

동시 통화검출 전용필터와 가중 Overlap-Add 기법을 적용한 서브밴드 음향 반향 제거기

Subband Acoustic Echo Canceller with Double-Talk Detector Using Weighted Overlap-add Method and Dedicated Filter

고 충 기*, 이 원 철*, 이 충 용**

(Chung Ki Ko*, Won Cheol Lee*, Chung Yong Lee**)

*숭실대학교 정보통신전자공학부, **연세대학교 기계전자공학부

(접수일자: 2000년 8월 28일; 채택일자: 2000년 10월 9일)

본 논문은 부밴드별 반향제거 필터 뱅크를 구현하기 위한 가중 Overlap-add 적응필터를 이용한 단일 채널 음향 반향 제거기를 제안한다. 기존 입력 신호의 고유치 분포율에 의존하여 수렴특성이 결정되는 NLMS (Normalize LMS) 알고리즘을 사용하여 전대역 처리 과정에서 발생하는 수렴 성능의 저하를 방지하고, 효율적인 블록 병렬처리가 가능한 부밴드 처리기법인 가중 Overlap-add 방식을 적용한 적응 반향제거기의 성능을 고찰한다. 또한 본 논문에서는 동시 통화 검출을 위한 전용 필터와 에너지 비교 방법을 동시에 사용하는 새로운 형태의 동시통화 검출 기법을 제안한다. 모의 실험을 통해 제안한 적응 반향 제거기와 동시 통화 검출기의 효과적인 근단 화자 검출 방법을 보인다.

핵심용어: 가중 Overlap-add, 반향 신호, 동시 통화 검출, 부 밴드, 적응 필터

투고분야: 음향 신호처리 분야 (1.1)

In this paper, we propose a subband acoustic echo canceller using the weighted Overlap-add adaptive filter bank to prevent the decrease of convergence speed in full-band NLMS processing, and make it possible to realize the adaptive filter in block-parallel processing, this paper introduces the weighted overlap-add technique for subband echo canceller. Moreover, we propose a new double-talk detector which employs dedicated filter in addition to the energy comparison method simultaneously. The computer simulation results show that the performance of the proposed subband adaptive echo canceller double-talk detection

Key words: Weighted overlap-add, Adaptive filter, Subband, Double talk detection

Subject classification: Acoustic signal processing (1.1)

I. 서 론

정보화 사회의 발전과 더불어 현재 단일 채널 음향 반향 제거기는 다양한 핸드 프리 통신상에 특히 원격회의 시스템에 폭 넓게 사용되어 왔다. 현재 대부분의 원격회의 시스템은 하나의 전이중 (full-duplex) 오디오 채널만을 이용하여 통신을 하고 있다. 음향 반향 신호는 수신된 상대방의 신호 (이하 원단화자 신호)가 스피커와 통화자가 위치한 실내공간을 거쳐 다시 마이크로폰으로 입력되는 경우를 말하며 이로 인해 원단화자는 자기가 한 말을 통화 왕복 시간 뒤에 듣게 되어 통화에 지장을 주게 된

다. 이때 수신실의 스피커와 마이크로폰 사이에 발생하는 상호교차 (cross-coupling)에 의해 반향 신호가 발생하며, 이를 제거하기 위하여 음향 반향 제거기가 사용된다. 이와 같은 음향 반향 신호를 통화에 지장을 주지 않고 적응 필터를 사용하여 효과적으로 제거하기 위한 단일 채널 음향 반향 제거기에 대한 많은 연구가 진행되고 있다 [1][2]. 본 논문에서는 단일 채널 원격회의 시스템 상에서 블록처리가 가능하고 우수한 수렴속도를 가지며 높은 ERLE (Echo Return Loss Enhancement)를 유지하는 가중 Overlap-add 구조에 NLMS 알고리즘을 이용한 음향 반향 제거기를 제안한다.

일반적인 동시 통화 검출 기법은 에너지 비교법과 추정된 신호와 마이크로폰 입력신호의 상관계수에 의한 방법[3][4]들이 있으며, 이들은 근단 화자의 신호 레벨이 작은 경우 검출 능력이 저하된다. 이를 위해 단일 음향반향

책임저자: 고충기 (fabre@amcs.ssu.ac.kr)

156-743 서울시 동작구 상도5동 숭실대학교 정보통신전자 공학부

통신신호처리 연구실

(전화: 826-7980; 팩스: 821-7653)

제거기의 문제점을 2장에서 살펴보고 3장에서는 일반적인 서브밴드 구조를 설명하고 4장에서 동시 통화 검출 방법을 소개하며 5장 가중 Overlap-Add 구조를 6장은 제안한 전용 필터에 대해 설명을 하고 7장에서 모의 실험을 통해 얻은 결과를 보이고 8장에서 결론을 맺는다.

II. 음향 반향 제거기의 문제점

음성 통신의 궁극적인 목적은 통화자의 음성을 왜곡 없이 상대방에게 전달하는 것이다. 그러나 핸드 프리 방식으로 이루어지는 단일 채널 통화 환경에서 원단 화자 신호 $x(n)$ 은 스피커, 사무실 공간 그리고 마이크로폰으로 구성되는 반향 경로를 경유하여 다시 원단 화자가 있는 곳으로 되돌아간다. 이로 인해 원단 화자는 근단 화자의 신호 $s(n)$ 뿐만 아니라 자기가 말한 음성을 통화 왕복 지연 시간 뒤에 듣게 되어 통화에 불편함을 느끼게 된다. 그림 1은 원거리 회의에서 발생하는 음향 반향 경로와 반향 제거기의 위치를 나타내고 있다.

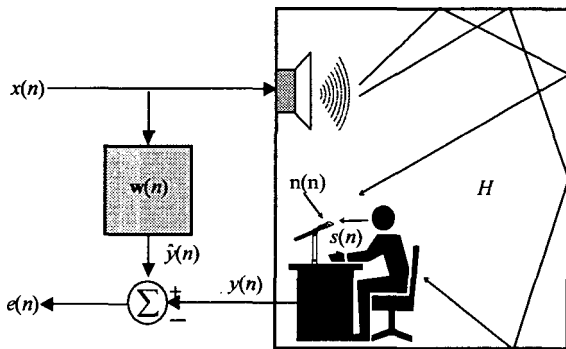


그림 1. 반향 제거기의 기본 구조
Fig. 1. Structure of the acoustic echo canceller.

여기서 $x(n)$ 은 스피커로 출력되는 원단 화자의 음성 신호이고, $y(n)$ 은 마이크로폰으로 유입되는 반향신호를 나타낸다. $s(n)$ 과 $n(n)$ 은 각각 근단 화자의 신호와 주변의 잡음을 나타낸다. 적응 반향 제거기의 적응 필터에서 추정된 신호는 다음과 같다.

$$\hat{y}(n) = w^T(n) x(n) \tag{1}$$

H 는 스피커로부터 특정 마이크로폰까지의 반향경로를 대표하는 전달함수를 나타내며 $w(n)$ 는 H 를 추정된 적응 필터의 필터 계수를 나타낸다. 적응 반향 제거 필터에서 추정된 신호와 반향 신호의 차가 결국 반향 제거기의 출력 값이다.

$$e(n) = y(n) - \hat{y}(n) \tag{2}$$

$e(n)$ 은 반향 제거기 출력인 에러값이며, 여기서 기준 입력신호 $x(n)$ 과 적응필터의 계수 $w(n)$ 은 다음과 같다.

$$x(n) = [x(n), x(n-1), \dots, x(n-L+1)]^T \tag{3}$$

$$w(n) = [w_0(n), w_1(n), \dots, w_{L-1}(n)]^T \tag{4}$$

이때 $[\cdot]^T$ 은 $[\cdot]$ 의 전치행렬 (Transpose Matrix)을 나타내며 L 은 필터 탭수이다. 일반적으로 원격 회의시 화자의 움직임이나, 스피커의 볼륨조절과 같은 주위 환경의 변화가 수시로 발생하게 되므로, 음향 반향 제거 필터는 일반적으로 적응 필터를 사용하게 된다. 식(5)는 NLMS (Normalized Least Mean Square) 알고리즘을 사용하는 적응 반향 제거 필터의 계수 갱신식이다. 여기서 μ 는 수렴 속도와 정상 상태에서의 오차 크기를 조정하는 수렴 상수이다.

$$w(n) = w(n) + \frac{\mu}{\|x(n)\|^2} x(n) e(n) \tag{5}$$

수신룸의 환경 변화에 반향 신호를 정확히 추정하면서 계산량을 적게 하는 적응 필터를 구현해야 하며 또한 수렴속도는 입력신호에 대해 적응 필터가 얼마나 빠르게 반향 신호를 추정하여 에러를 줄일 수 있는가 하는 문제이므로 효과적인 수렴 속도는 음성 통신상의 중요한 문제가 된다. 현재까지는 단일 채널 음향 반향 제거기로 계산량이 적고 구현이 용이하다는 장점 때문에 LMS 알고리즘을 적용하는 트랜스 버설 (Transversal) 형태의 적응 필터를 사용하였다. 그러나 일반적으로 수신룸상의 반향 현상은 수백 ms 동안 지속되며, 이를 제거하기 위해서는 고차의 FIR 형태의 필터를 필요로 한다. 결국 수렴성능의 저하가 유발되며 계산량의 증가로 인해 실시간 반향 신호 제거가 어렵게 된다. 이러한 과도한 계산량 및 수렴속도의 저하를 방지하기 위하여 서브밴드 형태의 적응 필터가 필요하다[5][6].

