

# 반사계수를 이용한 동시통화 검출기

## (A Double Talk Detector Using The Reflection Coefficients)

柳 在 河 \* , 趙 誠 鎬 \*\* , 尹 大 熙 \*\*\*

(Jae Ha Yoo, Sung Ho Cho, and Dae Hee Youn)

### 요 약

본 논문에서는 음향반향제거시스템의 성능을 향상시킬 수 있는 지능적인 동시통화검출기를 제안하였다. 기존의 방법들은 반향경로의 변화와 동시통화를 구분하지 못하거나 구분하더라도 판단 지연시간이 발생한다. 또한 적응필터가 수렴한 뒤부터 동시통화검출기가 정상적으로 동작할 수 있다. 제안된 알고리즘은 원단과 근단에 격자예측기를 설치하고 이로부터 얻어진 반사계수의 변화도를 이용하여 동시통화를 검출한다. 실제 음성신호를 사용한 다양한 모의실험 결과에 의해 제안된 알고리즘의 우수성이 입증되었다.

### Abstract

In this paper, we propose an intelligent double talk detector that can enhance the performance of the acoustic echo cancellation system. The conventional double talk detection methods often misunderstand the echo path changes as double talk. Although there exist several detection methods that can distinguish the echo path changes from the double-talks, they show poor tracking performance because of the excessive decision delay for the discrimination and can only be used after the adaptive digital filter converges. A new and more effective detection algorithm has been proposed, where the detection mechanism is performed by observing the change rate of the reflection coefficients of the two lattice predictors that are placed on the near-end and far-end terminals. The excellence of the proposed method is verified by extensive computer simulations using real speech signal.

### I. 서 론

통화자의 편리함과 안정성을 증가시키기 위하여 원격 회회의시스템, 스피커폰(speaker-phone), 탁상용 원격

회의 시스템과 같은 핸드프리(hand-free) 통화 방식이 점차 보편화되고 있다<sup>[1]</sup>. 그러나 핸드프리 통화 방식에서는 음향반향(acoustic echo)이 발생하여 통화 품질이 저하되므로 반향을 효과적으로 제어할 수 있는 장치가 필요하다. 음향반향제거기는 디지털필터를 사용하여 반향 발생경로의 임펄스응답을 추정하여 반향신호만을 빼주는 방식으로 가장 효과적인 방법이다.

\* 正會員, LG電子 멀티미디어 研究所

(Multimedia Research Laboratory, LG Electronics Inc.)

\*\* 正會員, 漢陽大學校 電子工學科

(Dept. of Electronic Eng. Hanyang Univ.)

\*\*\* 正會員, 延世大學校 電子工學科

(Dept. of Electronic Eng. Yonsei Univ.)

接受日字: 1997年5月23日, 수정완료일: 1997年8月4日

반향제거기의 입력 신호인 음성 신호와 반향경로는 시간에 따라 그 통계적 특성이 변화하기 때문에 반향제거기는 적응필터로 구현된다. 적응필터는 기준 입력 신호와 오차 신호를 사용하여 반향경로의 임펄스 응답을 추정한다. 근단화자 신호가 없는 구간에서는 오차 신호에 제거되지 못한 잔여 반향 성분만이 존재한다.

그러나 근단화자 신호가 존재하는 구간에서는 오차 신호에 잔여 반향 뿐만 아니라 근단화자 신호가 존재한다. 근단화자 신호는 적응필터의 계수 적응에 방해 요소로 작용하며 그 크기도 반향 신호에 비해 크기때문에 근단화자 신호가 존재하는 구간에서 필터 계수의 적응이 계속된다면 적응필터가 발산하기때문에 동시통화 구간에서는 필터 계수의 적응이 중단되어야 한다.

동시통화는 근단화자가 먼저 발성을 시작하고 원단화자가 나중에 발성을 하는 경우와 원단화자가 먼저 발성을 시작하고 근단화자가 나중에 발성을 하는 경우로 나눌 수 있다. 전자의 경우 원단과 근단에 설치된 음성구간검출기(VAD:Voice Activity Detector)로부터 쉽게 판단될 수 있으나 후자의 경우는 음성구간검출기로 판단되기 어렵다. 왜냐하면 원단화자 신호가 먼저 시작되는 경우 근단에는 근단화자 신호의 존재 유무에 관계 없이 반향 신호가 존재하기 때문에 근단의 음성구간검출기는 신호가 있는 구간으로 판단하게 된다. 그러므로 그림 1과 같이 음성구간검출기 이외에 동시통화검출기(DTD:Double Talk Detector)를 설치하여야 한다. 여기서  $x(n)$ 은 원단신호이며 적응필터의 기준입력신호가 된다.  $y(n)$ 은 마이크로폰에 유입되는 반향 신호이며  $v(n)$ 과  $\xi(n)$ 은 각각 근단화자의 음성 신호와 근단의 주변 잡음을 나타낸다.  $z(n)$ 은 반향 신호, 근단화자 신호 및 주변 잡음이 마이크로폰에 입력된 근단신호이며 적응필터의 주입력신호가 된다.  $\hat{y}(n)$ 은 음향반향제거기에 의해 계산된 반향추정신호이며,  $e(n)$ 은 근단신호로부터 반향추정신호가 제거된 오차 신호로 상대방에게 전송될 신호이다.

가장 단순한 동시통화 검출기는 근단신호  $z(n)$ 이나 오차신호  $e(n)$ 의 크기를 추정하여 크기가 갑자기 커지기 시작하는 순간을 동시통화가 시작되는 구간으로 판정하는 방법이다. 그러나 근단 신호와 오차 신호의 증가는 반향경로가 변화한 경우에도 일어날 수 있다. 이 경우 적응필터는 계수 적응을 중단해서는 안되며 오히려 변화된 반향경로의 임펄스응답을 빨리 추정할 수 있도록 계수 적응을 진행하여야 한다. 따라서 이러한 오동작을 피할 수 있도록 동시통화검출기는 동시통화와 반향경로의 변화를 구분할 수 있어야 한다. 그러므로 동시통화검출기의 성능은 근단화자 신호를 빠르고 정확하게 감지하는 것과 반향경로의 변화와 동시통화 상황을 빠르고 정확하게 구별하는 것에 좌우된다.

