대두되었지만, 아직도 명확하게 해결되지 않은 상태이다. 음향 시스템에서 이러한 현상은 특히 전화기와 같은 통신 시스템에서 커다란 문제가 된다.

최근 양방향 통신 방법 개발과 아울러 원격회의와 핸즈프리 시스템이 실용화되어 사용되고 있다. 이러한 시스템에서는 음향 피드백이 원활한 통신을 방해하므로 이를 제거하는 과정이 통신 시스템에 있어서 필수적이라고 할 수 있다. 음향 피드백을 제거하기 위해서 적응 필터를 이용하여 음향 피드백 경로의 충격응답을 추정하여 메아리 성분을 제거한다.

이때 음향 피드백 경로의 충격응답은 화자의 위치변화나 환경변화에 매우 민감하므로 실시간으로 추정을 해야 한다. 다시 말해서 실시간 시스템 모델링 기법이 음향 피드백 제거의 핵심기술이다.

이러한 실시간 시스템 모델링 기법은 크게 전대역 적응 필터를 이용하는 기법과 협대역 적응 필터를 이용하는 기법, 이 두 가지로 나뉘어진다.

본 연구에서는 위 두 가지 적응 필터의 AEC 적용의 장단점을 컴퓨터 모의 실험을 통해 조사 비교한다. 또한 LMS 적응 알고리즘에서의  $J$ 와 SEN의 비교를 통해 각각의 정확한 적용 분야를 알아본다.

\* 광주과학기술원 기전공학과  
E-mail : mcshin@kjist.ac.kr  
Tel : (062) 970-2409, Fax : (062) 970-2384

## 2. 전대역, 협대역 적응 필터

### 2.1 전대역 적응 필터 설계

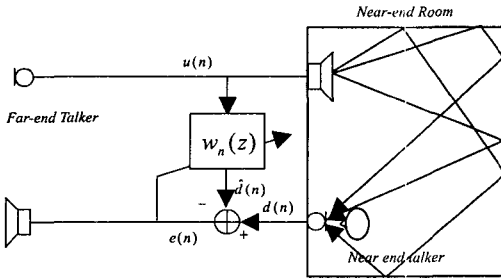


그림 2. 전대역 적응 필터의 구현

그림 2는 전대역 적응 필터 기법의 구현 시스템을 보인 것으로 주파수 영역별이 아닌 전체 대역을 대상으로 적응 필터 기법을 적용하는 방법이다. 여기서  $u(n)$ 는 Far-end 편 화자의 음성신호며 음향 피드백 제거 시스템의 입력신호로도 볼 수 있다.  $d(n)$ 는 Near-end 편 화자의 음성신호와 Near-end Room에 의한 음향 피드백 신호를 합한 신호가 된다. 그리고  $\hat{d}(n)$ 는 Near-end Room의 음향 피드백 경로를 추정한 적응필터  $w_n(z)$ 를 통과한 신호이며  $e(n)$ 은  $d(n)$ 에서  $\hat{d}(n)$ 을 뺀 신호다. 이상적으로  $e(n)$ 은 Near-end 편 화자의 음성신호만으로 이루어져야 한다.

적응필터  $w_n(z)$ 를 구하는 전대역 알고리즘으로 성능이 우수하고 구현이 용이한 Least Mean Square(LMS) 적응 알고리즘과 이의 변형인 Normalized Mean Square (NLMS) 알고리즘이 널리 쓰인다.

#### (1) LMS 적응 알고리즘[2]

$w(n)$ 이 시점  $n$ 에서의 적응 필터 계수의 벡터라고 하면 LMS 적응 알고리즘의 결과식 즉 적응 필터 계수의 업데이트를 위한 식은 아래와 같다.

$$w(n+1) = w(n) + \mu u(n)e^T(n) \quad (1)$$

여기서  $T$ 는 벡터의 transpose를 나타낸다. 식 (1)에서 step-size  $\mu$ 는 결과의 수렴 여부를 결정하게 된다.  $p$ 가 입력신호의 길이라고 하면  $\mu$ 가

$$0 < \mu < \frac{2}{(p+1)E[u(n)^2]} \quad (2)$$

식(2)의 조건을 만족해야  $w$ 는 수렴하게 된다. 여기서  $E[\cdot]$ 은  $\circ$ 의 기대값을 나타낸다.