다른 문제점은 근단 화자 신호가 마이크로폰의 입력으로 들어갈 때 적응 필터는 근단 화자 신호를 반향 신호와 함께 추정을 한다는 것이다. 따라서 근단 화자 신호가 존재할 때 적응 필터는 계수 적응을 방지하여 근단 화자 신호의 손실을 방지해야 한다. 반향 제거기의 구성요소의 하나로서 동시 통화 검출기는 정확하고 빠르게 근단 화자 신호를 검출하여 근단 화자 신호로 인하여 원단화자 신호의 제거를 위한 적응 디지털 필터의 계수가 수렴상태로부터 변화되는 것을 방지해 주어야 한다.

III. 서브밴드 적응 필터를 이용한 반향제거 필터

음향 반향 제거 시스템에서 제거되어야 하는 반향 신호의 스펙트럼은 음성 신호에 국한되기 때문에 매우 다이나믹한 성질을 갖고 있다. 따라서 전 (full)밴드 적응

반향 제거기는 많은 필터 계수들의 계산량과 기준 신호의 다이내믹한 입력 스펙트럼으로 인한 느린 수렴속도의 문제가 발생하게 된다. 이러한 문제들을 해결하기 위해 최근 원격회의 시스템에 효율적인 음향 반향의 기술로 적당한 서브 밴드 처리방법이 연구되고 있다[7][8]. 일반적인 LMS 형태의 전밴드 적응 필터 방법과 비교해 볼 때 서브 밴드 처리방법은 적어도 두 가지 장점 즉, 병렬 처리가 가능하므로 실제 하드웨어로 구현할 경우 실시간 처리가 용이하며, 입력신호를 대역 분할하므로 LMS 알고리즘과 함께 사용하는 경우 고유치 분포율이 감소되어 수렴속도를 증가시킬 수 있다.

그림 2는 K 개의 밴드를 갖고 $K:1$ 로 간축한 병렬처리 구조를 서브 밴드 적응 디지털 필터 구조를 나타낸다. 서브밴드 처리기법을 이용한 적응 반향 제거 필터는 크게 세 부분으로 나뉘는데 먼저 원단 화자측 스피커로부터 출력된 신호가 마이크로 입력되는 주 입력 신호 $y(n)$ 과 스피커 출력전에 취득하게 되는 순수한 원단 화자에 의한 기준 입력신호 $x(n)$ 를 각각 K 개의 부밴드별로 분할하기 위한 합성단과 이들을 사용하여 각 부밴드별로 반향 성분 제거를 위한 적응 반향 제거 필터뱅크와 각 대역 적응 필터의 입력으로 사용하는 과정, 그리고 각 밴드의 오차를 합성필터에 통과시켜 전대역 오차신호(error signal)를 발생시키는 과정이다. 일반적으로 전대역 신호를 K 개의 부밴드로 분할하기 위한 효율적인 방안으로 폴리페이즈 필터뱅크가 사용되었다[5]. 이때 K -밴드 폴리페이즈 필터뱅크의 설계 시 프로토타입 저역 통과 필터의 주파수 대역폭은 f_s/K (f_s : 표본화주파수)를 갖는다. 따라서 이 경우에 Critical Sampling에 의하여 신호를 K 배 간축함으로써 Nyquist의 샘플링 이론에 어긋나지 않도록 한다. 폴리페이즈 방식을 이용하여 신호를 프로토타입 저역 통과 필터에 통과시킨 후에 DFT 필터뱅크를 통과시켜 각 대역별 신호를 한 번의 저역 통과 필터링으로 줄였을 뿐만 아니라 DFT를 FFT로 수행하는 경우 계산량을 효과적으로 줄일 수가 있다.

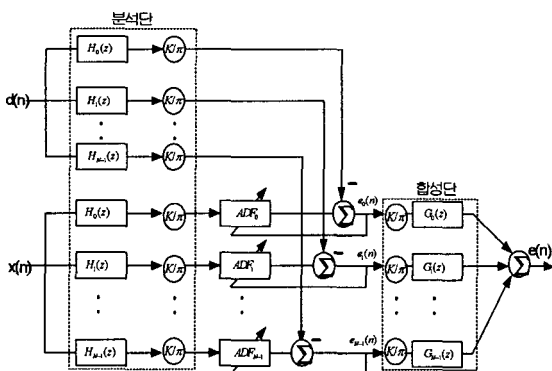


그림 2. 서브밴드 처리 기법을 적용한 적응 반향제거 필터의 구조
 Fig. 2. Structure of adaptive echo canceller which is based on subband process.

그림 2에서 반향 신호와 음향 반향 제거기에 입력되는 기준 입력 신호는 분석 필터 ($H_0(z), H_1(z), \dots, H_{(K-1)}(z)$)에 의해서 K 개의 인접한 주파수 서브 밴드로 나뉘어 진다. 여기서 K 개의 인접한 단-대역 필터의 출력들은 연속적으로 서브 밴드 신호를 만들어 내기 위해 K 라는 값으로 다운샘플링 된다. 적응 필터의 출력신호는 다시 K 이라는 값으로 업-샘플링 되며 합성 필터 ($G_0(z), G_1(z), \dots, G_{(K-1)}(z)$)에 의해서 반향 신호가 제거된 신호를 재구성된다.

IV. 근단 화자 검출

적용 필터는 기준 입력 신호와 오차 신호를 사용하여 반향 경로의 임펄스 응답을 추정하기 때문에, 근단화자 신호가 존재하는 경우에는 적응 필터의 필터 계수에 대한 추정 오차가 커지므로 오동작을 일으킬 수 있다. 그러므로 동시 통화 검출기(double-talk detection)를 설치하여 근단화자 신호가 존재하는 동안, 즉 동시 통화 기간에는 적응 필터의 계수 적응을 중단하여야 한다. 또한 반향 경로의 변화를 동시 통화 상황으로 잘못 판단하여 적응 필터의 계수 적응이 필요할 때 오히려 이를 중단하는 일이 없어야 한다. 동시 통화 검출을 위한 방법에는 기존의 에너지 비교 방법과 상호 상관 계수에 따른 비교법이 있다. 에너지 비교(energy comparison)에 의한 동시 통화 검출기는 마이크로폰 입력 신호와 원단화자의 크기 비교에 의한 식 (6)으로 표현된다. 계산량은 적지만 작은 크기의 근단 화자 신호에 대해서는 검출 능력이 떨어지는 단점이 있다.

$$|d(n)| \geq A \text{ Max } \{ |x(n)|, |x(n-1)|, \dots, |x(n-L+1)| \} \quad (6)$$

마이크로폰 입력 신호(근단 화자 신호+반향 신호)를 $d(n)$ 이라 할 때 원단화자는(기준 입력신호) $x(n)$ 이고, A 는 하이브리드의 감쇄를 고려하여 결정되는 상수이다. 작은 근단 화자신호에 대해 검출 지연이 없도록 근단 화자 검출을 위한 또 다른 방법은 상호 상관 계수(Cross-correlation)로 근단 화자 신호의 시작점을 검출하는 방법이다. 상관 신호는 적응 필터의 출력 신호 $\hat{y}(n)$ 와 마이크로폰 입력 신호인 $d(n)$ 신호와의 상호 상관 계수를 이용한다. 근단 화자의 신호가 없다면 마이크로폰의 입력 신호는 반향신호만 존재하게 된다. 근단 화자 신호가 존재한다면 반향 신호와 근단 화자 신호가 들어오게 된다. 이때 상호 상관 계수 값으로 근단 화자 신호를 구분 한다는 것이 상호 상관 계수를 이용한 근단 화자 신호 검출 방법이다. 그 식은 다음과 같다. $E[\cdot]$ 은 통계적 평균치를 나타낸다.