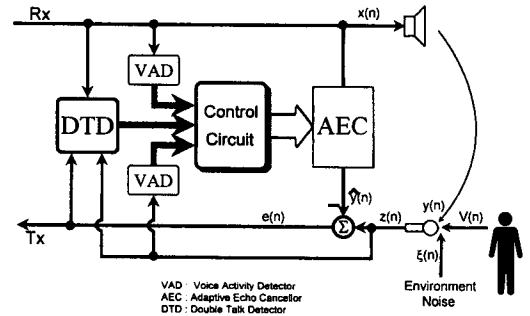


그림 1. 동시통화검출기  
Fig. 1. The double talk detector.

## II. 기존의 동시통화검출기

### 1. 크기 비교법

원단신호  $x(n)$ 의 현재와 과거값들로 이루어지는 기준 입력 신호 벡터의 원소중 가장 큰 절대값을 갖는 신호와 주입력 신호의 절대값을 비교하는 방법을 사용하여 동시통화를 검출할 수 있다<sup>[2]</sup>.

$$|z(n)| \geq \alpha \text{Max} \{ |x(n)|, |x(n-1)|, \dots, |x(n-L+1)| \} \quad (1)$$

여기서  $L$ 은 적응필터의 탭수를 나타내며  $\alpha$ 는 반향 귀환 감쇄(ERL : Echo Return Loss)를 고려하여 결정되는 상수이다. 음성 반향제거기에서는 하이브리드를 통해 최소 6dB의 감쇄가 일어나므로  $\alpha$ 는 1/2이 사용된다.

근단화자 신호가 시작되면 근단 신호  $z(n)$ 은 크기가 증가하게 되어 (1)의 부등식이 만족된다. 그러나 근단화자 신호의 레벨이 매우 작은 경우에는 (1)의 검출 알고리즘이 제대로 성능을 발휘하지 못하며 음향 반향제거기에 적용하기 위해서는 ERL을 미리 알고 있어야 한다. 그러나 반향경로가 많이 변화한 경우 특히, 사용자가 스피커의 볼륨을 올리거나 내릴 경우 ERL이 급격하게 변화하므로 오동작의 가능성이 매우 크다.

작은 근단화자 신호에 대해서도 정확히 동시통화를 검출하기 위해 주입력 신호를 사용하는 대신에 추정 오차 신호를 사용하는 방법이 있다<sup>[3]</sup>. 이 알고리즘에서는 다음과 같은 조건에 의해 동시통화가 검출된다.

$$Att(n) \leq Att_{opt} \quad (2)$$

이 때,  $Att(n)$ 은 반향 신호가 반향경로와 반향제거기에 의해 겪게되는 감쇄로서 다음과 같이 정의된다.

$$Att(n) = 10 \log_{10} \frac{E\{x^2(n)\}}{E\{e^2(n)\}} \quad (3)$$

$Att_{opt}$ 은 근단화자 신호가 존재하지 않는 구간에서 반향경로와 적응필터를 통하여 일어나는 반향 감쇄의 기대치로서 배경 잡음을 고려하여 신중하게 결정되어야 한다. 이 방법은 주입력 신호 대신에 오차 신호를 사용하기 때문에 근단화자 신호의 레벨이 작더라도 동시통화를 빨리 검출할 수 있다. 그러나 배경 잡음의 레벨과 반향경로의 성질은 시간에 따라 변화하기 때문에 적합한  $Att_{opt}$ 을 결정하는 일은 매우 어렵다.

반향 귀환 감쇄의 영향을 배제하고 반향 제거기에 의해 겪게되는 감쇄율만을 사용하여 동시통화를 검출할 수 있다<sup>[4]</sup>. 이 방법은 다음의 부등식이 만족되면 동시통화로 판단한다.

$$ERLE_{DTD}(n) \leq ERLE_{opt} \quad (4)$$

여기서  $ERLE_{DTD}(n)$ 은 (5)로 정의되며  $ERLE_{opt}$ 는 적응필터를 통하여 일어나는 반향 감쇄의 기대치로서 배경 잡음을 고려하여 신중하게 결정되어야 한다.

$$ERLE_{DTD}(n) = 10 \log_{10} \frac{E\{z^2(n)\}}{E\{e^2(n)\}} \quad (5)$$

이 방법은 근단 신호와 오차 신호를 사용하므로 근단화자 신호에 매우 민감하게 반응한다. 그러나 배경 잡음에 따라  $ERLE_{opt}$ 를 다시 선정하여야 하며 배경 잡음이 큰 경우에는 근단화자 신호에 대해 민감하게 반응하지 못한다. 반응 속도를 증가시키기 위해 서브밴드 구조를 사용하여 각 밴드별로  $ERLE_{DTD}$ 를 계산하여 하나의 밴드에서라도 동시통화가 검출되면 동시통화로 판정하는 방법이 있다<sup>[5]</sup>.

위의 3가지 방법은 동시통화를 검출할 수 있지만 반향경로의 변화도 동시통화로 판단하게 된다. 그 결과 실제로는 적응필터의 계수 적응이 더욱 필요함에도 불구하고 오히려 적응을 중단하게 된다. 그러므로 반향경로가 변화한 경우와 동시통화를 잘 구분해 낼 수 있는 지능적인 방법이 필요하다.

## 2. 상호 상관법

$Y_e$ 는 직교 성질을 이용한 동시통화 검출기를 제안

하였다<sup>[6]</sup>. 즉, 시간  $n$ 에서 적응필터가 최적치에 수렴한 경우 기준 입력 신호와 추정 오차 신호간에는 다음과 같은 직교 성질이 만족된다.

$$C_{Xe}(n) = E [ X(n) e(n) ] = 0_L \quad (6)$$

여기서  $X(n)$ 은  $L$ 개의 원소를 갖는 기준 입력 신호 벡터이고  $0_L$ 은  $L \times 1$  영벡터(zero vector)이다.