#### (2) NLMS 적응 알고리즘[2]

Normalized Least Mean Square (NLMS) 알고리즘은 LMS 알고리즘에서 문제가 되는 입력신호에 따른 Gradient 추정 에러의 영향을 입력신호의 표준화시켜 해결하였다. NLMS의 결과식은 식 (3)와 같으며

$$w(n+1) = w(n) + \frac{\tilde{\mu} u(n)e^T(n)}{a + \|u(n)\|^2} \quad (3)$$

여기서 임의의 작은 상수  $a$ 는 입력신호가 매우 적을 때 발산되는 현상을 막기 위해 추가한다.  $\|\cdot\|$ 은  $\circ$ 의 norm을 의미한다. 식 (3)의 수렴 조건은  $0 < \tilde{\mu} < 2$ 이다.

### 2.2 협대역 적응 필터 설계

협대역 적응 필터는 새로운 알고리즘이 아니라 기존의 적응 알고리즘을 주파수 대역별로 적용하는 구조이다. 다시 말하면, 입력신호와 시스템을 통과한 출력신호를 여러 주파수 대역으로 나눈 후 이를 각각의 대역별 적응 필터를 이용하여 각 대역에서 음향 피드백 시스템의 충격응답을 구하는 방법이다.

협대역 적응 필터의 구조로 기존의 지연을 함유하는 구조를 향상시킨 무지연 적응 필터[5]가 널리 쓰인다.

#### (1) 무지연 협대역 적응 필터

아래의 그림 3과 같이 무지연 협대역 적응 필터는 입력신호  $x(t)$ 와 시스템  $h$ 를 통과한 출력신호  $d(t)$ 를 polyphase 분해 필터[3]인  $F_0(z), F_1(z), \dots, F_{M-1}(z)$ 를 이용하여  $M$ 개의 협대역으로 분해한다. 분해된 신호들은  $D$ -Decimator를 거친 후에 각각의 대역별로 적응 알고리즘을 통해 충격응답을 추정하는 과정에 사용된다. 이러한 각 대역에서 추정된 충격응답을 바탕으로 weight-transform 기법을 통해 전대역 충격응답을 추정하는 구조로 되어있다.

각 대역에서 구한 대역별 충격응답을 전대역 충격응답으로 바꾸는 weight transform 기법은 여러

가지가 있지만 여기에서는 FFT-2 weight transform 기법 [6]이 사용되었다.

$$J(\%) = \frac{E[e^2(t)]}{E[d^2(t)]} \times 100 \quad (5)$$

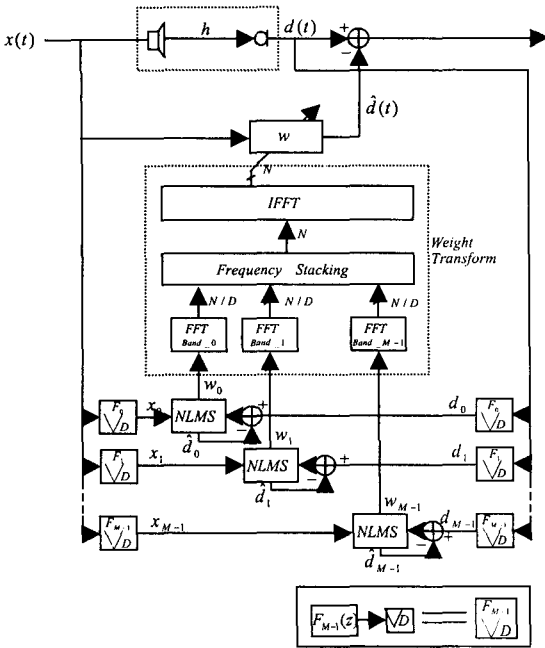


그림 3. 무지연 협대역 적응 필터의 구조

그림 3의 polyphase 분해 필터는  $D$ -Decimator와 하나의 블록다이어그램으로 나타내었다.