$$P_{\hat{y}, d} = \frac{E[\hat{y}(n)d(n)]}{\sqrt{E[\hat{y}^2(n)]E[d^2(n)]}} \quad (7)$$

여기서, $s(n)$ 는 근단 화자 신호이며, $y(n)$ 는 반향 신호를 나타낸다. 마이크로폰 입력 신호는 근단 화자 신호와 반향 신호로 구성된다.

$$d(n) = s(n) + y(n) \tag{8}$$

단일 채널 음향 반향 제거기가 수렴 상태에 도달했다고 가정하면 다음과 같은 관계가 성립한다. 근단 화자 신호가 없을시에 적응 필터의 계수 업데이트를 통하여 반향경로를 정확히 추정하였을 경우 다음과 같은 관계를 가진다.

$$\hat{y}(n) \cong y(n) \tag{9}$$

또한 원단 화자 신호와 근단 화자 신호가 서로 상관관계가 없다면 마이크로폰 입력신호 $d(n)$ 와 원단 화자 신호 $y(n)$ 과의 상호 상관값은 다음과 같다.

$$E[\hat{y}(n)d(n)] \cong E[y^2(n)] + E[s(n)y(n)] = \sigma_y^2(n) \tag{10}$$

근단 화자 신호와 단일 채널 음향 반향 제거기의 적응 필터의 출력 신호는 서로 상관 관계가 없고 평균값이 영이라고 가정하면, 마이크로폰 입력 신호의 자기 상관 값은 다음과 같다.

$$\begin{aligned} E[d^2(n)] &= E[s^2(n) + y^2(n)] \\ &= \sigma_s^2(n) + \sigma_y^2(n) : \text{원단화자가 존재할 경우} \\ &= E[y^2(n)] \\ &= \sigma_y^2(n) : \text{원단화자가 존재하지 않을 경우} \end{aligned} \tag{11}$$

근단 화자 신호의 분산은 항상 영보다 큰 값을 가지므로 근단 화자 신호가 존재 할 때와 존재 하지 않을 때의 상호 상관 계수 값은 항상 다음 관계를 만족한다. 결국, 상호 상관 계수 값은 다음과 같이 근단 화자의 유무에 따라 다른 값을 갖게 된다.

$$1 = \left[\frac{\sigma_y^2(n)}{\sigma_d^2(n)} \right] \geq \left[\frac{\sigma_y^2(n)}{\sigma_s^2(n) + \sigma_y^2(n)} \right]^{\frac{1}{2}} \tag{12}$$

위 식에서 나타난 것처럼 근단 화자 신호가 존재하지 않을 때의 상관 계수 값은 근단화자 신호가 존재할 때의 상호 상관 계수 값보다는 항상 상대적으로 큰 값을 가지므로 적당한 문턱치 (threshold)를 설정함으로써 근단 화자 신호를 검출 할 수 있다. 문턱치의 설정은 절대적인 값이 아니다. 근단 화자의 신호가 존재하지 않을 때는 상호 상관 계수 값이 1에 근접하게 된다. 근단 화자 신호가 존재하게 되는 경우에는 상호 상관 계수 값이 1이하로 떨어지게 되므로 문턱치 이하의 상호 상관 계수 값이 존재 할 때 근단 화자의 신호가 존재한다고 판단을 하고 적응 필터의 계수 추정을 멈추게 한다. 문턱치를 너무 높

게 설정을 하면 근단 화자의 작은 신호 부분을 검출하지 못하게 되고, 문턱치를 너무 낮게 설정을 하면 근단 화자 신호가 적응 필터에 의해 손상을 입게 된다.

V. 가중 Overlap-add 방법을 이용한 음향 반향 제거기

3장에서 소개한 부밴드별 적응 필터를 이용한 반향 제거기는 입력신호를 각 대역별로 나누어 NLMS를 사용한 음향 반향 적응 필터로 반향 신호를 각 밴드별로 제거함으로써 수렴 속도를 증가시킬 수 있다. 부밴드의 처리과정은 밴드별 병렬처리가 가능하여 실시간으로 단일 음향 반향 제거기를 구현 할 수 있다. 기존의 부밴드는 폴리페이즈 filter bank를 구성하여 부밴드별로 대역분할을 수행하며, 이때 입력신호가 filter bank와의 Convolution연산이 때문에 하드웨어로 실제 구현시 잦은 Shifting 동작으로 인해 복잡도가 증가하게 된다. 본 논문에서 음향 반향 제거 알고리즘으로 사용되는 가중 Overlap-add방식은 데이터를 블럭별 변환으로 처리를 하고 블럭의 신호를 DFT Filter bank를 통해 각 대역별로 분할을 한다. 표 1은 기존의 서스밴드 방식인 폴리페이즈 필터뱅크 방식과 가중 Overlap-Add 방식의 차이를 나타낸다.

표 1. 폴리페이즈 필터뱅크 방식과 가중 Overlap-add 방식의 비교

Table 1. Comparison of the polyphase filter bank and weighted overlap-add structure.

| | 폴리페이즈 필터 | 가중 Overlap-add |
|--------|--|---|
| 데이터 처리 | stream | block |
| 기능 | 순환처리 가능 | 병렬처리 가능 |
| 구조 | filtering 타입 | window 타입 |
| 연산 | convolution $K = MI, I \text{ integer}$ | multiplication+add $K = MI, I \text{ integer}$ 일 필요가 없음 |

이 구조는 DFT filter bank를 기본으로 블럭별 관점에서 신호를 분석하고 (합성) 한다. 가중 Overlap-add 구조는 입력신호를 블럭 단위로 대역을 분할하는 분석단과 각 부밴드별 NLMS (Normalized LMS)적응 필터를 사용하여 각 밴드의 반향 신호를 제거한다. 또한 일반적으로 분리된 부밴드간의 Cross-Talk현상을 방지하기 위하여, 과표본화를 적용하기 위해 부밴드수 K 와 간속비 M 을 서로 다르게 ($K > M$) 설정한다. 이때 폴리페이즈 필터방식은 부밴드수 K 가 간속비의 정수배가 되어야 한다는 조건이 있으나 가중 Overlap-add 구조는 이러한 정수배의 조건을 만족시킬 필요가 없으며 부밴드수가 간속비의 임의의 실수배가 되어도 상관없이 장점을 지닌다.[11] 그림 3은 K 번째 부밴드 신호를 생성하기 위한 Overlap-add 분석 구조를 나타낸다.

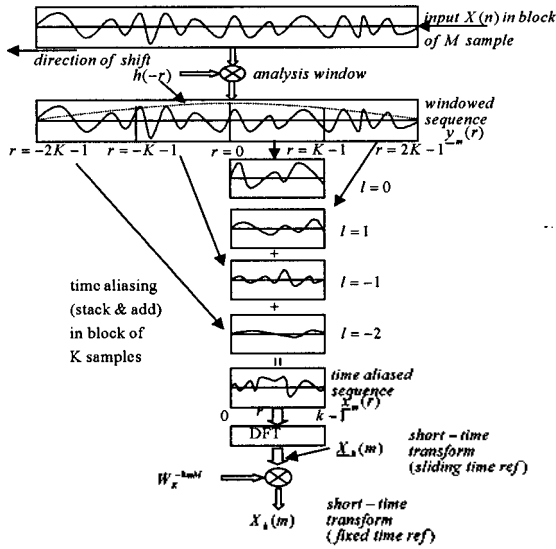


그림 3. 가중 Overlap-add 분석 구조에 의한 K번째 부밴드 신호
Fig. 3. Kth subband signal using the weighted overlap-add analysis.