동시통화가 발생하는 경우 적응필터의 주입력 신호는 급격히 레벨이 커지게 된다. 그러나 근단화자 신호  $v(n)$ 은 기준 입력 신호와 상관성이 없는 신호이므로  $C_{Xe}$ 는 직교 성질이 계속 만족된다. 반면, 반향경로의 특성이 변화한 경우에는  $C_{Xe}$ 의 직교 성질이 더 이상 만족되지 못한다. 즉, 기준 입력 신호와 추정 오차 신호간의 상관도는 동시통화시에는 변화가 없고 반향경로가 변화한 경우에는 매우 민감하게 증가한다는 사실을 이용하는 것이다. 그러므로 적당한 문턱치(threshold)를 선택하므로써 동시통화와 반향경로의 변화를 효과적으로 구분할 수 있다. 그러므로 상관도의 변화는 다음과 같이 분류된다. 처음 수렴하는 동안과 반향경로가 변화한 경우에는 큰 값을 갖고 수렴후 정상 상태에서는 근단화자 신호의 존재 유무에 관계 없이 작은 값을 갖는다.

한편 두개의 문턱치를 사용하므로써 [6]에서 발생할 수 있는 잘못된 동시통화 검출을 피할 수 있고, 기준 입력 신호와 추정 오차 신호간의 상호 상관도를 추정하는 대신 (7)과 같이 필터 출력 신호와 추정 오차 신호 간의 상호 상관도를 추정하여 사용하므로써 계산량 및 필요한 메모리 소자수를 줄일 수 있다<sup>[7]</sup>.

$$C_{\hat{y}(n)e(n)} = E [ \hat{y}(n) e(n) ] = 0 \quad (7)$$

위의 방법은 동시통화와 반향경로의 변화를 구분해 낼 수 있지만 반향제거기의 운용 방식에 문제점이 있다. 즉, 필터 계수가 수렴한 후에는 필터 계수를 적응시키지 않고 반향경로가 변화한 경우에만 적응을 수행한다는 것이다. 그러나 음향 반향경로는 사용자의 머리카락 등 신체부위의 움직임이나 주변 환경의 변화가 자주 일어나므로 수렴 후에도 계속적으로 필터 계수를 적응시켜야 한다.

## 3. 동시통화와 반향경로 변화간의 구분이 가능한 방법

음향 반향제거기에 적합한 방식으로 필터 계수를 적응시키면서도 동시통화와 반향경로의 변화를 구분할

수 있는 방법이 있다.

대화의 통계적 특성을 사용하여 동시통화가 발생한 경우와 반향경로가 변화한 경우를 구분할 수 있다<sup>[8]</sup>. 이 방법은 동시통화는 일정 시간 이상 지속되지 않는다는 것을 이용한 방법이다. 즉 오차 신호의 크기를 관찰하여 문턱치를 넘는 경우 일단 필터 계수의 적응을 중단하고 계속해서 오차 신호의 레벨을 검토하다가 일정 시간이 지나도 계속 문턱치를 넘는 경우 반향경로의 변화로 판단하여 적응을 재개한다. 즉, 일정시간 이상 계속해서 오차 신호가 문턱치를 넘는 것은 반향경로가 변화한 경우에만 발생한다고 판단하는 것이다. 그러나 이 방법은 반향경로가 변화한 경우 일정시간([8]에서는 3초)이 지난뒤에야 적응을 재개하므로 추적 성능이 매우 떨어진다.

음향 반향제거기는 서브밴드 적응필터로 구성하고 동시통화와 무관하게 계속해서 필터 계수를 적응시키는 1개의 서브밴드 적응필터  $H_{DT}$ 를 별도로 마련하여 동시통화를 검출하는 방법이 제안되었다<sup>[9]</sup>. 이 방법은 다음과 같은 성질을 이용하여 동시통화와 반향경로의 변화를 구분한다. 동시통화인 경우에는  $H_{DT}$ 가 계속해서 필터 계수를 적응시키므로 ERLE가 계속해서 문턱치 아래에 머무르는 반면 반향경로가 변화한 경우에는 일정 시간 뒤에는 ERLE값이 문턱치 위로 올라온다. 또한 2개의 문턱치를 사용한다. 즉, 상위 문턱치는 동시통화를 빠르게 검출하기 위해 사용하고 하위 문턱치는 빠른 추적 성능을 갖기 위해서 사용한다. 상위 문턱치는 30~35dB로 하위 문턱치는 20~25dB로 설정한다. 그러나 ERLE는 입력 신호의 크기에 따라서도 값이 변화하므로 실제 음성 신호가 쓰일 경우 동시통화나 반향경로의 변화가 발생하지 않더라도 ERLE 값은 자주 변화하여 잘못된 동시통화 검출을 유발시킬 수 있다. 그러므로 실제 시스템 구현시에는 문제가 발생한다.

### III. 제안된 동시통화검출기

현재까지 연구된 동시통화 검출 방법들은 동시통화가 발생한 경우와 반향경로가 변화한 경우를 구분하지 못하거나 구분이 가능하여도 두 상태를 구분하는데 필요한 지연시간이 많이 필요하다. 또한 적응필터의 수렴을 가정한 상태에서의 검출 방법이 대부분이므로 적응필터가 수렴 상태에 도달하기 전에 동시통화가 일어

나는 경우에는 오동작을 하게된다. 그러므로 다음과 같은 성질을 갖는 동시통화 검출 방법이 필요하다. 첫째, 정확하고 빠르게 동시통화를 검출하고 둘째, 반향경로의 변화와 동시통화를 적은 판단 지연 시간 내에 구분해 낼 수 있어야 하며 셋째, 적응필터의 수렴 여부에 무관한 검출 방법이어야 한다.

음성 신호의 시작 부분과 끝 부분을 검출하는 끝점 검출(endpoint detection) 과정은 음성 인식 시스템의 인식율의 성능을 좌우하는 중요한 과정이다. 이웃한 프레임(frame) 간의 스펙트럼의 변화도를 이용하면 끝점을 검출할 수 있다. 왜냐하면 음성 신호가 시작되는 부분과 끝나는 부분에서는 그 특성이 변화하므로 스펙트럼의 변화도 크게 나타나기 때문이다. Soong<sup>[10]</sup>은 스펙트럼의 변화도를 이용하여 음성 신호의 음소 단위 분리를 위한 분할(segmentation) 과정을 수행하였으며, [11]에서는 이웃 프레임 간의 LPC 케스트럼(cepstrum)의 변화도를 끝점 검출 과정에 적용하였다.