### 3. 에러 분석

#### 3.1 모델링 에러 ( $J$ ) [7]

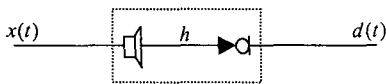


그림 4. 1 채널 음향 시스템

그림 4와 같은 1 채널 음향 시스템에서 시스템  $h$ 의 입력신호를  $x(t)$ , 출력신호를  $d(t)$ 라 하자.  $\hat{h}(t)$ 를 음향 시스템의 추정된 충격응답이라고 하면 이는 수학적으로 식 (4)와 같이 표현된다.

$$d(t) = \hat{h}(t) * x(t) + e(t) \quad (4)$$

여기서  $*$ 는 선형중첩적분을 나타내고  $e(t)$ 는 에러 신호이다. 이를 토대로 모델링 에러( $J$ )를 식(5)과 같이 정의한다.

#### 3.2 System Error Norm (SEN) [8]

SEN은 두 충격응답 간의 직접적인 차이를 가늠하는 에러로서 다음 식과 같이 정의된다.

$$SEN(dB) = 10 \log \left( \sum_{i=0}^{L-1} [h(i) - \hat{h}(i)]^2 \right) \quad (6)$$

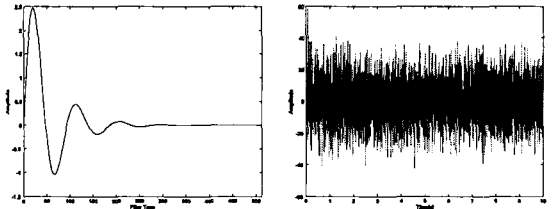
여기에서  $L$ 은 충격응답의 길이이며  $h(t)$ 와  $\hat{h}(t)$ 는 각각 음향 시스템의 실제 충격응답과 추정된 충격응답이다.

#### 3.3 LMS 알고리즘에서의 $J$ 와 SEN의 비교

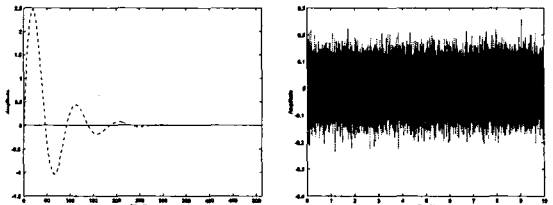
보편적으로  $J$ 와 SEN은 같은 의미를 갖는 것으로 생각하기 쉽다. SEN 값이 작으면  $J$  값 역시 작다고 볼 수 있다. 역으로  $J$ 가 작으면 출력신호 간의 차이가 적고 이는 시스템의 충격응답 추정이 잘 된 것으로 볼 수 있으므로 SEN 값 역시 적다고 할 수 있다.

하지만 전자는 맞지만 후자는 항상 맞는다고 할 수 없다. 그 이유는 LMS 알고리즘이 에러의 평균의 제곱값을 목적 함수로 하고 이를 최소화시키는 이론이기 때문이다.

다음의 예를 통해 이해할 수 있다.