일반적인 DFT Filter bank 분석단에서 K번째 밴드성분 $X_k(m)$ 은 다음과 같다[10].

$$X_k(m) = \sum_{n=-\infty}^{\infty} y_m(n) W_K^n \sum_{n=-\infty}^{\infty} h(mM-n)x(n) W_K^{-kn} \quad (13)$$

$k = 0, 2, 3, \dots, K-1$

식 (13)에서 M 은 슬라이딩 Shift샘플수 이고, $W_K^{-kn} = e^{-j\frac{2\pi kn}{K}}$ 이다. K 는 블록 안의 밴드 수를 나타내고 여기서 $h(n)$ 은 슬라이딩 분석창으로 입력신호 $x(n)$ 을 분석하기 위해 시간에 따라 M 샘플수 만큼씩 이동하며 입력신호와 곱해짐을 알 수 있다. 식 (13)에서 입력신호 $x(n)$ 을 슬라이딩 분석창에 의해 얻게 되는 신호 $y_m(n)$ 은 다음과 같이 나타낼 수 있다.

$$y_m(n) = h(mM-n)x(n) \quad (14)$$

따라서 $y_m(n)$ 은 블록 인덱스 m 에 따라 가중처리된 신호이며, 실제 수행하는 프로세서에서는 고정된 타임 프레임에서 슬라이딩 타임 프레임으로 변환이 필요하다. 이때 M 을 창함수의 슬라이딩 Shift간격이고 r 을 슬라이딩 타임 프레임 인덱스라고 하면 다음과 같이 표현된다.

$$r = n - mM \quad (15)$$

식 (15)를 식 (14)에 대입하여 식을 전개하면,

$$X_k(m) = \sum_{r=-\infty}^{\infty} h(-r)x(r+mM) W_K^{-k(r+mM)} = W_K^{-kmM} X_k(m) \quad (16)$$

식 (16)에서 $X_k(m)$ 은 슬라이딩 타임 프레임 인덱스 r 에 대한 관점에서의 일종의 short-time 푸리에 변환 형태를 가지며 다음과 같이 정의된다.

$$X_k(m) = \sum_{r=-\infty}^{\infty} h(-r)x(r+mM) W_K^{-kr} \quad (17)$$

슬라이딩 타임 프레임 분석창 $h(-r)$ 은 고정되어 있고 데이터가 오른쪽에서 왼쪽으로 슬라이딩되고 있음을 식 (17)에서 나타내고 있다. 식 (16)에서와 같이 선형적으로 증가하는 위상 부분 W_K^{-kmM} 은 고정된 타임 프레임 $X_k(m)$ 에 대해 슬라이딩 타임 프레임으로 인덱스 상에서 창함수에 의해 슬라이딩된 입력신호를 가중처리 하여 변환된 성분과 곱해지게 된다. 슬라이딩 타임 프레임에서 고정된 타임 프레임으로 변환하는 식 (16)은 간축된 신호를 창함수로 가중처리 되어 W_K^{-kmM} 에 의해 출력값을 변조함으로써 같아지게 된다. 즉, short-time 푸리에 변환에 관련되어진 슬라이딩 타임 프레임 $X_k(m)$ 은 W_K^{-kmM} 에 의해 변조되기 전에 간축된 복소 중간 대역 필터 $h_k(n)$ 에 의해 필터링된 결과 값과 같다. 슬라이딩 타임 프레임 해석은 short-time spectral analysis를 block 변환의 관점에서 부가적으로 얻어진다. 이런 특별한 실행으로 식 (17)의 슬라이딩 변환은 FFT (fast Fourier transform) 같은 빠른 알고리즘을 통하여 효과적으로 실현 될 수 있다. 슬라이딩 타임 프레임 인덱스 r 상에서 변환되는 샘플수는 사실 상 FFT를 위해 요구되는 K 개의 샘플 데이터 보다 많으므로 K -point FFT 형태가 식 (17)같지 않다. 이 문제는 시간상에서 신호를 에일리어징 하는 방법을 거쳐서 FFT를 계산한다. 슬라이딩 타임 프레임 인덱스 r 에 대한 window와 가중된 입력 신호를 식 (18)으로 표현하면,

$$y_m(r) = h(-r)x(r+mM) \quad (18)$$

식 (18)을 식 (17)에 대입하면 슬라이딩 타임 프레임 인덱스 r 에 대하여 나타낼 수가 있다.

$$X_k(m) = \sum_{r=-\infty}^{\infty} y_m(r) W_K^{-kr} \quad (19)$$

시퀀스 $y_m(r)$ 는 window와 대응되는 신호를 K 샘플씩 순환하며 더해지는 과정을 통해 K -Point 시퀀스로 나뉘어진다. 이때 그 블록들의 형태는 다음과 같다.

$$y_m(r) |_{r=r+K} = x_m(\rho) = \sum_{\rho=-\infty}^{\infty} y_m(\rho+K), \quad \rho=0, 1, 2, \dots, K-1$$

$$= \sum_{\rho=-\infty}^{\infty} h(-\rho-K)x(\rho+K+mM) \quad (20)$$

슬라이딩 타임 프레임 인덱스 r 에서 l 은 $y_m(r)$ 신호는 K 개 샘플들로 블록화 할 경우 블록 인덱스가 된다. 결국,

식 (20)은 창함수에 의해 가중 처리된 신호를 K -Point 블록별로 나누어 처리하는 과정이다. 따라서 식 (19)는 다음과 같은 FFT처리과정과 동일하게 된다. FFT를 거쳐 대역별로 분할을 하면 가중 overlap-add 구조의 분석단의 출력 신호이자 동시통화 검출기의 입력 신호가 된다. 같은 방법으로 반향 신호를 처리한다. 근단 화자의 신호가 존재하지 않는다면 합성단을 통과하기 전에 반향 경로를 추정하여 단일 채널 음향 반향 제거기가 통화의 음질을 개선시킬 수 있다. 부밴드의 분석단과 합성단에 반향 제거기를 설치하면 수렴속도를 향상시킬 수 있기 때문이다.

$$X_k(m) = \sum_{p=0}^{K-1} x_m(p) W_k^{-kp} \quad (21)$$

식 (21)은 K -Point FFT 알고리즘을 위한 적합한 형태가 되며, $h(n)$ 은 유한한 길이를 갖는 FIR filter 형태이므로 식 (20)도 유한한 블록들의 합으로 볼 수 있다. 블록 타임 시간 인덱스 m 에 따른 동작을 하기 위해서 입력 신호 $x(n)$ 은 M 만큼 Shift되어 가중처리 된 후 FFT 처리를 거친다. N_s 는 sample shift 레지스터이며 분석창 $h(n)$ 의 탭수이다. Window의 크기는 $N_w = mK$ 가 되며, 만약 분석창의 크기가 128이라고 가정하면 $K = IM$ 에서 $I = 2$ 인 ($K = 32 - Point$)경우에 총 8번의 window Shifting 과정을 거치면서 입력신호 128샘플에 대한 분석이 가능해진다. Shift 레지스터의 테이터는 식 (18)에 따라 short-time 시퀀스 $y_m(r)$ 을 구할 수 있다. 이때 $y_m(r)$ 은 블록 K 샘플로 K -Point DFT를 통하여 $X_k(m)$ 는 슬라이딩 타임 프레임이 된다. 최종적으로 W_k^{-kmM} 에 의해 대역별로 다중화 된 신호는 원하는 고정 타임 프레임 $X_k(m)$ 이 된다. 타임 블록 시간 m 에 따라 새로운 M 만큼의 입력 데이터를 같은 방법으로 처리한다.

$X_k(m)$ 의 신호는 K 샘플수의 블록으로 블록타임 m 에 따라 순차적으로 부밴드 반향 제거기의 각각의 입력 신호가 된다. 그림 3에서 마이크로폰 입력 $d(n)$ 도 마찬가지로 슬라이딩 타임 프레임을 FIR계수의 windowing을 통해서 time-에일라이징 되고 K -Point FFT를 수행하고 W_k^{-kmM} 에 의해서 고정 타임 프레임으로 변환한다. 반향신호와 기준신호가 부밴드 수만큼의 샘플수로 구성되는 각 블록들이 적응 필터의 입력들로 사용되며 이에 대한 반향 신호 제거 오차 성분도 각 밴드별로 K 개 만큼씩 생성된다.

또한 근단 화자 전용 필터를 이용한 동시 통화 검출기(double-talk detection)를 이용하여 근단 화자 신호 검출 시 계수의 적응을 방지하여 근단 화자 신호 성분을 추정하지 못하게 한다. 단일 채널 음향 반향 제거기에 입력은 $x(n)$ 과 반향신호 $y(n)$ 는 K -Point FFT를 거쳐 변조가 중치 W_k^{-kmM} 로 곱혀져서 다중화 된 채널의 각 밴드로 구성된 1-by- K 인 벡터의 형태이다. NLMS역시 벡터의

연산으로 수행하게 된다. 적응필터의 구성 역시 각 밴드별로 수렴을 해야 하기 때문에 밴드 수에 따른 적응 필터가 필요하게 된다.