본 논문에서는 그림 2와 같이 근단 신호를 입력 신호로 하는  $P$ 차의 격자 예측기를 설치하고 샘플간의 반사계수의 변화도를 사용하여 동시통화를 검출하는 방법을 제안한다.

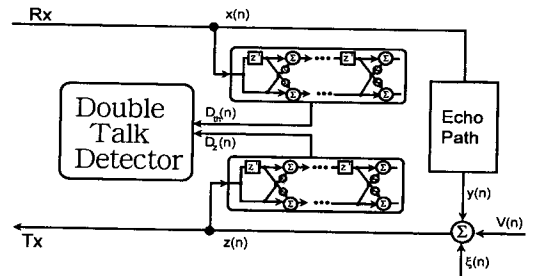


그림 2. 제안된 동시통화 검출기  
Fig. 2. The proposed double talk detector.

근단화자 신호가 존재하는 경우 적응필터의 주입력 신호인  $z(n)$ 에는 반향 신호와는 통계적 특성이 매우 다른 신호가 존재하게 된다. 또한 반향 신호에 비해 근단화자 신호는 큰 레벨을 갖기 때문에 근단화자 신호가 시작될 때는 통계적 특성이 크게 변화하여 반사계수의 변화도도 커지게 된다. 한편, 반향경로가 변화한 경우에는 그 특성의 변화가 미미하므로 작은 변화도를 갖게 된다. 그러므로 (8)과 같은 반사 계수의 변화도  $D(n)$ 을 추정하고 이를 적절하게 선택된 문턱치

와 비교하므로써 동시통화와 반향경로의 변화를 구분한다.

$$D(n) = \frac{\sum_{i=1}^{\Gamma} h_i (k_i(n) - k_i(n-T))^2}{\sum_{i=1}^{\Gamma} h_i (k_i(n))^2} \times 100 \quad (2 \leq \Gamma \leq P) \quad (8)$$

여기서  $\Gamma$ 와  $T$ 는 각각  $P$ 차로 구성된 격자 예측기에서 유효한 값을 갖는 반사계수가 있는 차수와  $T$ 샘플 이전의 반사계수를 나타내며  $h_i$ 는 가중치 함수이다.  $\Gamma$ 는 실험적으로 얻어지는 값이다. 음성 신호에 대한 반사계수의 통계적 분포 특성<sup>[12]</sup>을 고려하여 앞 단의 반사계수에 더욱 큰 비중을 준다.

동시통화가 발생할 경우  $D(n)$ 은 갑작스럽게 커지게 된다. 그러나 반향경로가 변화한 경우에도 그 값이 크게 변화한다면 제안된 동시통화 검출기는 반향경로의 변화도 동시통화로 간주하여 필터계수의 적응을 멈추게 되므로 성능 저하가 매우 심각해진다. 음향 반향 제어 시스템에서 반향경로가 변화하는 경우는 사용자가 좀 더 잘 듣기 위해 스피커의 볼륨을 올리는 경우, 사용자가 움직이거나 문의 개폐등으로 인한 경우이다. 스피커의 볼륨을 올리더라도  $z(n)$ 의 통계적 특성은 변화하지 않기 때문에  $D(n)$ 은 거의 변화하지 않는다. 또한 주변환경의 변화로 인한 경우도 신호의 통계적 특성이 반향경로에 의한 부분에 의해서만 일어나기 때문에 근단화자 신호가 시작되는 경우보다는 매우 작은 변화가 일어나게 된다.

본 논문에서는 4m×3m 크기의 실내 공간에서 PC를 테이블 위에 설치하고 PC용 스피커와 일반적인 콘덴서 마이크를 사용하여 측정된 8가지 경우의 임펄스응답을 얻고 이에 대해 실험한 결과  $D(n)$ 이 매우 작은 값을 나타냄을 관찰할 수 있었다. 특히, 하나의 임펄스응답으로 부터 다른 임펄스응답으로 한샘플만에 급격히 변화하는 경우에 대해 실험하였으므로 실제로는 시간적 여유를 두고 임펄스응답이 변화하므로 이보다는 더욱 작게 나타난다. 표 1은 8가지 임펄스응답의 측정시 환경을 나타낸다.

또한  $D(n)$ 의 증가는 원단 신호의 통계적 특성이 변화한 경우에도 즉 이웃 음소간에 통계적 특성이 급격하게 변화한 경우에도 일어날 수 있다. 그러므로 원단 신호  $x(n)$ 에도 격자 예측기를 설치하여 원단 신호의 통계적 특성이 급격하게 변화하였는지에 대한 정보

를 근단의 격자 예측기에 전달하여야 한다.

표 1. 8가지 임펄스응답

Table 1. 8 kinds of impulse response.

1	한사람이 앉아 있을 때
2	두사람이 앉아 있을 때
3	두사람이 앉아 있고 한 사람이 뒤에 서 있을 때
4	한사람이 앉아 있고 다른 사람이 등 뒤로 지나갈 때
5	문이 열릴 때
6	마이크로폰을 5cm 가량 이동시킨 후
7	마이크로폰을 10cm 가량 이동시킨 후
8	마이크로폰을 화자쪽으로 가깝게 이동시킨 후

근단 신호에 대한 반사계수의 변화도를  $D_x(n)$ 이라 하고 원단 신호에 대한 반사계수의 변화도를  $D_x(n)$ 이라 하면 제안된 동시통화 검출기는 다음과 같이 동작한다. 먼저 원단 신호로부터  $D_x(n)$ 을 계산한다. 이로부터 동시통화 검출을 위한 문턱치  $D_{th}(n)$ 을 계산한다. 근단 신호로부터  $D_x(n)$ 을 계산한다.  $D_x(n)$ 을  $D_{th}(n)$ 과 비교하여 큰 경우 동시통화로 판단하고 필터계수의 적응을 멈추고 그렇지 않은 경우 계속해서 필터계수를 적응시킨다.  $D_{th}(n)$ 은 다음과 같이 구해진다.

$$D_{th}(n) = \gamma * \max\{D_x(n), D_x(n-1), \dots, D_x(n-M+1)\} \quad (9)$$

현재의 반사계수 변화도  $D_x(n)$ 과  $M-1$ 개의 과거 반사계수 변화도중 가장 큰 값을 구하고 상수  $\gamma$ 를 곱하므로써 얻어진다. 여기서  $\gamma$ 는 1보다 큰 값이며 반향 신호대 잡음비를 고려하여 설정되어야 한다.  $\gamma$ 를 작게 설정하면 동시통화에 민감하게 반응하지만 음성 신호의 통계적 특성의 변화도 동시통화로 잘못 판단할 수 있으므로 약간 높게 설정한다.