(a) white noise 입력신호



(b) 고역 통과된 white noise 입력신호

그림 5. 입력신호에 따른 LMS의 수렴경향, 출력오차

그림 5은 간단한 1 자유도 시스템의 충격응답 (512 taps)을 LMS 알고리즘으로 추정한 컴퓨터 모의 실험 결과이다. 그림 5 (a)의 경우 white noise를 입력신호로 사용한 경우이고 그림 5 (b)의 경우 고역 통과( $F_s=8000\text{Hz}$ , Highpass

cutoff frequency = 2000Hz)된 white noise 를 입력신호로 사용한 경우이다. 각각에서 왼쪽은 추정된 충격응답(실선)과 원래의 1 자유도 시스템(점선)을 나타내었고 오른쪽은 출력신호  $d(t)$  (검은색)와 에러 신호  $e(t)$  (회색)를 나타낸 그림이다. 그림 5(a)는 쉽게 예측할 수 있는 결과이지만 그림 5(b)의 경우는 시스템 충격응답의 추정 오차는 크지만(SEN:22.2183dB) 출력신호의 오차는 적다( $J$ :1.089%). 이러한 경우  $J$ 값으로 추정된 충격응답의 정확도를 예측하기 곤란하다. 또한 적응 필터의 길이가 예측하고자 하는 충격응답의 길이보다 적은 경우에도 같은 현상을 보인다[9].

즉 SEN 은 시스템간의 충격응답의 추정 분야에 그리고  $J$  는 출력신호간의 오차를 최소화할 때 사용되어야 한다. AEC 의 성능은 출력오차와 깊은 관계에 있으므로  $J$  로 판단해야 하고 적응 필터의 시스템 충격응답의 추정 정도는 SEN 으르 판단하는 것이 타당하다.

#### 4. 컴퓨터 모의 실험

컴퓨터 모의 실험에서 샘플링 주파수는 8000Hz 이고 이 실험에 사용된 충격응답은 실제의 보통방의 충격응답을 사용하였다(그림 6).

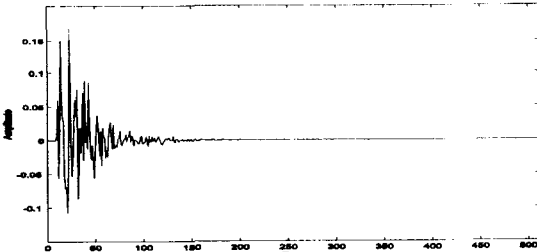
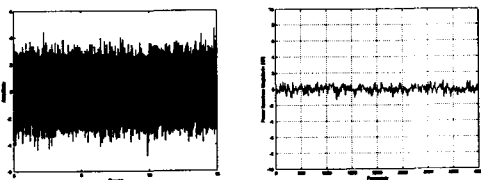
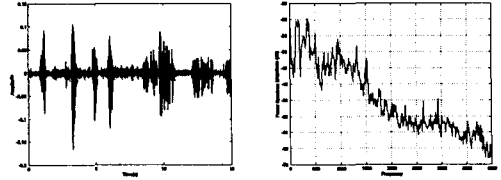


그림 6.모의 실험에 사용된 충격응답(512 taps)

이 모의 실험에서는 전대역 적응 필터 구조의 성능을 알아보기 위해 NLMS 알고리즘을 사용하였고 협대역 적응 필터 구조에서도 적응 알고리즘으로 NLMS 를 사용하여 대역 수에 따른 비교를 수행하였다. 입력신호로는 white noise 신호와 음성신호를 사용하였다(그림 7).



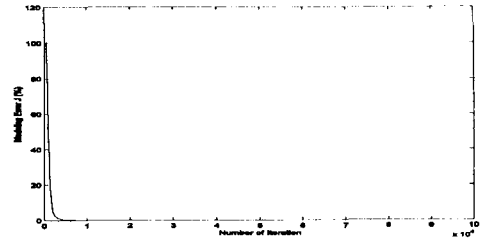
(a) white noise



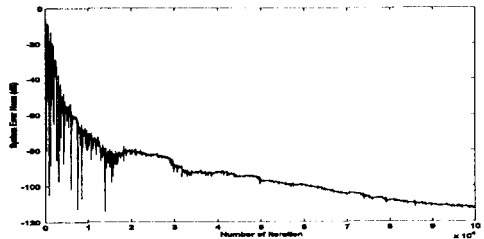
(b) 음성신호

그림 7. 모의 실험에 사용된 입력신호

#### 4.1 전대역 적응 필터 구조의 학습 선로

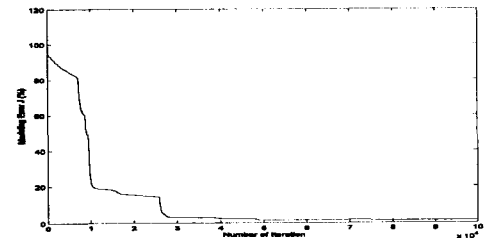


(a) Modeling Error  $J$  (%)