$y(n)$ 는 실제 반향 경로 시스템을 통과하여 얻은 스피커의 출력에 의해 만들어지는 신호이며 이런 반향 신호의 특성을 제거하기 위해 적응 필터는 실제 반향 경로를 추정하여 전송실로 되돌아가는 반향 신호를 제거한다. 이때 $\hat{y}(n)$ 은 각 부밴드에서 적응 필터를 통한 추정된 반향 신호이며 다음과 같이 표현된다.

$$\hat{y}(n) = \text{diag}(\hat{W}(n))^T X(n) \quad (22)$$

식 (22)에서 $\text{diag} A$ 는 $N \times N$ 행렬 A 의 diagonal 성분만으로 구성된 $N \times 1$ 벡터이다. 적응 필터의 가중벡터는 입력 신호와 콘벌루션이므로 각 밴드별로 현재 값을 추정하기 위해서 과거의 값이 필요하다. 여기서 T 는 전치행렬, 탭수 L 인 적응 필터의 계수와 입력 신호를 다음과 같이 나타낼 수 있다.

$$\hat{W}(n) = [w_1(n), w_2(n), \dots, w_K(n)] \quad (23)$$

$$w_i(n) = [w_{i1}(n), w_{i2}(n), \dots, w_{iL}(n)]^T, i = 1 \rightarrow K \quad (24)$$

$$\hat{X}(n) = [x_1(n), x_2(n), \dots, x_K(n)] \quad (25)$$

$$x_i(n) = [x_i(n), x_i(n-1), \dots, x_i(n-L+1)]^T, i = 1 \rightarrow K \quad (26)$$

적용 필터는 각 밴드별의 수 만큼 필요하게 되고 기준 입력 신호 역시 밴드별 블록 단위로 수행을 하기 때문에 그 값을 담아놓을 수 있는 버퍼를 사용한다. 버퍼의 크기는 입력 신호 벡터와 NLMS의 필터 탭 수를 고려하여 설정한다. 만약, 512탭 수로 이를 32밴드별로 처리한다면 버퍼의 크기는 32-by-16이 된다. 따라서 적응 필터의 각 밴드별 탭 수는 16 이다. i 번째의 부밴드 처리 과정에서 생성되는 잔여 반향 에러 신호는 다음과 같이 나타낼 수 있다.

$$e(n) = [e_1(n), e_2(n), \dots, e_K(n)] \quad (27)$$

$$e(n) = y(n) - \hat{y}(n) = y(n) - \text{diag}(W^T(n)) X(n) \quad (28)$$

에러 역시 각 밴드별로 잔여 반향 신호를 최소화하기 위한 적응 필터를 구성해야 한다. 32밴드별 적응 필터의 계수 업데이트 식은 다음과 같다.

$$W^T(n) = \begin{bmatrix} w_1^T(n) \\ w_2^T(n) \\ \vdots \\ w_K^T(n) \end{bmatrix} = \begin{bmatrix} w_1^T(n-1) + \mu e_1(n) x_1^T(n) / Z_1 \\ w_2^T(n-1) + \mu e_2(n) x_2^T(n) / Z_2 \\ \vdots \\ w_K^T(n-1) + \mu e_K(n) x_K^T(n) / Z_K \end{bmatrix} \quad (29)$$

결국 각 밴드별로 에러가 존재하고 입력신호를 저장한 버퍼에 의해 계수 업 데이트가 진행된다. Z_l 는 l 번째 부밴드 입력신호 블록에 대한 이득을 나타낸다. NLMS를 사용한 전 밴드 필터는 큰 에너지를 갖는 주파수 성분부터 수렴을 하기 때문에 상대적으로 에너지가 작은 신호에 대해서는 수렴이 저하된다. 부밴드를 적용한 단일 음향 반향 제거기를 사용하는 두 가지 특징은 첫째, 병렬처리 (블록별 처리)를 통한 실시간 데이터의 처리와 둘째, 입력신호를 대역 분할하므로 NLMS 알고리즘과 함께 사용할 경우 고유치 분포율이 감소하여 수렴속도를 증가시킬 수 있기 때문이다. 기존의 폴리페이즈 필터 बैं크와 비교하여 가중 Overlap-add 구조는 블록 연산을 한다는 점과 에일리어징을 방지하기 위해 over sampling factor I 가 2^n 의 형태가 될 필요가 없다. 또한 필터의 convolution 연산이 아닌 window를 사용한 데이터의 처리로 복잡도를 줄였으며, stream 연산이 K 개의 샘플씩 블록 처리할 수 있다는 장점이 있다. 제안한 가중 overlap-add 적용 반향 제거기의 합성과정은 분석과정의 역 (반전)으로 수행한다.

$$\hat{x}(n) = \sum_{l=-\infty}^{\infty} f(n-lM) \frac{1}{K} \sum_{k=0}^{K-1} \hat{X}_k(l) W_K^{kn} \quad (30)$$

식 (30)식은 NLMS의 적응 필터를 통과하여 반향 신호를 제거한 성분을 나타낸다. 슬라이딩 타임 프레임 관점에서 식 (30)를 변형하면, $r=0, 1, 2, \dots, M-1$

$$r = n - lM, \quad n = r + lM \quad (31)$$

식 (31)을 식 (30)에 적용하면 다음과 같다. $m = m_0$ 일 경우

$$\hat{x}(r + m_0 M) = f(r) \frac{1}{K} \sum_{k=0}^{K-1} \hat{X}_k(m_0) W_K^{km_0 M} W_K^{kr} + (\text{term for } l \neq m_0) \quad (32)$$

식 (32)의 오른쪽 식은, $\hat{x}(r + m_0 M)$ 은 m_0 일 때의 short-time 스펙트럼 $\hat{X}_K(m_0)$ 에 의한 출력이며 블록 타임 l 에 따라 $l = m_0$ 일 때와 $l \neq m_0$ 부분의 합으로 이루어진다. 고정된 시간 프레임에서 슬라이딩 타임 프레임으로 $\hat{X}_K(m_0)$ 을 변환하면, 식 (33)이 되며, $\hat{x}_{m_0}(r)$ 을 $\hat{X}_K(m_0)$ 의 Inverse DFT 라고 하면 식 (34)로 나타낼 수 있다.

$$\hat{X}_K(m_0) = \hat{X}_K(m_0) W_K^{km_0 M} \quad (33)$$

$$\hat{x}_{m_0}(r) = \frac{1}{K} \sum_{k=0}^{K-1} \hat{X}_k(m_0) W_K^{kr} \quad (34)$$

Inverse DFT를 식 (32)에서 따로 분리하여 나타내면

$$\begin{aligned} & \frac{1}{K} \sum_{k=0}^{K-1} \hat{X}_k(m_0) W_K^{k(r+m_0 M)} \\ &= \frac{1}{K} \sum_{k=0}^{K-1} (\hat{X}_k(m_0) W_K^{km_0 M}) W_K^{kr} \quad (35) \\ &= \text{IFFT}(\hat{X}_k(m) W_K^{km_0 M}) \end{aligned}$$

결국 식 (35)를 이용하여 블록 타임 m_0 를 m 으로 변환하면 식 (32)는 다음과 같이 표현할 수 있다.

$$\begin{aligned} \hat{x}(r + mM) &= \{f(r) \cdot \text{IFFT}(\hat{X}_k(m) W_K^{kmM})\} \\ &+ \sum_{l \neq m}^{\infty} f(r + mM - lM) \frac{1}{K} \sum_{k=0}^{K-1} \hat{X}_k(l) W_K^{k(r+mM)} \end{aligned} \quad (36)$$

$$\begin{aligned} f(r) \cdot \text{IFFT}(\hat{X}_K(m) W_K^{kmM}) &= f(r) \text{IFFT}\{\hat{X}_k(m)\} \quad (37) \\ &= f(r) \hat{x}_m(r) \end{aligned}$$

식 (35)-식 (37)에 의해서 합성단 출력은 블록 타임 m 에 따라 $r=0, 1, 2, \dots, M-1$ 신호를 합성한다. 다음식은 DFT filter bank를 통해 슬라이딩 타임 프레임의 신호를 고정 타임 프레임으로 변환한 식 (30)를 간단히 나타내고 있다.

$$\begin{aligned} \hat{x}(r + mM) &= f(r) \hat{x}_m(r) \\ &+ \sum_{l \neq m}^{\infty} f(r + mM - lM) \frac{1}{K} \sum_{k=0}^{K-1} \hat{X}_k(l) W_K^{k(r+mM)} \\ &r = 0, 1, 2, \dots, AK-1 \end{aligned} \quad (38)$$

식 (33)을 통해서 Short-time 변환 $\hat{x}_k(m)$ 은 선형적으로 증가하는 위상 변조 W_K^{kmM} 를 통해 변환한다. IFFT를 통해 $\hat{x}_m(r)$ 에 해당하는 K -sample 시퀀스를 만든다. 식 (34) $\hat{x}_m(r)$ 는 주기적으로 $f(r)$ 의 크기만큼 생성된 후 windowing 된다. 이 과정이 그림 3에서의 분석 window의 시간적 aliasing의 과정에 대응된다. 그 결과 블록타임 m 마다 생성되는 시퀀스 $f(r) \hat{x}_m(r)$ 이 된다. $f(r)$ 은 프레임 단위의 합성창이 되며, 블록 타임 m 에 따라 각각의 새로운 블록은 출력에 더해진다. window의 길이는 분석창의 길이와 같으며 windowing된 신호는 N_f/M 부분 만큼 출력 $\hat{x}(m)$ 에 대응된다. 그림 4는 블록 타임 m 에서 DFT filter bank 합성단에 대한 가중 Overlap-add 구조의 합성과정이다.

