

협대역 동시통화 검출기를 이용한 음향반향 제거 시스템

백수진*, 박규식*, 김기만**

*단국대학교 정보·컴퓨터학부, 컴퓨터과학 전공

** 한국해양대학교 전파공학과

Acoustic Echo Cancellation System using Narrow-Band Double Talk Detector

Sujin Paek*, Kyusik Park*, Kiman Kim**

*Division of Information and Computer Science, Dankook Univ.

**Dept. of Radio Science & Eng., Korea Maritime University

E-mail : sion22kk@dankook.ac.kr, kspark@dankook.ac.kr

E-mail : kimkim@hanara.kmaritime.ac.kr

요약

본 논문은 원거리 회의 시스템이나 차량 내 핸즈프리 통화 시 필연적으로 발생하는 음향 반향을 제거하기 위해 새로운 알고리즘을 제안한다. 본 논문에서 제안된 음향반향 제거 시스템은 Delayless 서브밴드 음향반향 제거기와 협대역 동시통화 검출기로 구성된다. Delayless Subband 적응 음향 반향 제거기는 적은 계산 량과 높은 수렴속도로 음향 반향 제거 성능이 뛰어난 것으로 알려져 있으며 본 논문은 이를 이용해 안정적인 음향 반향 제거를 위해서 협대역 Subband 내에서 동시통화 검출기를 구현한다. 기존의 광대역 동시 통화 검출기에 비해 본 논문에서 제안된 협대역 동시통화 검출기는 저주파 Subband 대역에서만 동시통화 검출을 수행하여 down-sampling으로 인한 계산 량 감소와 저주파 특성을 가지는 Subband 대역의 신호 특성으로 인한 신뢰성 있는 통화 상태 정보를 제공함으로써 전체적인 음향반향 제어 시스템의 성능을 향상시킬 수 있도록 하였다. 본 연구에서 제안된 음향반향 제어 시스템의 성능은 다양한 컴퓨터 시뮬레이션을 통하여 입증하도록 한다.

I. 서론

그림 1은 밀폐된 공간 안에서 근단화자와 원단화자 사이의 통신 환경을 나타내고 있다. 마이크로폰을 통해 전송된 원단화자의 음성 신호($x(n)$)는 전송로를 통해 근단화자 측 스피커로 출력되고 이 신호가 반향 되어 다시 원단 화자에게 전달됨으로서 원단화자는 자신이 전송했던 음성(Echo)을 다시 듣게 된다. 이러한 현상은 통신상의 불편함을 초래하게 되는 데, 이를 방지하기 위해 음향 반향 제거 시스템이 이용된다. 음향 반향 제거 시스템은 반향 된 신호($d(n)$)를 적응 적으로 추정하여 이 신호와 유사한 형태의 반향 신호($y(n)$)를 생성해 내는데,

본 연구는 산업자원부의 산업기초기술연구개발사업의 연구비 지원으로 수행되었습니다.

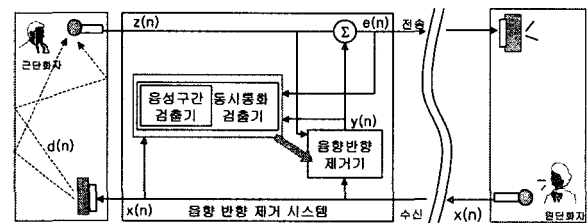


그림1. 음향반향 제어 시스템

($x(n)$):원단 화자 신호, $e(n)$:잔여 반향 신호, $z(n)$:반향 신호, $d(n)$:원단화자 반향신호, $y(n)$:추정 반향 신호)

이를 원 반향신호에서 빼줌으로서 반향 된 신호를 제거 하게 된다. 하지만 음향 반향 제거는 현재 근단화자의 통화 상태에 따라 원하지 않는 결과를 얻을 수도 있어, 동시 통화 검출기(Double Talk Detector:DTD)를 이용하여 근단화자의 통화 상태에 대한 정보와 함께 음향반향 제거기의 동작을 제어한다.