#### IV. 실험 및 결과 고찰

본 절에서는 제안된 동시통화 검출기의 성능을 컴퓨터 모의 실험을 통하여 검증한다. 성능 평가는 동시통화가 발생하는 경우와 반향경로가 변화하는 경우에 대하여 수행하였고 사용된 신호는 실제 음성신호이다. 비교 대상으로 사용된 기존의 동시통화 검출 방법은 크기 비교 형태의 동시통화 검출 알고리즘중 (4)의 방법이다. 공정한 성능 평가를 위해서 기존의 방법과 제

안된 방법 모두 동일한 적응 방향 제거기를 사용하였으며 사용된 적응디지털필터는 LMS알고리즘을 사용하는 트랜스버설 필터이다. 또한 모의 실험에 사용된 임펄스응답은 표 1의 첫번째 임펄스응답으로 512 샘플이다.

그림 3은 모의 실험에 사용된 원단 신호이다. 그림 4는 반향 신호대 잡음비를 10dB, 15dB, 18dB, 20dB로 바꾸어가며 반향제거기를 동작시킬 경우의 기존 방법의 검출인자인 ERLE\_DTD(n)을 나타낸 것이다. 이 때 적응디지털필터는 정상상태에 도달한 이후이다. 음성 신호가 큰 구간에서는 값이 크게 나타나는 반면 음소의 경계 구간에서는 작은 값을 나타내고 있다. 또한 반향 신호대 잡음비가 좋을수록 큰 값을 나타내며 나쁠수록 작은 값을 갖는다. 그러므로 실제 음성 신호를 사용하는 경우 기존의 방법은 음소의 경계에서 일어나는 감쇄와 반향 신호대 잡음비가 나빠지는 경우까 지 고려하여 문턱치를 결정하여야만 한다.

본 논문에서는 반향 신호대 잡음비가 20dB인 경우에 대해 모의 실험을 수행하였다. 실제 환경에서 잡음은 시간에 따라 변화하므로 반향 신호대 잡음비가 17dB 까지 나뉠 수 있는 경우를 고려하여 기존의 방법의 문턱치는 7.5dB를 사용하였다. 제안된 방법에서는 (9)의 방법으로 동시통화를 검출하게 되며 이 때 M은 64로 하였다. 비례상수  $\gamma$ 는 음성신호의 시간에 따른 특성변화와 반향경로가 변화한 경우 발생할 수 있는 오동작을 피하기 위해 4.5를 사용하였다. 격자 예측기의 차수는 4차이며 가중치 함수는 rectangular 함수를 사용하였다.

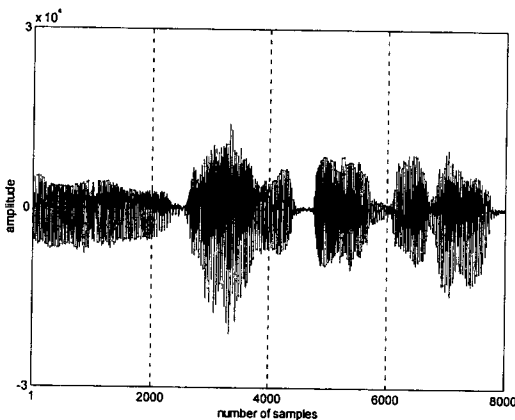


그림 3. 원단 신호  
Fig. 3. The far end signal.

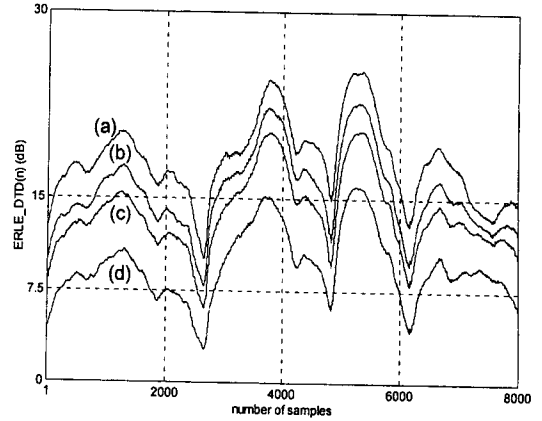


그림 4. 반향 신호대 잡음비에 따른 ERLE\_DTD(n)  
(a) 20dB (b) 18dB (c) 15dB  
(d) 10dB

Fig. 4. ERLE\_DTD(n) according to variation of the echo to noise ratio.  
(a) 20dB (b) 18dB (c) 15dB  
(d) 10dB

1. 반향경로 변화

그림 5는 표 1의 1번 임펄스응답에서 6번 임펄스응답으로 반향경로가 변화한 경우의 기존의 방법과 제안된 방법과의 성능을 나타낸다. 이때 반향경로의 변화는 실제 상황과 유사하도록 80샘플 동안 점진적으로 변화하도록 하였다. (a),(b)는 각각 기존의 방법과 제안된 방법의 성능을 나타내는 ERLE 곡선이다. (c)는 기존의 검출방법에서 사용하는 ERLE\_DTD(n) 값을 나타내며 (d)는 제안된 검출 방법에서 사용하는 검출인자 값으로 근단 신호  $z(n)$ 을 입력으로 하는 격자 예측기로부터 얻어지는  $D(n)$  값과 (9)에서 계산된  $D_x(n)$ 과 M-1개의 과거 반사 계수중 가장 큰 값의 비율로서  $Ratio(n)$ 으로 정의한다.