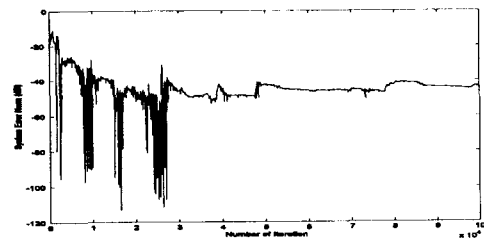


(b) System Error Norm (dB)

그림 8. White noise 입력신호의 경우



(a) Modeling Error  $J$  (%)



(b) System Error Norm (dB)

그림 9. 음성신호가 입력신호인 경우

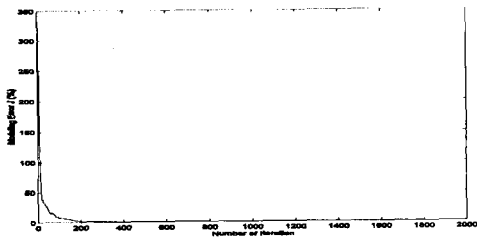
NLMS 적응 알고리즘을 사용하는 전대역 적응 필터 구조는 그림 8 과 같이 입력신호가 white noise 신호이면  $J$ 는 5000 iteration 후에 약 0%로 수렴하게 되고 SEN은 120000 iteration 후 수렴값은 약  $-115\text{dB}$ 이다. 그림 9 과 같이 음성신호가 입력신호이면  $J$ 는 50000 iteration 후에 약 0%로 수렴하게 되고 SEN은 30000 iteration 후 약  $-50\text{dB}$ 로 수렴한다.

전대역 적응 필터 구조에서는 입력신호에 따라  $J$ 와 SEN 모두 편차가 컸으며, AEC의 실질적인 입력신호인 음성신호에 대한 학습선로는 white noise 입력신호의 경우에 비해 그 효율이 떨어진다.

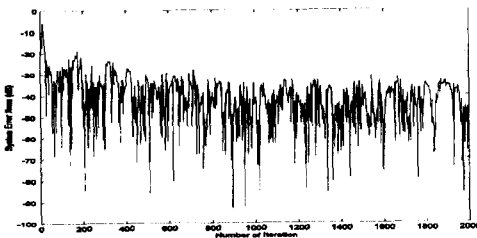
## 4.2 협대역 적응 필터 구조의 학습 선로

협대역 적응 필터 구조의 컴퓨터 모의 실험에서는 그림 3의 무지연 협대역 적응 필터 구조를 사용하여 주파수 대역을 각각 64 대역과 128 대역으로 나눈 두 가지 경우에 대해서  $J$ 와 SEN의 학습 선로를 그려 보았다.

### (1) 64 대역 적응 필터의 경우

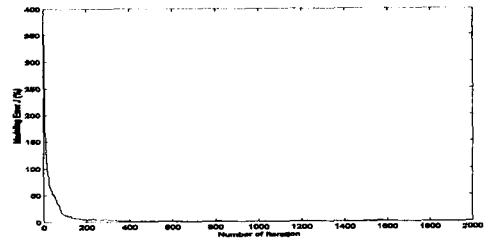


(a) Modeling Error  $J$  (%)

