

# 핸즈프리 전화통신을 위한 잔여반향제거

박선준, 조점군, 이지하, 차일환, 윤대희  
연세대학교 전기·컴퓨터공학과

## Residual Echo Cancellation for Hands-Free Telephony

Seon Joon Park, Chom Kun Cho, Ji Ha Lee, Il Whan Cha, Dae Hee Youn  
Dept. of Electrical and Computer Engineering, Yonsei University  
E-mail: seonjoon@lethe.yonsei.ac.kr

### 요약

본 논문에서는 차량 환경에서 핸드프리 단말기를 위한 잔향반향제거 방법을 제안한다. 제안된 방법은 기존의 음향반향제거와 잡음제거의 결합구조에 근거하며, 음성신호의 스펙트럼 특성을 배경잡음화함으로써 잔여반향제거 성능을 향상시킨다. 일반적으로 음향반향제거에서 실제 충격응답보다 적은 차수의 적응필터를 이용할 경우 잔여반향의 전력이 증가하며, 잡음제거기법을 적용하여 잔여반향성분을 줄일 수 있다. 음성신호가 입력되는 음향반향제거기의 잔여반향을 효과적으로 제거하기 위해 음성신호의 AR 스펙트럼에 따른 역필터링을 수행함으로써 잡음제거기에 의한 잔여반향제거 성능을 향상시킬 수 있다. 제안된 기법은 현재 상용화되고 있는 이동통신용 음성부호화기에 포함된 잡음제거기법과 결합하여 사용할 경우 매우 적은 부가 계산량만으로 구현할 수 있다.

### 1. 서론

이동통신 기술을 이용한 전화통신이 사회 전분야에 확산되면서 핸드프리 통신에 적합한 음질개선방법이 크게 요구되고 있다. 이러한 관점에서 해결해야 할 두가지 문제가 있다. 첫째로 스피커와 마이크로폰의 커플링에 의하여 음향반향이 발생하는 문제가 있으며, 둘째로 전송해야 하는 근단화자의 신호가 배경잡음에 의하여 손상되는 문제가 있다[1].

음향반향은 원단화자 음성이 근단화자 핸드프리 단말기의 스피커와 마이크로폰 사이의 경로를 통해 귀환됨으로써 발생한다. 음향반향은 통화 품질을 저해하는 가장 주요한 요소로, 2선으로 연결된 스피커폰에서의

하이브리드/음향 루프에서의 하울링을 방지하고 원단화자가 자신의 반향신호를 듣는 것을 방지하기 위하여 반드시 제거되어야 한다. 배경잡음제거는 운행 중인 차량 실내와 같이 잡음이 심한 상황에서 핸드프리 단말기를 사용하여 통화를 할 경우에 특히 중요한 문제로 대두된다.

최근 들어 핸드프리 상황에서 이 두가지 문제를 동시에 해결하기 위하여 음향반향제거와 배경잡음제거를 결합하는 문제가 크게 대두되고 있다[2][3]. 특히 반향제거기와 잡음제거기를 결합하여 사용할 경우에 반향제거기의 추정오차로 발생하는 잔여반향성분을 추가로 제거할 수 있다는 장점을 갖게 된다. 일반적으로 음향반향제거기에서는 계산량 문제로 인하여 실제 충격응답보다 적은 차수의 적응필터를 이용할 수밖에 없고, 따라서 잔여반향(residual echo)의 전력이 증가하게 된다. 이 경우 증가한 잔여반향의 전력을 잡음제거기에 반영하여 잔여반향성분을 줄일 수 있다. 본 논문에서는 잔여반향의 AR 분석을 통해 역필터링을 수행함으로써 잡음제거기에 의한 잔여반향제거 성능을 향상시켰다.

### 2. 반향제거기와 잡음제거기의 결합구조

그림 1은 일반적인 반향제거기와 잡음제거기의 결합구조를 보여주고 있다. 여기서  $H$ 는 FIR 필터 구조를 가지며 NLMS 알고리즘을 이용하여 계수적용을 하는 기존의 반향제거기이고,  $G$ 는 일반적인 배경잡음제거기이다. 여기서  $x(k)$ 는 원단화자의 신호,  $y(k)$ 는 근단화자의 신호,  $n(k)$ 는 잡음신호를 의미한다. 따라서, 원단화자의 신호가 반향경로를 통과한 반향신호를  $d(k)$ 라고 할 경우, 근단화자 측의 마이크로폰으로 입력되는 신호  $y(k)$ 는 (1)과 같이 표현된다.