블록타임 m 에 따라 M 샘플씩 Shift 한 자리엔 zero를 오른쪽에 넣어 주며, 각각의 블록 타임 m 에서 새로운 terms $f(r) \hat{x}_m(r)$ 은 병렬 형태로 shift register의 성분들을 더한다. 그리고 M 샘플만큼 shift를 한다. $f(r)$ 에 의해 시퀀스 $\hat{x}_m(r)$ 가 가중 처리된 후 Overlapping 되고 shift register에 의해 그 성분들이 더해지며, 이때 Shift register의 길이는 $f(r)$ 과 N_f 가 같아야 한다. 합성창의 범위를 넘어 shift register를 지나간 데이터도 식 (33)에

의한 유한한 블록의 합으로 완전해진다. 출력 $\hat{x}_m(r)$ 은 shift register에서 왼쪽 M 샘플만큼 갖고 데이터는 shift 된다. 분석구조의 경우와 같이 합성구조 역시 필터의 구성에 따라 채널을 나누고 IFFT 알고리즘을 이용하여 변조하는 DFT filter bank의 합성을 효과적으로 실행한다.

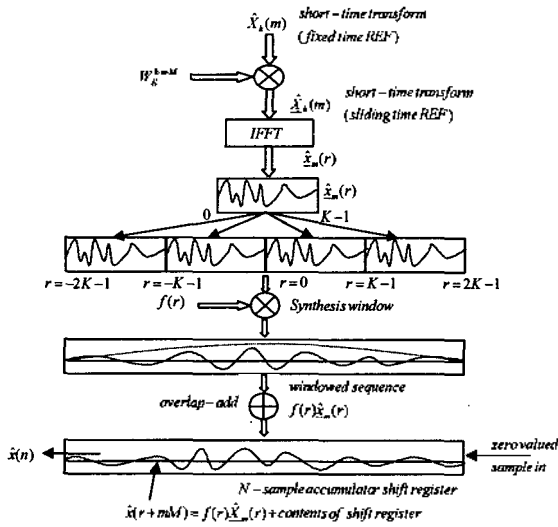


그림 4. 블록 타임 m 에 따른 합성 과정
Fig. 4. Synthesis process due to block time m .

VI. 제안한 전용필터를 사용한 동시통화 검출 기법

본 논문에서 제안한 동시 통화 검출 방법에선 근단 화자의 신호만을 검출 할 수 있도록 전용 필터를 구현하여 근단 화자 신호를 좀더 정확하고 빠르게 검출하고 반향 경로의 변화와 동시 통화를 짧은 구분 지연 시간 내에 구분하고 적응 필터의 수렴 여부에 무관한 검출 방법을 위해 두 개의 적응 필터를 사용하였다. 전 단의 적응 필터는 반향 신호만을 제거하기 위한 전용 필터로서 후 단의 근단 화자 전용 필터에서 근단 화자 신호를 검출하면 즉시, 전단의 적응 필터의 계수 업데이트를 중단시켜 근단 화자 신호의 손실이 없도록 하였다. 또한 근단 화자 신호를 검출하는 방법에서 32개의 밴드 중에서 첫 번째 밴드로 근단 화자 신호의 검출 유무를 판단하기 때문에 계산량이 감소한다. 근단 화자 신호를 검출하는 방법은 근단 화자 전용 필터 $D(n)$ 를 사용하여 근단 화자 신호의 유무를 판단하여 전단의 적응 필터 $W(n)$ 의 계수 적응을 방지한다. 다음 그림은 가중 overlap-add 구조에서 근단 화자 검출을 위한 동시 통화 검출기를 나타낸다.

$x(n)$ 은 원단 화자 신호이며, $y(n)$ 은 마이크로폰 입력 신호이다. 전단 필터 $W(n)$ 의 경우 반향 경로를 따라서 반향신호를 추정하게 되고, 후단의 필터 $D(n)$ 는 근단 화자의 문턱치를 결정하기 위해서 마이크로폰의 입력 신호를 추정한다.

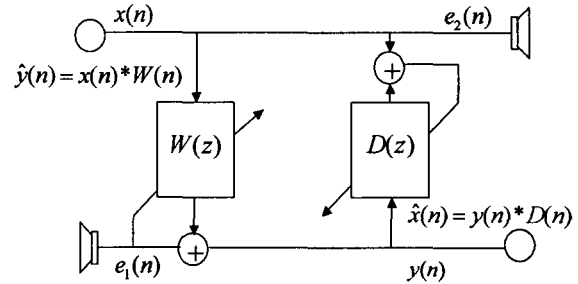


그림 5. 근단 화자 검출을 위한 단일 채널 적응 전용 필터
Fig. 5. Single channel adaptive dedicated filter for double talk detection.

$$\hat{x}(n) = \text{diag}(D^T(n) Y(n)) \tag{39}$$

근단 화자 전용 필터 $D(n)$ 에서 입력신호를 추정한 신호이며 필터에 따른 에러를 구해보면 다음과 같다.

$$e(n) = x(n) - \hat{x}(n) = x(n) - \text{diag}(D^T(n) Y(n)) \tag{40}$$

근단 화자의 검출은 상호 상관 계수 값을 이용하여 다음과 같이 One-Loop LPF (Low Pass Filter)[3]를 사용한 순환식으로 추정할 수 있다. 본 논문에서는 반향 제거용 필터 $W(n)$ 과 동시 통화 검출용 필터 $D(n)$ 의 추정 신호들로 표현된 다음의 지수들에 의해 구해진다.

$$P_x(n) = \lambda P_x(n-1) + (1-\lambda) x^2(n) \tag{41}$$

$$P_y(n) = \lambda P_y(n-1) + (1-\lambda) y^2(n) \tag{42}$$

$$P_{\hat{y}_{e_1}}(n) = \lambda P_{\hat{y}_{e_1}}(n-1) + (1-\lambda) \hat{y}(n) e_1(n) \tag{43}$$

$$P_{\hat{x}_{e_2}}(n) = \lambda P_{\hat{x}_{e_2}}(n-1) + (1-\lambda) \hat{x}(n) e_2(n) \tag{44}$$

여기서 forgetting factor인 λ 의 범위는 $0.9 < \lambda < 1$ 이다. 식 (41)~(44)에서 $P_x(n)$ 은 적응 필터 원단화자 신호에 대한 전력 값이며, $P_y(n)$ 은 마이크로폰 입력신호에 대한 전력 값이고, $P_{\hat{y}_{e_1}}$ 은 $W(n)$ 에서 추정된 신호와 에러의 correlation 전력 값이며, $P_{\hat{x}_{e_2}}$ 은 $D(n)$ 에서 추정된 신호와 에러의 상관 전력 값이다. 동시 통화 검출기는 근단 화자 신호의 존재 여부를 빠르고 정확하게 감지 할 수 있어야 한다. 제안한 방법은 두 개의 적응 필터를 구현함으로써 전단의 적응 필터는 반향 경로만을 추정하고 후단의 전용 필터는 근단 화자를 검출하기 위한 검출기로서의 기능을 갖게 된다. $D(n)$ 에서 추정된 $\hat{x}(n)$ 은 근단 화자 신호가 존재하지 않는 경우에 원단 화자 신호

$x(n)$ 을 추정하게 된다. 여러와의 상관계수를 원단화자로 나뉜 일정한 문턱치를 설정할 수 있게 된다. 근단 화자 신호와의 상관 관계가 전혀 없기 때문에 가능한 것이다.

$$R = - \left[\frac{P_{\hat{y}e_1}}{P_y} \right] \cdot \left[\frac{P_{\hat{x}e_2}}{P_x} \right] \quad (45)$$

후단의 근단 화자 전용 적응 필터는 계속해서 필터 계수를 적용하며 근단 화자 신호의 입력시 전단 필터의 계수 적용을 중지하여 근단 화자 신호 성분의 손실을 방지한다.