그림1에서 동시 통화 검출기안의 음성 구간 검출기는 본 연구의 협대역 Subband DTD에서 이용되는 것으로 원단화자의 신호 존재여부를 판단하는 기능을 가지며, 동시 통화 검출기내에서 보다 신뢰성 있는 통화 상태 정보를 얻기 위한 보조적인 역할을 한다.

II. Delayless 서브밴드 음향 반향 제거 알고리즘과 동시통화 검출

본 논문에서 사용되는 Delayless 서브밴드 음향 반향 제거 알고리즘은 앞서 설명한 서브밴드 적응 디지털 필터의 적은 계산 량과 높은 수렴속도에 대한 잇점은 그대로 유지하면서 필연적으로 발생하게 되는 앨리어싱(Aliasing)과 시간지연의 보완이 가능하다[1][2][3]. 그림 2는 Delayless subband 알고리즘의 구조를 보인다.

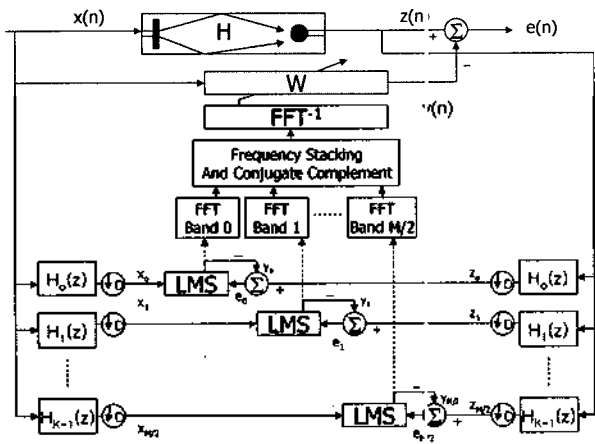


그림 2. Delayless 서브밴드 음향반향 제거 알고리즘

그림 2에서 $x_{0 \sim M/2}$, $z_{0 \sim M/2}$, $y_{0 \sim 2/M}$ 은 각각의 서브밴드에서의 원단 화자 신호와 반향 신호, 추정 반향 신호를 나타낸다. Delayless 서브밴드 알고리즘은 각 서브밴드별로 협대역 적응 filter의 계수를 갱신한 후, 광대역(wideband) 필터 계수로 재구성하여 적응 필터링을 구현하는 방법이다. 먼저 Delayless 서브밴드 음향 반향 제거기는 원단화자 신호와 반향신호를 K개의 서브밴드 필터 뱅크 $\{H_0(z), H_1(z), \dots, H_{K-1}(z)\}$ 를 통과시켜 각 밴드별로 NLMS (Normalized Least Mean Square) [4] 알고리즘을 이용하여 각 밴드별 적응 계수를 갱신한다. 각 밴드내의 NLMS계수들은 특정 대역내의 시간 영역에서의 값들로서 FFT를 수행하여 주파수 영역으로 변환한 후, 동작 대역에 해당하는 위치에 저배치(Frequency Stacking)를 하게되고 FFT를 수행하여 광역 필터 계수로 변환을 하게 된다. 각 밴드별 적응 계수의 갱신은 수식 (1)과 같다.

$$W_m(n+D) = W_m(n) + \mu X_m^* e_m(n) \quad (1)$$

수식(1)에서 $e_m(n)$ 은 추정 반향 신호와 반향신호간의 각 m번째 서브밴드에서의 신호 차이를 나타내고, $W_m(n)$ 과 $X_m(n)$ 은 각 m번째 서브밴드의 필터 계수와 대역 분할된 원단화자 신호를 나타낸다. 수렴 계수 μ 는 수식(2)와 같이 각 대역별 서브밴드 신호의 에너지에 의해 정규화 시켜 사용한다.