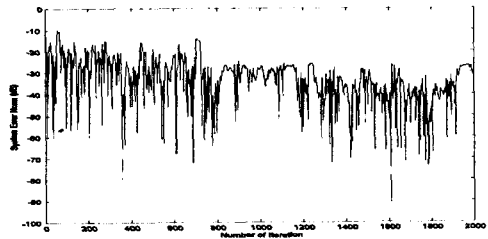


(b) System Error Norm (dB)

그림 10. White noise 입력신호의 경우



(a) Modeling Error  $J$  (%)

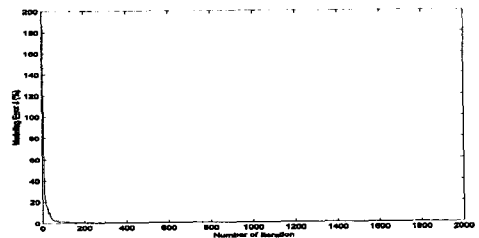


(b) System Error Norm (dB)

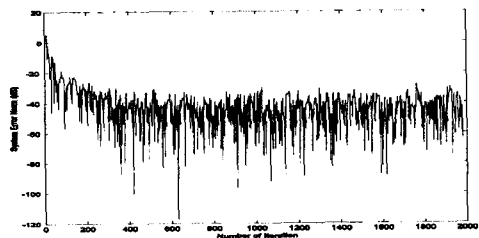
그림 11. 음성신호가 입력신호인 경우

64 대역 적응 필터 구조에서는 입력신호가 white noise 신호일 경우  $J$ 는 약 500 iteration 후에 수렴하게 되고 SEN은 1000 iteration 후  $-50\text{dB}$  정도로 수렴한다. 입력신호가 음성인 경우에는  $J$ 는 500 iteration 후에 수렴하게 되고 SEN은 1500 iteration 후 약  $-50\text{dB}$ 로 수렴한다.

### (2) 128 대역 적응 필터의 경우

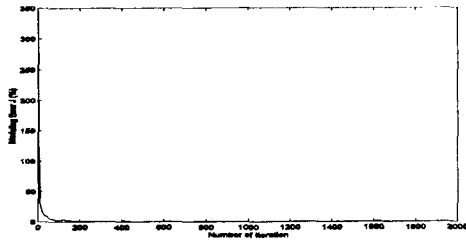


(a) Modeling Error  $J$  (%)

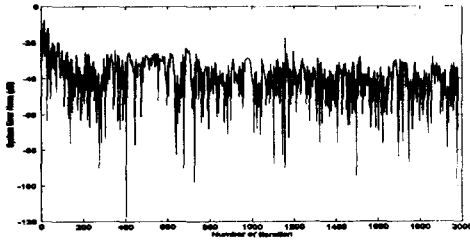


(b) System Error Norm (dB)

그림 12. White noise 입력신호의 경우



(a) Modeling Error  $J$  (%)



(b) System Error Norm (dB)

그림 13. 음성신호가 입력신호인 경우

128 대역 적응 필터 구조에서는 입력신호가 white noise 신호일때  $J$ 는 약 200 iteration 후에 거의 0%로 수렴하게 되고 SEN 은 500 iteration 후  $-50\text{dB}$  정도로 수렴한다. 음성신호가 입력신호인 경우  $J$ 는 약 200 iteration 후 수렴하게 되고 SEN 은 1000 iteration 후 약  $-50\text{dB}$ 로 수렴한다.

협대역 적응 필터 구조의 경우 전대역 필터의 경우와 비교하여 입력신호에 따른 성능의 편차가 작다. 실질적인 AEC 의 입력신호인 음성신호의 경우  $J$ 는 전대역 구조에 비해 64 대역의 경우 수렴 iteration number 가 1/10 배, 32 대역의 경우 1/25 배 감소된다. SEN 은 전대역, 협대역 경우 모두 비슷한 수렴값( $-58\text{dB}$ )을 갖고 수렴 iteration 수는 현저하게 작다.