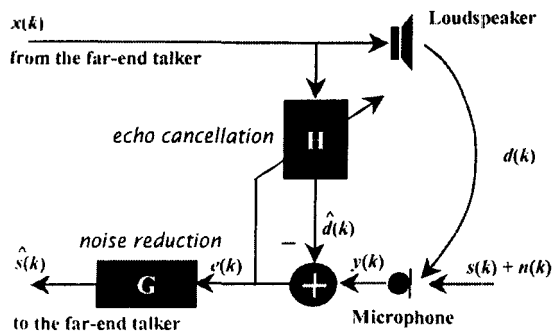


그림 1. 반향제거와 잡음제거의 결합구조

$$y(k) = s(k) + n(k) + d(k) \quad (1)$$

오차신호  $e(k)$ 은  $y(k)$ 로부터 추정된  $\hat{d}(k)$ 를 빼줌으로써 (2)와 같이 얻을 수 있다.

$$\begin{aligned} e(k) &= y(k) - \hat{d}(k) \\ &= s(k) + n(k) + d(k) - \hat{d}(k) \\ &= s(k) + n(k) + r(k) \end{aligned} \quad (2)$$

여기서  $r(k) = d(k) - \hat{d}(k)$ 는 반향제거기의 성능에 따라 전력의 크기가 좌우되는 잔여반향을 나타낸다.

잔여반향은 적응알고리즘이 불일치(misadjustment)하는 경우와 실제 시스템에서 제한된 길이의 필터를 사용해야하는 제약 조건이 있는 경우에 발생하게 된다. 이를 제거하기 위한 기존의 방법은 이득 스위칭 제어 (gain switching control)에 의해 잔여반향을 제거 혹은 억제하는 것이다[4]. 하지만 이 경우 동시통화(double talk) 구간동안 신호를 전송하지 않는 half-duplex 상황이 되고 잡음 대비(noise contrast)가 발생하는 단점이 있다.

이러한 문제점을 개선하기 위하여 그림 1과 같이 후처리를 사용하여 잔여반향신호와 배경잡음을 동시에 제거하는 방법이 제안되고 있다[5][6]. 일반적으로 잡음제거기법에서는 배경잡음이 WSS(Wide Sense Stationary) 신호이며 시간에 따라 천천히 변화한다고 가정한다. 음성신호에 천천히 변화하는 백색 혹은 유색 잡음이 섞여 있는 경우 잡음만을 추정하여 제거함으로써 깨끗한 음성신호를 얻을 수 있다. 그러므로 잔여반향신호가 WSS 특성을 갖는 백색 혹은 유색신호일 경우에 이를 배경잡음추정에 반영시켜 제거할 수 있다. 하지만 반향신호의 발생원이 음성신호인 경우에는 반향제거기의 출력으로 발생하는 잔여반향신호가 음성신호의 특성을 갖게되므로 기존의 잡음제거기만으로는 효과적으로 잔여반향을 제거할 수 없다.

원단화자의 신호가 음성신호이고 근단화자의 신호와 배경잡음이 없는 경우 잔여반향신호의 시간파형과 스펙트로그램을 그림 2에 보여주고 있다. 스펙트로그램에서 볼 수 있듯이 잔여반향은 신호 레벨이 작지만 음성 신호와 유사한 특성을 갖는다. 따라서 잡음제거기에서는 잔여반향성분을 배경잡음이 아닌 음성신호로 판단

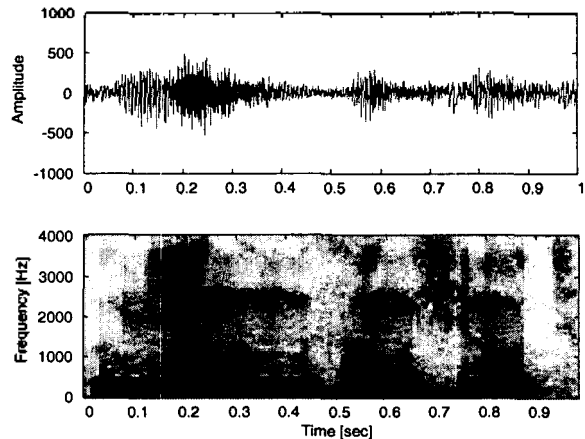


그림 2. 잔여반향신호의 시간파형 및 스펙트로그램

하게 되고 잔여반향신호를 그대로 원단화자에게 전송하게 된다. 그러므로 배경잡음을 WSS 신호로 가정하여 추정한 후 제거하는 일반적인 잡음제거기를 후처리기로 이용할 경우 잔여반향을 효과적으로 제거할 수 없음을 쉽게 알 수 있다.