$$\mu = \mu(n) = \frac{\alpha}{X_m^T(n)X_m(n)}, \quad 0 < \alpha < 1 \quad (2)$$

그림 2의 적응 음향 반향 제거기는 근단화자의 존재 여부에 따라 시스템 동작 여부가 결정되어야 한다. 수식 (3)과 같이, 원단화자 신호의 반향 신호에 근단화자 신호가 더해지는 경우, 즉, 동시통화의 경우, 반향신호와 추정된 반향 신호간 근단화자 신호크기 만큼 차이가 나게되어 음향반향 제거기는 $e(n)$ 을 최소 값으로 만들기 위해 적응 적으로 반향 신호를 추정하면서 근단화자의 신호마저 제거되어 원하지 않는 결과를 얻게 된다.

$$\begin{aligned} e(n) &= z(n) - y(n), \\ z(n) &= d(n) + v(n), \quad y(n) \approx d(n) \end{aligned} \quad (3)$$

수식 (3)에서 $d(n)$ 은 반향신호, $v(n)$ 은 근단화자 신호를 나타낸다. 또한 근단화자 신호만 있게 되는 경우는

수식 (3)에서 $d(n)$ 이 존재하지 않아 동시통화일 경우와 유사한 결과를 초래하게 된다. 이와 같이 음향 반향 제거기는 근단화자와 원단화자의 통화 상태에 따라 그 동작 여부가 민감하게 결정되어야 한다. 이러한 이유로 음향 반향 제거 시스템에서는 근단화자의 통화 상태 정보를 제공하는 동시통화 검출기(DTD)를 사용하는데, 동시통화 검출기는 근단화자의 통화 상태 정보를 음향 반향 제거기에 제공함으로써, 동시통화의 경우 계수적응은 막으면서, 원단화자의 반향신호를 제거하기 위한 필터링을 하게되며, 스피커의 볼륨을 올리는 경우, 몸을 움직였을 경우 등, 반향경로가 변경되었을 때는 동시통화 상태가 아니기 때문에 계속적인 계수적응이 이루어지도록 한다.

III. 협대역 동시통화 알고리즘

음향 반향 제거기의 효율적인 동작을 위해 통화상태 정보를 제공하는 동시 통화 검출 알고리즘에는 크기 비교법, 상호 상관법 [8] 등이 존재한다. 크기 비교법은 반사 경로에 의한 감쇄나 적응필터에 의한 감쇄, 또는 이 두 가지 감쇄 모두를 이용하여 동시통화의 발생 여부를 판단하는 알고리즘으로 근단 화자 신호가 작은 경우, 동시통화 임에도 불구하고 동시통화가 아닌 것으로 판단될 수 있다는 단점을 가진다. 반면 상호 상관 법은 원단 화자 신호와 잔여반향신호간의 직교성질을 이용한 알고리즘으로, 입력되는 신호 크기에 큰 제약을 받지 않는다. 이 상호 상관 법을 정규화 시켜 통화 상태의 구분을 명확하게 해주기 위한 수식(4)의 정규화 된 신호 상관 법]이 제시되기도 하였다.

$$C_{xe}(n) = \frac{E[x(n)e(n)]}{\alpha \sigma(x(n)) E[x(n)^2]}, \quad 0 < \alpha < 1 \quad (4)$$

상호 상관법을 이용한 동시통화 알고리즘은 원단화자 신호와 잔여 반향신호간의 직교성질을 이용하는 것으로 $C_{xe}(n)$ 값에 대해 적절한 경계치 값을 설정하여 동시 통화 검출을 하게된다. 또한 근단화자의 신호와 원단화자의 신호간 서로 상관도가 낮다는 것을 이용하여 반향경로로 변경 시, 잔여 반향 신호와 원단화자 신호간의 상관도가 동시통화의 경우보다 상대적으로 높아 반향경로가 변경되었을 경우와 동시통화의 경우를 보다 효과적으로 구분할 수 있다는 점에서 좋은 성능을 보인다. 수식(4)에서 원단화자 신호를 추정 반향신호로 이용해도 무관하며, 본 연구에서는 수식(4)의 신호 상관 법을 이용해 통화 상태 정보를 결정한다.