### 5. 전대역, 협대역 적응 필터의 계산량의 비교

전대역 적응 필터의 구조에서 샘플당 필요로 하는 계산량은 적응필터의 길이를  $N$  이라고 하면  $2 \times N$ 이다. 그림 3 의 무지연 협대역 적응 필터 구조를 기초로 한 계산량을 알아보면 식 (7-10)과 같다.

$$C_1 = \frac{4K}{M} + 4 \log_2 M \quad (7)$$

$$C_2 = 2 \frac{8N}{M} \quad (8)$$

$$C_3 = 2 \log_2 \frac{2N}{M} + \log_2 N \quad (9)$$

$$C_4 = N \quad (10)$$

$C_1 \sim C_4$ 는 각각 협대역 필터링 과정, 대역별 적응 필터의 계수를 업데이트 하는 과정, weight transform 과정, 그리고 전대역 총격응답과 입력 신호의 필터링 과정에서의 계산량을 나타낸다. 여기서  $M$ 은 대역수를 나타내고,  $K$ 는 prototype-polyphase 분해 필터의 길이이다. 그러므로 전체 계산량은 다음 식 (11)과 같게 된다.

$$C_{total} = C_1 + C_2 + C_3 + C_4 \quad (11)$$

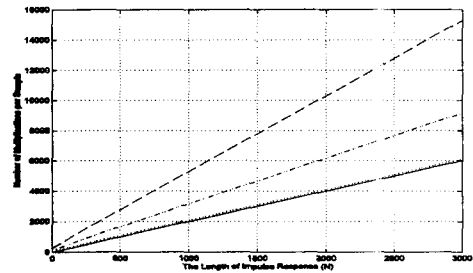


그림 14. 전대역, 협대역 적응필터의 계산량 비교 (실선: 전대역, dash : 4 대역, dash-dot : 8 대역, dot-dot : 16 대역)

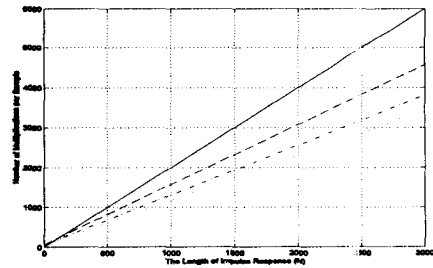


그림 15. 전대역, 협대역 적응필터의 계산량 비교 (실선: 전대역, dash : 32 대역, dash-dot : 64 대역, dot-dot : 128 대역)

그림 14 와 15 은 추정하고자 하는 총격응답의 길이에 따른 전대역 적응 필터의 계산량(실선)과 협대역 적응 필터의 계산량을 비교한 그림이다. 그림 14 와 같이 대역의 수가 16 대역 이하일 때는 전대역 적응 필터 방법의 계산량이 오히려 적었지만 그림 15 과 같이 32 대역 이상일 경우 총격응답의 길이가 증가할수록 계산량이 감소한다. 특히 128 대역의 경우에는 50%에 가까운 계산량의 감소를 기대할 수 있다. 하지만 다음 그림 16 과 같이 적응 필터의 길이가 짧을 경우에는 전대역 적

응 필터 방법의 계산량이 더 적을 수 있다.

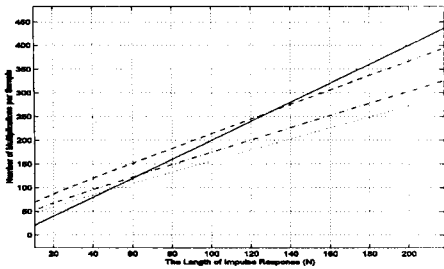


그림 16. 적응 필터의 길이가 짧을 경우 (실선: 전대역, dash : 32 대역, dash-dot : 64 대역, dot-dot : 128 대역)

## 6. 모델링오차( $J$ )와 모델오차( $SEN$ )의 비교

실제적인 AEC 의 성능은 출력신호간의 오차에 의해 결정되기 때문에  $J$  의 값이 중요한 요소가 된다. 4 절에서 알아본 바와 같이 전대역 적응 필터구조와 협대역 적응 필터 구조에서 실질적인 입력신호인 음성신호에 대한 성능이  $J$  의 측면에서는 협대역 적응 필터의 경우에 뛰어났다. 물론 SEN 은 3 절에서 언급한 것과 같이 시스템 충격응답 간의 오차를 명확히 보여준다.