### 3. 잔여반향제거

본 논문에서는 근단화자 신호가 없는 경우 잔여반향신호의 특성을 배경잡음화시켜 잡음제거기의 입력으로 사용함으로써 잔여반향제거성능을 향상시키는 방법을 제안한다. 제안한 시스템의 구조는 그림 3과 같다.

앞장에서 언급한 바와 같이 원단화자신호가 음성신호인 경우 잔여반향신호  $r(k)$ 는 음성신호의 특성을 가지므로  $P$ 차의 AR 모델로 근사화할 수 있고, 또한 근사화한 AR 계수를 이용하여 역필터링함으로써 다음과 같이 백색화(whitening)할 수 있다.

$$r(k) = \sum_{i=1}^P a(i)r(k-i) + Gu(k) \quad (3)$$

$$w_r(k) = - \sum_{i=0}^{P-1} a(i)r(k-i), \quad a(0) = 1 \quad (4)$$

여기서  $a(i)$ 는 AR 계수를 나타내며,  $w_r(k)$ 는  $r(k)$ 가

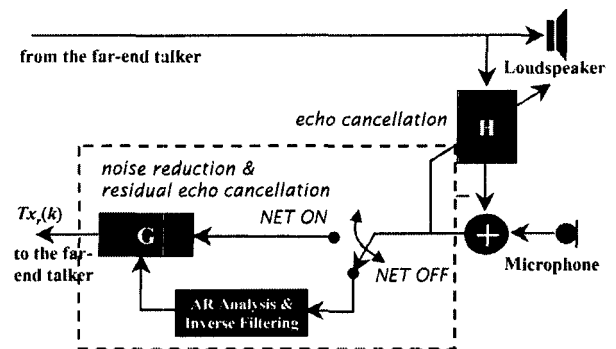


그림 3. 제안된 잔여반향제거 기법

백색화된 신호를 나타낸다. 이 때, 근단화자의 음성신호가 왜곡되어 전송되는 현상을 방지하기 위하여 근단화자 신호의 유무를 판단하여 근단화자의 신호가 없는 경우에만 AR 분석과 역필터링과정을 수행한다.

제안된 방법을 차량용 핸즈프리 단말기에 적용할 경우 일반적인 이동통신용 음성부호화기의 부가기능으로 규정되어 있는 배경잡음제거기법을 사용할 수 있다. 이 경우에는 (3)과 (4) 과정만을 추가하여 구현할 수 있으므로 아주 적은 계산량을 부가함으로써 잔여반향효과적으로 제거할 수 있다.

#### 4. 모의실험

모의실험을 위하여  $N$ 차의 FIR 적응필터를 구동하기 위한 적응 알고리즘으로는 NLMS 알고리즘을 사용하였으며, 잡음제거를 위해서는 IS-127에 규정되어 있고 EVRC 음성부호화기의 부가기능으로 포함되어있는 잡음제거방법을 사용하였다. 제안된 잔여반향제거기법에 사용된 AR 분석을 위해서는  $P=10$ 을 사용하였다. 반향신호는 8kHz 표본화율에서 1024탭의 음향반향경로를 이용하여 모델링하였다.

그림 4(a)에는 잡음제거기의 출력을 통과하지 않은 상태에서 반향제거기의 출력에 발생한 잔여반향신호와 그

스펙트로그램을 보여주고 있다. 이때 반향경로를 추정하기 위하여 512탭의 적응필터를 사용하였으며, 근단화자단의 마이크로폰으로 입력되는 잡음신호의 반향 대 잡음비(ENR; Echo-to-Noise Ratio)는 10dB로 하였다.