본 논문에서 제안하는 협대역 동시통화 검출기는 Delayless 서브 밴드 음향반향 제거 알고리즘에서 원단 화자 신호와 반향신호를 일정 간격의 주파수대로 나누어 적응 필터를 이용한다는 점과, 음성신호의 정보는 낮은 주파수 밴드에 집중되어 있다는 점을 착안하여 저주파 대역에서만 동시 통화 검출을 한다. Delayless 서브밴드 음향반향 제거기와 정규화된 상호 상관 법을 기반으로 하여 협대역 동시통화 검출기를 적용시킨 음성반향 제거 시스템의 구조는 그림3과 같다.

그림 3에서 e_0 는 최저 주파수대의 잔여반향신호, y_0 는 최저 주파수대의 추정반향신호를 나타낸다. 우선 협대역 동시통화 검출기는 Delayless 서브밴드 음향 반

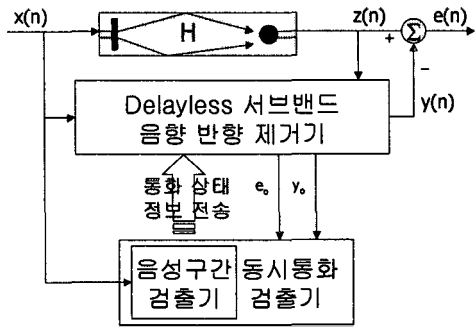


그림3. 제안된 음향 반향 제거 시스템

향 제거기에서 최저 주파수대의 잔여 반향 신호와 추정 반향신호를 입력받아 둘의 상관도를 계산하여 근단화자 신호의 존재여부를 결정하고, 음성구간 검출기를 통해 원단화자의 신호 존재여부를 판단해, 최종 통화 상태 정보를 결정한다. 이렇게 얻어진 정보는 Delayless 서브밴드 음향반향 제거기에 전달이 되어 현재 입력 음성 sample에 대해 필터링만을 할 것인지, 계수적응을 할 것인지 결정한다. 협대역 동시통화 검출기가 상관도를 계산하기 위해 이용하는 두 입력, 최저주파수대의 잔여 반향신호와 추정반향신호는 음향 반향 제거 알고리즘에서 협대역 연산을 하기 위해 down-sampling된 상태의 신호이다. 즉, sample수가 down-sampling비율만큼 줄어들게 된다. 그러므로 협대역 동시통화 알고리즘을 이용하는 경우, 전 대역(wideband)에서 동시통화 검출을 하는 경우보다 down-sampling된 비율만큼 계산량의 감소가 이루어 질 수 있게 된다. 실제 컴퓨터 모의 실험에서는 40000sample에 대해 32개 대역 분할과 2배 over-sampling을 하는 Delayless 서브밴드 음향 반향 제거 알고리즘을 이용하게 되는데, 이로서 협대역 DTD에서는 16배의 계산량 감소를 가져 올 수 있다. 또한 두 신호 $y_0(n), e_0(n)$ 는 최저 주파수대의 신호로서 부드러운 신호의 특성으로 인해 수식(4)의 $C_{xx}(n)$ 값의 변동치가 작아져서 경계치 결정에 안정성을 제공할 수 있다.