기존의 방법에서는 반향경로가 변화한 경우 ERLE\_DTD(n)이 감소하여 문턱치인 7.5dB 이하로 까지 감소하게 되어 동시통화가 판정되어 필터 계수의 적응을 중단하게 된다. 그림 5를 살펴보면 기존의 방법도 문턱치를 4dB정도로 매우 낮게 선정하면 잘못된 판정을 막을 수 있다. 그러나 이 경우에는 실제 동시통화가 발생하면 동시통화 검출이 매우 늦어지게 되어 필터 계수가 발산할 가능성이 커진다.

반면 제안된 방법은 반향경로가 변화하여도  $Ratio(n)$  값은 어느 정도 이상 증가하지 않기때문에

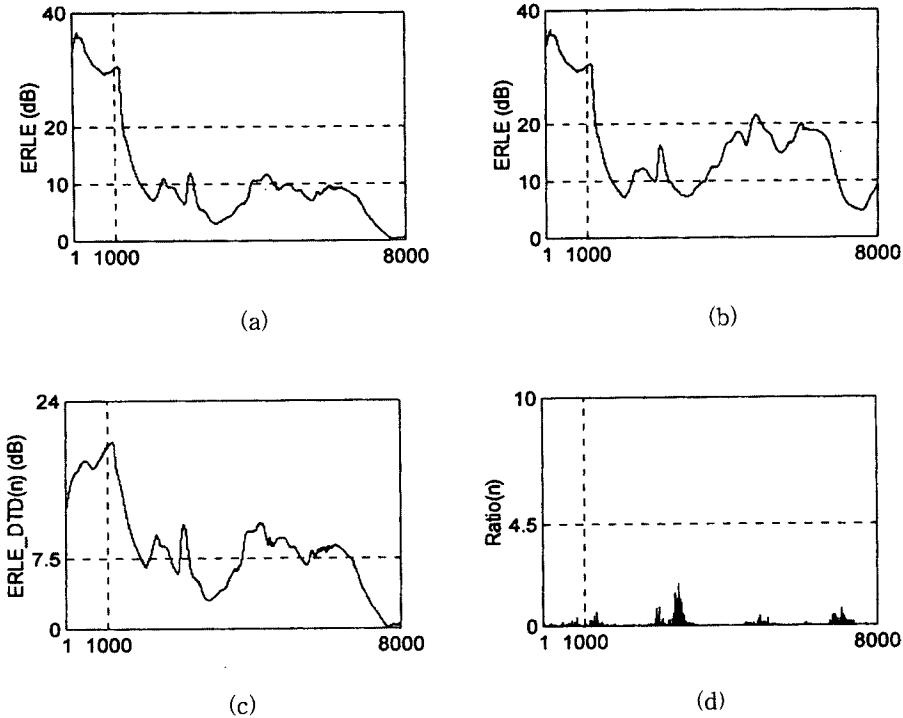


그림 5. 반향경로 변화시 성능  
 (a) ERLE(기존의 방법) (b) ERLE(제안된 방법) (c) ERLE\_DTD(n)  
 (d) Ration(n)  
 Fig. 5. The performance of the echo path variation.  
 (a) ERLE(conventional) (b) ERLE (proposed) (c) ERLE\_DTD(n)  
 (d) Ration(n)

계속해서 필터계수를 적응시킬 수 있다. 제안된 방법은 3000번째 샘플 부근에서 Ratio(n)이 증가하는 것을 볼 수 있는데 이것은 음성 신호의 특성이 급격하게 변화하기 때문이다. 그러므로 제안된 방법에서는 이를 고려하여 비례상수  $\nu$ 를 설정하였다.

2. 동시통화

동시통화시의 성능은 근단화자 신호가 무성음으로 시작하는 경우와 유성음으로 시작하는 경우로 나누어 실시하였으며 시작되는 위치를 변화시켜가며 수행하였다. 이 때 근단화자 신호대 반향 신호의 비는 10dB이다. 그림 6은 근단화자 신호의 파형을 나타낸다. (a)는 무성음으로 시작하는 경우이고 (b)는 유성음으로 시작하는 경우이다. 그림 7은 그림 6의 (a)에 나타난 무성음으로 시작하는 근단화자 신호에 의해 동시통화가 발생하는 경우의 성능을 나타낸다. 제안된 방법이 기존의 방법보다 빨리 근단화자 신호가 시작되었다고 판단하기 때문에 동시통화 구간에서 우수한 성능을 나타낸

다.

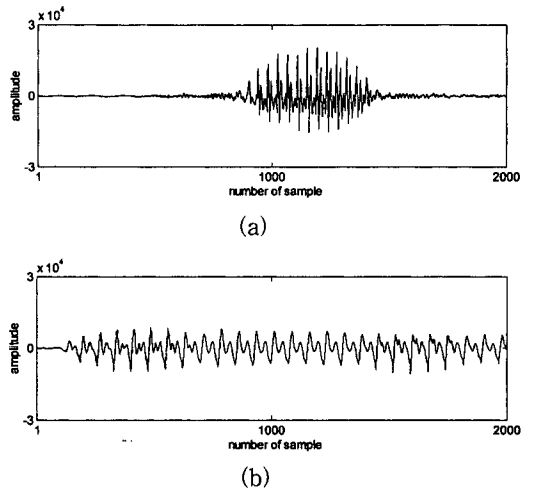


그림 6. 근단화자 신호  
 (a) 시작 : 무성음 (b) 시작 : 유성음  
 Fig. 6. The near end talker signal.  
 (a) start : unvoiced (b) start : voiced