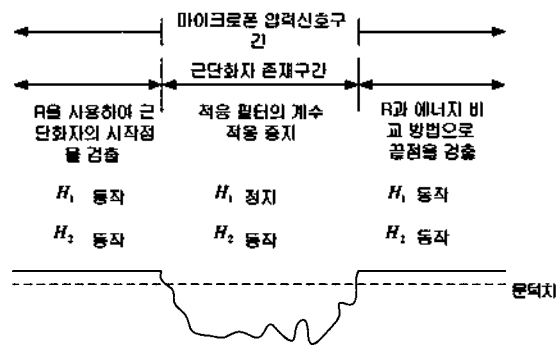


그림 6. 제안된 근단 화자 검출 구성도
Fig. 6. Structure of the proposal double talk detector.

동시 통화 검출기의 목적은 원단 화자 신호가 반향 경로를 통해 다시 마이크로폰으로 유입되는 상황에서 단일 채널 음향 반향 제거기의 적응 필터를 통해 반향 신호가 제거하고 근단 화자 신호가 마이크로폰에 입력되면 그때 적응 필터의 계수 적용을 중지시켜 근단 화자 신호의 손실을 억제하기 위함이다. 결론적으로 근단 화자의 신호를 얼마나 빠르고 정확하게 반향 경로의 변화와 구분하여 어떻게 검출을 하는가 하는 문제이다. 본 논문은 근단 화자 신호를 검출할 수 있는 전용 필터를 구현하여 근단 화자 신호의 검출을 좀더 효율적이고 정확하게 검출할 수 있는 방안을 제안하였다.

VII. 모의 실험

이 장에서는 제안된 방법의 우수성을 입증하기 위해 AR 필터에 통과 시켜 만든 유색 잡음 신호와 실제 음성 신호를 통해, 단일 채널 음향 반향 제거기의 적응 알고리즘에 따른 성능 분석 등을 컴퓨터 모의실험을 통해 입증한다. 또한 근단 화자 전용 필터를 이용하여 근단 화자의 시작점을 검출하고 에너지 비교와 상호상관 계수값을 이용하여 근단 화자 신호의 끝점을 검출하여 적응 필터의 계수 적용을 방지한다. 먼저 제안한 가중 overlap-add 적응 반향 제거기의 성능상의 향상을 보이기 위해 ERLE 곡선을 이용하였다. 컴퓨터 모의 실험에서 사용한 유한한 임펄스 응답들은 8kHz로 샘플링 되었고 크기 $4m \times 5m \times 3m$ 인

실제의 사무실에서 측정된 실제 임펄스 응답 시퀀스 중 앞부분 256샘플만을 추출하여 롬 전달함수의 임펄스 응답으로 사용하였다.

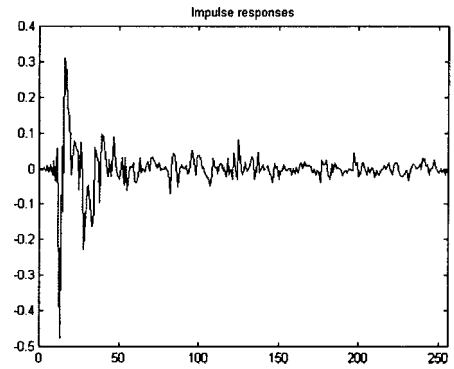
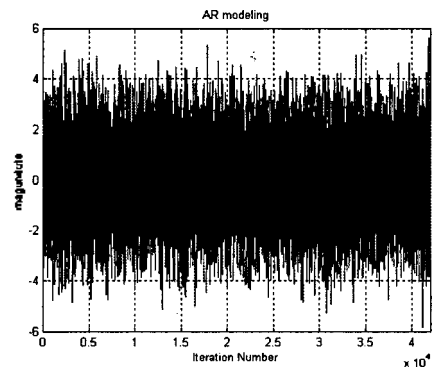


그림 7. 모의 실험에 사용한 임펄스 응답
Fig. 7. Impulse responses for simulation.

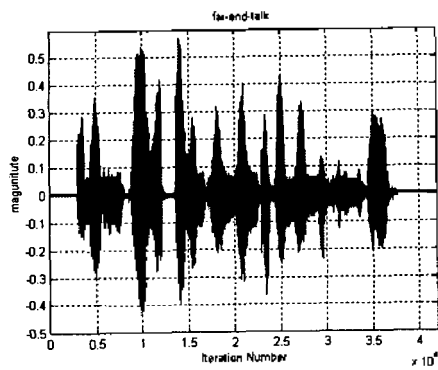
그림 7은 수신실의 반향 경로인 H 를 나타낸다. 모의 실험에서 사용한 원단 화자 신호 $x(n)$ 으로 유색잡음과 음성신호를 사용하였다. 유색잡음은 평균이 0이고 분산이 1인 백색 가우시안 잡음 (white gaussian noise) 신호를 음성의 스펙트럼 특성과 유사한 AR (auto-regressive) 필터에 통과시켜 만든 신호이다. 또한 음성 신호는 11kHz로 샘플링한 것이다. 모의 실험에 사용한 반향신호 $y(n)$ 은 원단 화자 신호를 그림 7의 임펄스 응답과 컨볼루션을 통해 신호를 생성하였다. 그림 8의 (a)는 원단화자 신호로 유색잡음 신호이며 (b)는 음성 신호를 나타낸다. (c)는 반향 신호를 나타내고 (d)는 근단 화자 신호를 나타낸다.

모의 실험에서 사용한 성능 평가는 음향 반향 제거기의 성능 평가에서 널리 사용되는 ERLE(Echo Return Loss Enhancement)값으로 비교 하였다.

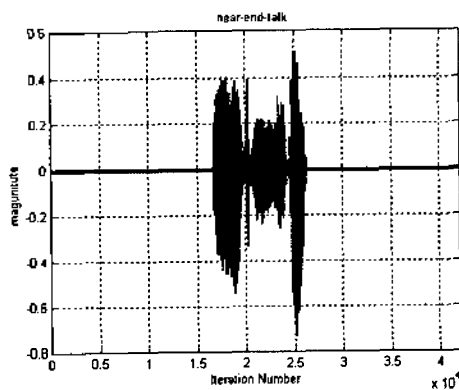
$$ERLE(n) = 10 \log_{10} \frac{\hat{\sigma}_y^2(n)}{\hat{\sigma}_e^2(n)} [dB] \quad (46)$$



(a)



(b)



(c)

그림 8. 모의 실험에 사용된 신호
(a) 유색잡음 신호 (b) 원단 화자 신호
(c) 근단 화자 신호

Fig. 8. Signals for simulation,
(a) Colored noise signal, (b) Far-end-talk,
(c) Near-end-talk.

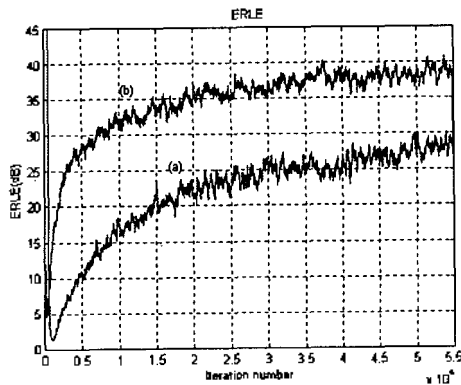
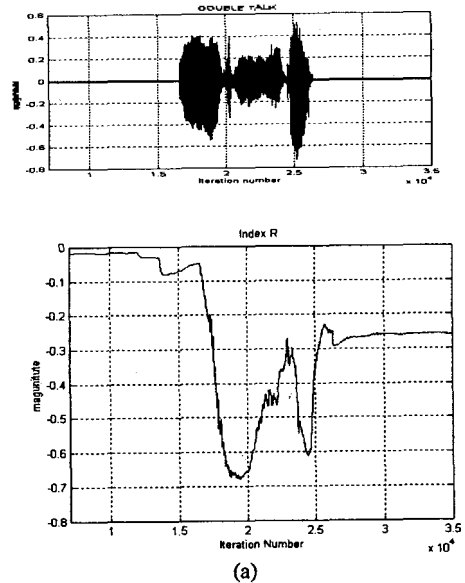


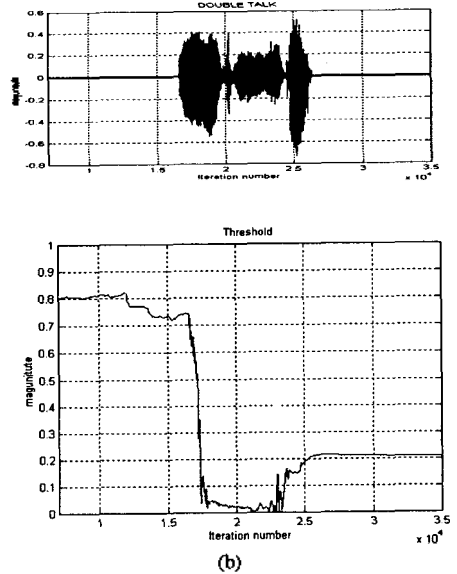
그림 9. 유색 잡음에 대한 ERLE 곡선
(a) 전대역 적응 반향 제거기
(b) 제안한 부분대역 적응 반향 제거기

Fig. 9. ERLE curve for colored noise,
(a) Full band adaptive echo canceller,
(b) Proposal subband adaptive echo canceller.