V. 컴퓨터 모의 실험

본 장에서는 제안된 협대역 동시 통화 알고리즘의 성능을 알아보기 위해 통화상태에 따라 음향반향제거 시스템이 협대역 DTD를 이용했을 경우와 광대역 DTD를 사용했을 경우를 비교하여 모의 실험을 수행하였다. 본 실험을 위해서 Delayless 서브밴드 음향 반향 제거 알고리즘에서 32개 서브밴드, 512개 tap을 가지는 광역 계수 필터를 이용하고, 각 서브밴드별로 32tap을 가지는 NLMS 필터를 적용시켜 결과를 유도한다. DTD로는 Normalized Correlation을 이용한다.

모의 실험에서 사용된 원단화자신호와 반향신호는 표 1과 같이 4가지(IDLE:신호존재하지 않는 경우, NETst:근단화자 신호만 존재하는 경우, 동시통화, FETst:원단화자 신호만 존재하는 경우)의 통화 상태가 고려되도록 인공적으로 생성하였다. 협대역의 경우, FETst 상태 결정을 위해 신호 존재 여부를 결정하는 VAD(Voice Activity

Detector)를 보조적으로 이용하여 FETst를 구별한다. 그리고, 각 통화 상태에 따른 음향 반향 제어기의 동작에 대해 필터링과 계수적응을 구분하여 제어되도록 하였다.

입력 sam	$(len/8)*0+1$ $\sim (len/8)*1$	$(len/8)*1+1$ $\sim (len/8)*2$	$(len/8)*2+1$ $\sim (len/8)*3$
통화 상태	IDLE	NETst	동시통화
음향반향 제거 동작	X	X	△
$(len/8)*3+1$ $\sim (len/8)*4$	$(len/8)*4+1$ $\sim (len/8)*5$	$(len/8)*5+1$ $\sim (len/8)*6$	$(len/8)*6+1$ $\sim (len/8)*7$
동시통화	FETst	IDLE	FETst
△	o	X	o
$(len/8)*7+1$ $\sim (len/8)*8$		필터링	계수적응
FETst(*)	△	on	off
o	X	off	off

표1. 입력신호정보에 따른 음향반향제거기 동작상황

표1에서 FETst(*)는 바로 이전 FETst의 상태에서 반향 경로가 변경 되었을 때의 FETst를 나타내며, len은 입력되는 신호의 총 길이를 나타낸다. 표1과 같은 입력에 대해 음향 반향 제거기의 수렴성능은 수식 (5)의 ERLE (Echo Return Loss Enhancement)를 이용하여 평가하게 된다.

$$ERLE(n) = 10 \log_{10} E \left[\frac{|z^2(n)|}{e^2(n)} \right] [dB] \quad (5)$$

ERLE는 반향 신호 대 잔여 반향 신호의 비율 나타내는 parameter로서 음향 반향 제거기의 성능 평가에 주로 사용되며 잔여 반향 신호 $e(n)$ 에 민감하게 반응한다. 반향 제거기의 필터 계수가 반향 경로에 수렴할수록 ERLE는 일정한 값을 유지하게 되지만 반향 경로가 변하는 경우에는 ERLE가 떨어지게 된다. 실제로 동시 통화의 경우 이전 ERLE값을 일정하게 유지되어야 하며, 원단화자 신호만 존재하게 될 경우나 반향 경로가 변경될 경우는 빠른 속도로 수렴되어야 할 것이다. 이외의 경우는 음향 반향 제거기가 동작될 필요성이 없게 된다.

그림4는 전체 신호길이(len)를 40000sample로 설정하고 입력 신호를 표1의 상태대로 적용 시켰을 경우, 음향 반향 제거 시스템의 ERLE에 대한 모의 실험 결과이다.