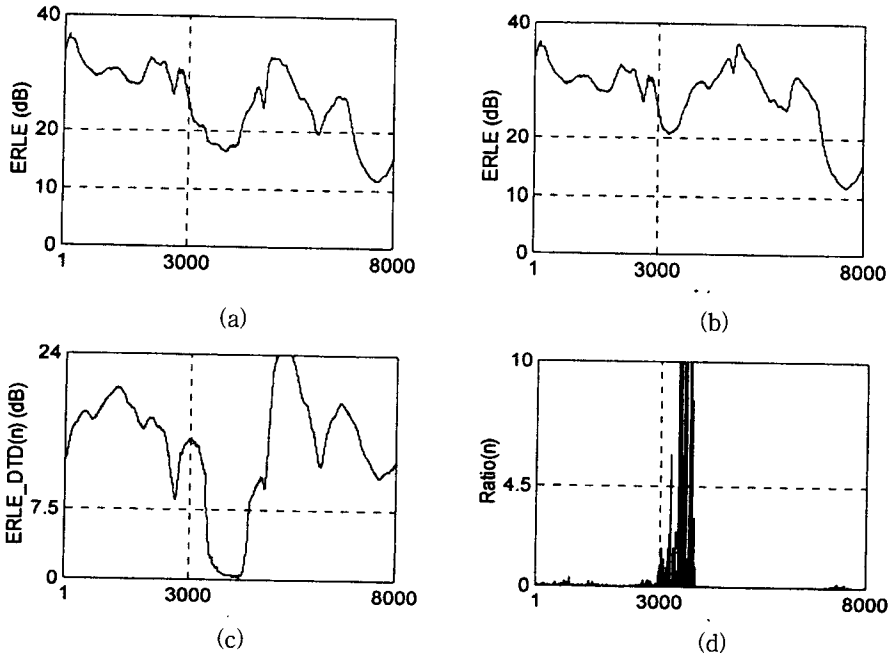


그림 7. 동시통화 성능(시작 : 무성음)

(a) ERLE(기존의 방법) (b) ERLE(제안된 방법) (c) ERLE\_DTD(n) (d) Ration(n)

Fig. 7. The performance of the double talk(start : unvoiced).

(a) ERLE(conventional) (b) ERLE(proposed) (c) ERLE\_DTD(n) (d) Ration(n)

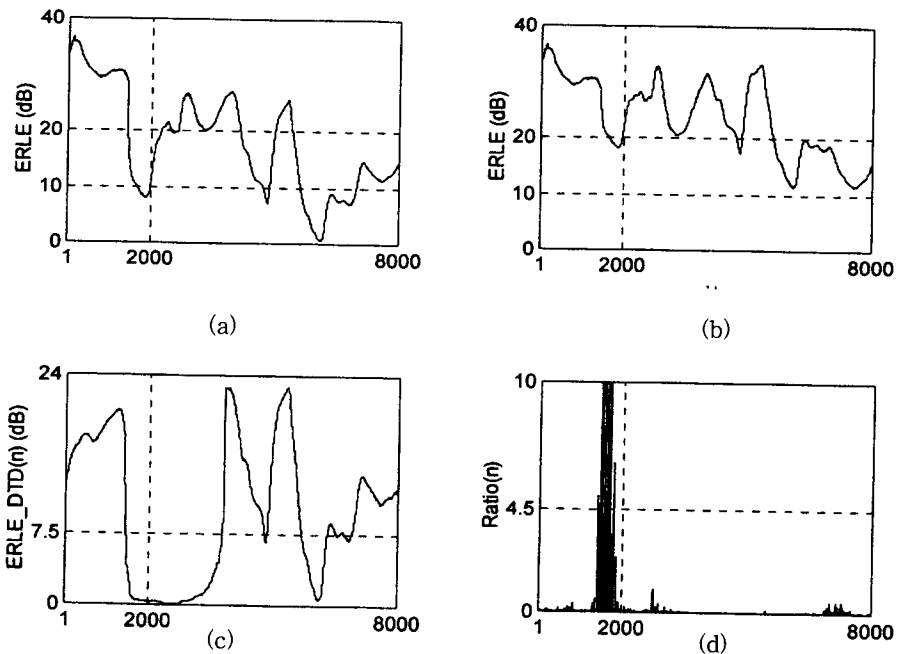


그림 8. 동시통화 성능(시작 : 유성음, 경우 1)

(a) ERLE(기존의 방법) (b) ERLE(제안된 방법) (c) ERLE\_DTD(n) (d) Ration(n)

Fig. 8. The performance of the double talk(start : voiced, case 1).

(a) ERLE(conventional) (b) ERLE (proposed) (c) ERLE\_DTD(n) (d) Ration(n)



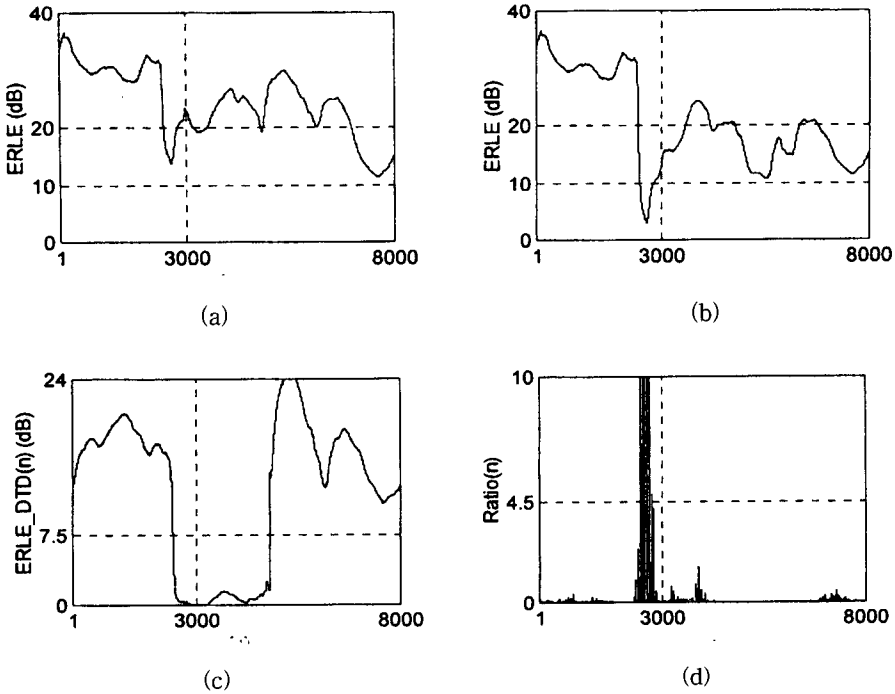


그림 9. 동시통화 성능(시작 : 유성음, 경우 2)  
 (a) ERLE(기존의 방법) (b) ERLE(제안된 방법) (c) ERLE\_DTD(n)  
 (d) Ration(n)  
 Fig. 9. The performance of the double talk(start : voiced, case 2).  
 (a) ERLE(conventional) (b) ERLE (proposed) (c) ERLE\_DTD(n)  
 (d) Ration(n)

그림 8은 그림 6의 (b)에 나타낸 유성음으로 시작하는 근단화자 신호에 의해 동시통화가 발생하는 경우의 성능을 나타낸다. 이 경우에도 제안된 방법이 기존의 방법보다 빨리 동시통화를 검출하기 때문에 동시통화구간에서 더 우수한 성능을 나타낸다. 그림 9는 그림 8에서 사용한 근단화자 신호와 동일한 신호에 의해 동시통화가 발생하나 동시통화가 일어나는 구간이 1000샘플 뒤에 일어나는 경우의 성능이다. 이 경우에는 기존의 방법이 제안된 방법보다 빨리 동시통화를 검출하여 더 우수한 성능을 나타냄을 볼 수 있다.