그림 9의 (a)은 유색잡음 신호에 대한 전대역 필터의 LMS알고리즘을 이용한 적응 필터의 ERLE를 나타내고 (b)는 제안한 가중 Overlap-Add 방식을 이용한 적응 반향 제거기의 ERLE를 나타내고 있다. 기존의 적응 반향 제거기 보다 성능이 우수함을 알 수 있다. 이 때 기존의 적응 반향 제거기의 수렴계수는 0.8이고 제안한 방법에서 사용한 수렴계수는 0.7이다. 서브 밴드의 구조는 전 밴드 구조와 비교 할 때 같은 NLMS 알고리즘으로도 향상된 수렴 특성을 보인다. 제안한 방법은 전용필터의 입력신호와 출력신호 마이크로 입력신호들의 상관 계수값들로 근단 화자 신호의 시작점을 검출하고 에너지 비교 방법을 동시에 사용하여 근단 화자 신호의 끝점을 검출한다.



(a)



(b)

그림 10. 근단 화자 검출 문턱치와 상관 계수
(a) 전용 필터를 이용한 문턱치
(b) 기존의 방법을 이용한 문턱치

Fig. 10. Threshold for near-end-talk detector and coefficient,
(a) Threshold for using dedicated filter,
(b) Threshold for using conventional method.

그림 10은 근단 화자 신호가 존재하는 구간이 문턱치보다 작은 값으로 변화하는 것을 알 수 있다. 그림 (a)는 근단 화자 전용 필터를 통한 상호 상관 계수 값이다. 그림 (b)는 기존의 방법으로 근단 화자 신호가 존재하지 않을 때의 상관 계수 값은 근단화자 신호가 존재할 때의 상호 상관 계수 값보다는 항상 상대적으로 큰 값을 가진다는 방법을 이용한 것이다. 그림 8(c)의 근단 화자 신호와 비교할 때 제안한 방법에서 근단 화자 신호 검출에 검출 지연이 나타나지 않았음을 알 수 있다. 또한 근단 화자 신호의 끝점을 에너지 비교와 상관계수로 검출하여 반향 신호에 대한 추정을 계속해서 할 수 있도록 하였다.

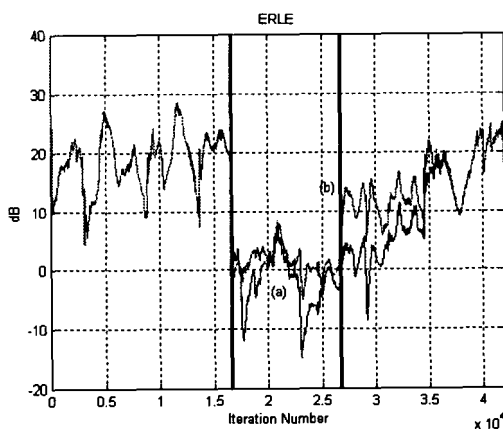


그림 11. 동시 통화 구간을 포함한 ELRE 비교
(a) 기존의 방법 (b) 제안한 방법
Fig. 11. Comparison of ERLE include double talk period,
(a) Conventional method, (b) Proposal method.

그림 11의 (a)는 기존의 상호 상관 계수로 근단 화자 검출을 했을 경우이며 (b)는 전용필터를 이용하여 근단 화자 신호를 검출했을 때 각각의 ERLE를 비교한 그림이다. 동시 통화 구간 이후의 수렴속도가 상호 상관 계수를 이용한 동시 통화 검출 방법 보다 개선되었음을 알 수 있다. 결국, 근단 화자 신호 검출 후 반향 신호를 효과적으로 제거한다는 것을 보인다.

VIII. 결 론

본 논문은 부밴드별 반향제거 필터 뱅크를 구현하기 위한 가중 Overlap-add 적용필터를 이용한 단일 채널 음향 반향 제거기와 전용 필터를 이용한 동시 통화 검출기에 대해서 제안하였다. 음성 신호와 같이 고유치 분포율이 큰 신호에 대해 기존의 TDL구조를 갖은 LMS 알고리즘 구조는 자주 변화하는 수신통의 환경에 민감하게 추정하지 못한다. 또한 근단 화자 신호가 존재하는 경우 음향 반향 제거기를 정상적으로 동작시키기 위해 효과적인 동시통화 검출기가 필요하게 된다. 따라서 제안한 단일 음향 반향 제거기는 수렴 속도의 향상과 계산량을 줄이

기 위해서 가중 Overlap-add 구조를 이용한 부 밴드 적용 알고리즘이다. 이는 NLMS 알고리즘과 서브 밴드 알고리즘을 결합한 형태로 각 밴드별로 필터 계수의 갱신을 수행한 후 이 필터 계수로 반향 신호를 추정하여 반향 신호를 제거하는 형태로 되어 있다. 이 필터 계수들은 특정 대역내의 시간 영역에서의 값들이므로 FFT를 수행하여 주파수 영역으로 변환한 후 동작 대역에 해당하는 위치에 주파수를 재배치한다. 근단 화자 전용 필터를 사용하여 반향 경로를 추정한 필터와 독립적으로 동작을 하여 근단 화자 검출시 계수의 적용을 방지하여 근단 화자 신호의 손실을 최소한으로 줄였으며 근단 화자 검출에도 검출지연이 발생하지 않았음을 보였다.

감사의 글

본 연구는 "한국과학재단의 99특정 기초 사원지원" 과제의 일환으로 수행된 연구 결과이며, 한국과학재단 관계자와 관련 연구원의 도움에 감사 드립니다.

참 고 문 헌

1. S.H.Jensen, "Acoustic Echo Canceller for Hands-Free Mobile Radio telephony," Proc.EUSIPCO, Belgium, pp. 1629-1632, 1992.
2. K. Murano, S. Unagami and F. Amano, "Echo Cancellation and Applications," IEEE Comm. Mag., pp. 49-55, Jan. 1990.
3. 김남선, 임용훈, 유재하, 차일환, 윤대회, "향상된 수렴속도와 근단화자 신호 검출능력을 갖는 적응 반향제거기," 신호처리 학술대회, vol.5, no.1, PP 159-163, sep. 1992.
4. 이행우, 은명희, 김종교, "음향반향제거기의 동시통화처리 알고리즘," 한국음향학회지 1998, 4 v.17, n.3, pp. 10-16 ISSN 1225-4428.
5. Stephan Weiss and Robert W.Stewart, "On The Optimality Of Subband Adaptive Filters," Proc. 1999 IEEE Workshop, New Paltz, New York, Oct. 17-20, 1999.
6. M.Taherzeshadi and J. LIU "A Subband Approach To Adaptive Acoustic Echo Cancellation," computer.Elect.Engng.vol. 23, NO.4, pp 205-215, 1997.
7. B.widrow and S.D.Stearns. Adaptive Signal Processing, Prentice-Hall Inc., 1985.
8. H.Perez and F.Amano, "A New subband Echo canceler structure," trans. of the EICE.vol.E73, PP. 1625-1631, oct. 1990.
9. 류차희, 유재하, 차일환, 윤대회, "SUB-BAND 적용 디지털 필터 실시간 시스템 구현에 관한 연구," 한국음향학회지 1993 v.12, n.6, pp. 13-20 1225-4428
10. S. Haykin, Adaptive Filter Theory, Prentice-Hall Inc.,1986.
11. Ronald E.crochiere, Lawrence R.Rabiner, Multirate Digital Signal Processing, Prentice-Hall Inc.

▲ 고 충 기 (Chung Ki ko)

1973년 6월 20일생



1999년 2월 : 호서대학교 공과대학
전자공학과 학사

1999년 3월~현재 : 숭실대학교 대학원
보통신전자공학부 석사
과정

※ 주관심분야: 디지털시스템 인지,
적용신호처리

▲ 이 원 철 (Won Chul Lee)

한국 음향학회지 제18권 제6호 참조

▲ 이 충 용 (Chung Yong Lee)

한국 음향학회지 제18권 제7호 참조