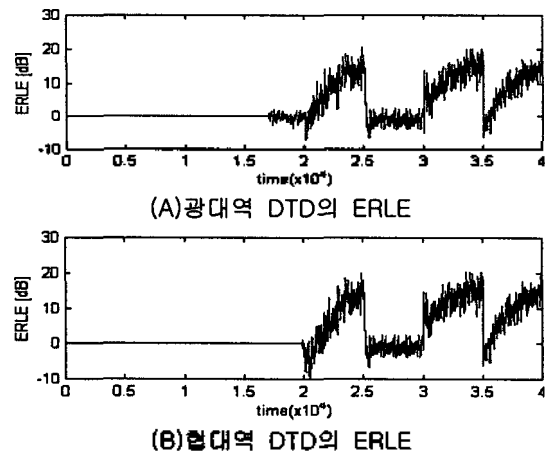
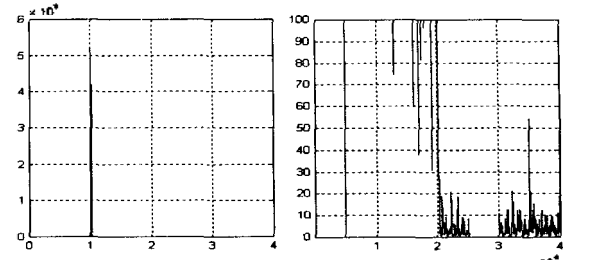


그림4. len=4000 sample에 대한 모의 실험 비교

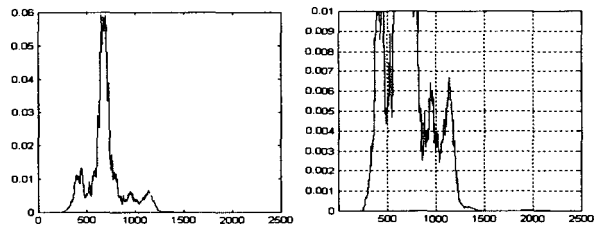
그림 4에서 1~10000sample까지 IDLE한 상태의 신호와 NETst신호가 들어오는데, 이 경우 모뎀 음향 반향 제거기가 동작하지 않아야 하는 부분이 된다. 이로 인해 위에서 제시했던 대로 입력 신호는 필터링만을 거치게 되어, 현재 계수가 0으로 초기화된 상태에서, ERLE의 초기값도 0의 값을 가지게 되어, 필터링을 거치면서 ERLE의 값은 0의 값을 가지게 된다. 다음 20000sample까지는 동시 통화 부분이므로 현재 저장된 계수로 필터링을 하게 되어 이전의 경우와 같이 ERLE는 0의 값을 가지게 된다. 하지만 그림 4(A)의 광대역 DTD의 경우, 17000 sample과 20000 sample사이 협대역 DTD에서와는 다르게 0이 아닌 값의 계수로 필터링이 이루어지고 있음을 발견할 수 있다. 이는 앞서 설명했던 경계치 결정에 있어서의 문제로 동시통화 구간이지만, 광대역 DTD의 경우 $C_{xx}(n)$ 의 큰 변동률 때문에 갑작스런 peak값으로 통화 구간이 FETst으로 잘못 판단되어 순간 계수 적응이 일어나 0으로 초기화되었던 계수 값이 0이 아닌 값으로 변경되었기 때문이다. 다음 25000sample까지는 FETst구간으로 계수 적응이 일어나 반향(Echo)이 적응적으로 제거되어야 하는 구간으로 그림 4에서와 같이 ERLE값이 수렴되어야 하는 부분이다. 그리고 다음 30000sample까지는 IDLE 구간으로 현재 계수 값으로 필터링만 하게 된다. 또한, 35000sample까지와 40000sample까지는 FETst구간이지만 반향환경이 변화된 것뿐이므로 새로운 반향 환경에 새롭게 계수 적응이 일어나야 하므로 그림 4에서와 같이 ERLE가 새롭게 수렴과정을 거치는 것이 옳다.