모델링 에러  $J$  는 실질적인 음향 피드백 제거기의 효율과 관련이 있고 SEN 은 만약 적응 필터를 시스템의 특성을 알아보는 용도로 쓰는 경우에 시스템 충격응답의 추정 정도와 관련이 있는 에러이다.

본 논문에서 정의된 두 가지 오차 즉 모델링 오차( $J$ )와 모델오차( $SEN$ )는 각각 다른 성격을 가지고 있으며 AEC 의 성능 판단은  $J$  로 해야 한다는 결론을 내릴 수 있다.

## 7. 결론

본 논문에서는 전대역 적응 필터 구조와 협대역 적응 필터 구조를 모델링 에러  $J$  와 System Error Norm( $SEN$ )의 관점에서 고찰하였고 또한 실질적인 계산량을 알아봄으로써 두 구조를 비교하여 어느 구조가 더 실질적인 구현에 효율적인가 하는 것을 밝혔다. 실질적인 음향 피드백 제거 기법에의 적용 효율은 구조상의 복잡함으로 인해 프로그래밍이 용이하지 않다는 점을 제외하고는 협대역 구조가 더 효율적인 구조라는 것을 알 수 있다.

실질적인 음향 피드백 제거기의 구현을 위해서는 범용 프로세서가 아닌 디지털 시그널 프로세서가 필요한데 이의 구현에 있어서도 단위 시간안에 계산되는 계산량에 프로세서의 가격이 비례하게 된다. 협대역 적응 필터 구조를 쓰게 되면 같은 가격에서는 좀 더 정확한 결과를 기대할 수 있고 또한 같은 성능이라면 좀 더싼 가격에 구현할 수 있다는 장점이 있다.

## 후 기

본 연구는 한국과학재단 목적기초연구(R05-2003-000-11999-0) 지원으로 수행되었음.

## 참고문헌

- [1] Monson H. Hayes, 1996, Statistical Digital Signal Processing and Modeling, John Wiley & Sons, Inc., New York.
- [2] Widrow B and Stearns S.D, 1985, Adaptive Signal Processing, Prentice Hall
- [3] E. R. Ferrara Jr, 1985, Frequency-domain Adaptive filtering in Adaptive Filters, C.F.N.Cowen and P.M.Grant, Eds. Englewood Cliffs, NJ : Prentice-Hall
- [4] A. Gilloire and M. Vetterli, Jan 1992, Adaptive filtering in subbands with critical sampling : Analysis, experiments, and application to acoustic echo cancellation, IEEE Trans. Signal Processing, vol. 40, pp. 1862-1875
- [5] Dennis R. Morgan, James C. Thi , August 1995, A Delayless Subband Adaptive Filter Architecture, IEEE Trans. Signal Processing, Vol. 43, No.8.
- [6] J. Huo, S. Nordholm and Z. Zang, Nov 2001, New weight transform schemes for Delayless Subband Adaptive Filtering, Proceedings of IEEE Global Telecommunications Conference, Texas, USA, vol.1, pp. 197-201.
- [7] S. Haykin, 1996, Adaptive Filter Theory, 3<sup>rd</sup>. ed., Prentice Hall
- [8] Christina Breining, Pia Dreiseitel, July 1999, Acoustic Echo Control, IEEE Signal Processing Magazine.
- [9] Sang-Myeong Kim and Se-Myong Wang, Jan 2003, A Wiener Filter Approach to the Binaural Reproduction of Stereo Sound, Active Acoustics Lab Tech. Paper 01, Mechatronics, K-jist.