모의 실험 결과 동시통화 검출 성능 측면에서는 근단화자 신호의 시작위치에 따라 기존의 방법이 우수할 수도 있고 제안된 방법이 우수할 수도 있다. 그러나 제안된 방법은 반향경로가 변화한 경우에는 오동작을 하지 않을 뿐만 아니라 시간 지연 없이 두 상태를 구분할 수 있기 때문에 기존의 방법보다 훨씬 우수한 추적 성능을 갖는다.

### V. 결 론

제안된 동시통화 검출 알고리즘은 근단신호와 원단신호에 격자예측기를 설치하여 반사계수를 구하고 그 변화도를 이용하여 동시통화와 반향경로가 변화한 경우를 구분한다. 동시통화가 발생하면 반사계수의 변화도가 급격히 커지며 반향경로가 변화한 경우에는 정상 상태에서와 유사한 변화도를 갖게 된다. 두 상태의 구분을 위한 문턱치는 원단에 설치된 격자예측기의 반사계수 변화도로 부터 구해진다.

모의 실험 결과 제안된 동시통화 검출 알고리즘은 반향경로가 변화한 경우 판단지연이 없기때문에 기존의 동시통화 검출 알고리즘을 사용하는 적응필터 보다 항상 우수한 추적 성능을 갖는다. 동시통화의 경우에는 근단화자 신호의 시작 위치에 따라 기존의 크기 비교 형태의 알고리즘이 우수한 경우도 있고 제안된 방법이 우수한 경우도 있었다. 그러므로 제안된 방법을

기존의 크기 비교법과 함께 사용할 경우 제안된 방법만 쓰는 경우보다 어느 정도 더 우수한 성능을 기대할 수 있다.

참 고 문 헌

[1] 유재하, 조성호, 윤대희, 차일환 “축소격자필터 구조를 사용한 음향반향제거기”, 대한전자공학회 논문집 제 32권 B편 제11호, pp. 1473~1480, 1995년 11월

[2] P. C. Yip and D. M. Etter, “An Adaptive Multiple Echo Canceller for Slowly Time-Varying Echo Paths,” *IEEE Trans on Commun*, vol. 38, no. 10, pp. 1693~1698, Oct. 1990.

[3] Nobuji Furuya, Yuhzoh Fukushi, Yoshinori Itoh, Junji Tanabe and Takashi Araseki, “High Performance Custom VLSI Echo Canceller,” *Proc. ICC*, pp. 1470~1476, 1985.

[4] Padma P. Mallela, Mansoor A. Chishtie and Sen M. Kuo, “A TMS320C5x DSP-Based Full-Duplex Speaker Phone,” *Proc. International Conference on Signal Processing Applications and Technology*, vol. 1, pp. 1~6, 1994.

[5] Peter L. Chu, “Weaver SSB Subband Acoustic Echo Canceller,” *Proc. IEEE Workshop on Application of Signal*

*Processing to Audio/Acoustics*, 1993.

[6] Hua Ye and Bo-Xiu Wu, “A New Double-Talk Detection Algorithm Based on the Orthogonality Theorem,” *IEEE Trans on Commun*, vol. 39, no. 11, pp. 1542~1545, Nov. 1991.

[7] 유재하, 윤대희, 조성호, “효율적인 동시통화 검출을 위한 새로운 알고리즘,” 한국통신학회 논문지 제20권 제 7 호, pp. 1814~1821, 1995

[8] S. Minami and T. Kawasaki, “A Double Talk Detection Method for an Echo Canceller,” *Proc. ICC*, pp. 1492~1497, 1985.

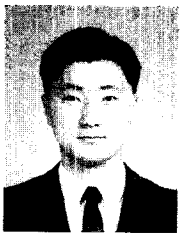
[9] F. Amano, H. P. Meana, A. D. Luca and G. Duchen, “A Multirate Acoustic Echo Canceller,” *IEEE Trans on Commun*, vol. 43, no. 7, pp. 2172~2176, Jul, 1995.

[10] F. Soong, “On the Automatic Segmentation of Speech Signals,” *Proc. ICASSP*, pp. 77~80, 1987.

[11] P. Alexandre and P. Lockwood, “Root Cepstral Analysis : A Unified View. Application to Speech Processing in Car Noise Environments,” *Speech Communication*, vol. 12, pp. 277~288, 1993.

[12] A. M. Kondoz, *Digital Speech(Coding for Low Bit Rate Communication Systems)*, John Wiley & Sons., 1994.

저 자 소 개



柳 在 河(正會員)

1966년 11월 21일생. 1990년 2월 연세대학교 전자공학과 졸업(공학사). 1992년 2월 연세대학원 전자공학과 졸업(공학석사). 1996년 2월 연세대학교 대학원 전자공학과 졸업(공학박사). 1996년 1월 ~ 현재

LG전자 멀티미디어 연구소 선임연구원. 관심분야는 적응 디지털 필터 응용, 음성 신호처리



趙 誠 鎬(正會員)

1960년 2월 21일생. 1982년 2월 한양대학교 전자공학과 졸업(공학사). 1984년 12월 University of Iowa 전기 및 컴퓨터공학과 졸업(공학석사). 1989년 8월 University of Utah 전기 및 컴

퓨터공학과 졸업(공학박사). 1989년 8월 ~ 1992년 8월 한국전자통신연구소 후호기술부 선임연구원. 1992년 9월 ~ 현재 한양대학교 전자공학과 조교수, 부교수. 관심분야는 디지털통신, 무선통신, 신호처리, 정보통신 시스템