그림 5는 각 DTD에서의 시간별 $C_{xx}(n)$ 값이며, 오른쪽의 그림은 왼쪽의 그림을 더욱 자세히 작은 scale 값으로 $C_{xx}(n)$ 값의 범위를 구체화시킨 것이다. 그림 5에서 가로축은 sample수를 나타내며, 세로축은 수식 (4)에서 보였던 $C_{xx}(n)$ 의 값을 나타낸다. 모의 실험에 이용된 경계치 값은 광대역에서 동시 통화를 구분하기 위한 경계치 값 50, FETst 경계치 값 0.3, NETst를 경계치 값 3000을 이용하며, 협대역에서 동시통화를 구분하기 위한 경계치 값 0.0025, 원단 화자 신호의 존재 유무를 구분하기 위한 경계치 값 50(VAT에서 이용하는 경계치 값)을 이용한다. 그림 5에서 보듯이 광대역일 경우는 동시통화의 경우와 FETst, 그 이외의 경우를 분명히 구분할 수 있지만, sample간 값의 변화도가 크며 많은 peak 값을 가져 순간적으로 잘못된 통화 상태 정보를 제공하게 될 우려가 있다. 그림 4(A)는 이러한 peak 때문에 17000sample에서 순간적으로 동시 통화 제거기가 FETst 구간으로 판단하여 계수적응을 한 것이다. 반대로 협대역일 경우는 근단화자 신호의 존재 유우만을 판별 가능하지만, 이는 VAT를 부가적으로 이용함으로써 보완이 가능하며, 광대역에서의 peak값이 거의 존재하지 않는 곡선 형태를 형성하여, 이로 인해 보다 안정적인 경계치 값 결정으로 신뢰성 있는 통화 상태 정보를 제공할 수 있다는 것을 알 수 있다.

본 모의 실험과정에 있어서 실제 인산되는 시간을 관측해 본 결과 광대역 DTD의 경우와 협대역 DTD의 경우 각각 165.0370, 2.2930 (MATLAB:etime()이용)으로 협대역 DTD의 경우가 훨씬 적은 연산시간을 가졌다.



(A)광대역의 경우(왼쪽: 큰 scale,오른쪽: 작은 scale)



(B)협대역의 경우(왼쪽: 큰 scale,오른쪽: 작은 scale)

그림 5. $C_{xx}(n)$ 값의 변화

V. 결론

원거리 회의 시스템이나 차량 내 핸즈프리 통화 시 필연적으로 발생하는 음향 반향은 음향반향 제거기를 적응적으로 사용하여 제어할 수 있다. 이러한 음향반향 제거기를 정상적으로 동작시키기 위해서는 근단화자의 통화 상태를 판단하기 위한 효율적인 동시통화 검출기가 필요하다. 본 논문에서 제안된 음향반향 제거시스템은 Delayless 서브밴드 음향반향 제거기와 정규화된 신호 상관 법을 이용한 협대역 동시통화 검출기로 구성된다. 본 연구에서는 기존의 광대역 동시 통화 검출기 대신에 협대역 동시통화 검출기를 사용하여 음성신호의 저주파 Subband 대역에서만 동시통화 검출을 수행함으로써 down-sampling으로 인한 계산 량 감소와 저주파 subband 대역의 신호 특성으로 인해 보다 신뢰성 있는 동시 통화 검출 성능을 보장함으로써 전체적인 음향반향 제어 시스템의 성능을 향상시킬 수 있었다.

참고 문헌

- [1] R. E. Crochiere and L. R. Rabiner, Multirate Digital Signal Processing, Englewood Cliffs, N. J.: Prentice-Hall, 1985
- [2] E. R. Ferrara, Jr., "Frequency-domain adaptive filtering", in Adaptive Filters, C. F. N. Cowen and P. M. Grant, Eds. Englewood Cliffs, N.J.: Prentice Hall, 1985
- [3] D. Morgan and J. Thi, "A Delayless Subband Adaptive Filter Architecture", IEEE Trans on Signal Processing, vol. 43, no. 8, August 1995
- [4] J. R. Treichler, C. R. Johnson, Jr and M. G. Larimore, Theory and Design of Adaptive Filters, John Wiley & Sons, New York, 1987