그림 4 (b)와 (c)에는 각각 기존결합구조와 제안된 방법을 사용한 경우에 원단화자에게로 돌아가는 잔여반향신호의 시간파형과 스펙트로그램을 보여주고 있다. 잡음제거기만을 사용한 경우에 전송되는 잔여반향신호의 레벨이 많이 감소하였지만, 잡음제거기가 잔여반향신호  $x(k)$ 를 음성신호로 판단하여 음성신호의 특성이 많이 남아있는 것을 볼 수 있다. 하지만 제안된 방법을 사용하여  $x(k)$ 를 처리한 경우에는 음성신호의 특성이 많이 제거되었음을 볼 수 있다. 따라서 원단화자 신호가 불연속이 없는 자연스러운 통화환경에서 잔여반향신호를 편안한 잡음(comfort noise)으로 인지할 수 있으며, 간단한 주관적 청취실험 결과 이를 확인할 수 있었다.

제안된 잔여반향제거기법의 객관적 성능을 평가하기 위하여 ERLE(Echo Return Loss Enhancement)를 측정하였다. 그림 5와 6에는 ENR이  $\infty(n(k)=0)$ 인 경우와 10dB인 경우에 대하여 ERLE를 나타내었다. 두 상황에 대하여 반향경로의 길이와 같은 탭수의 적응필터만을 반향제거기로 사용한 경우, 그림 1의 기존 결합구조에

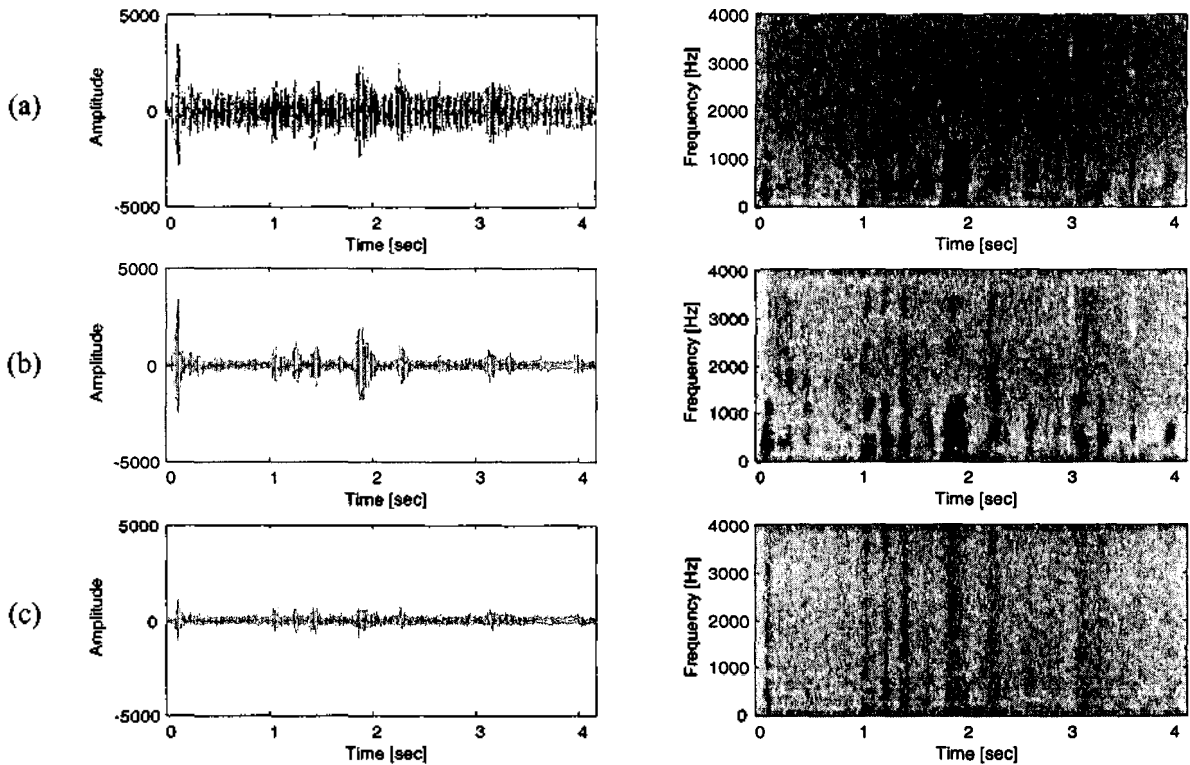


그림 4. ENR이 10dB인 경우 잔여반향신호의 시간파형 및 스펙트로그램  
(a) 잡음제거기를 통과하지 않은 경우, (b) 기존 방법(그림 1), (c) 제안된 방법(그림 3)

512탭의 적응필터를 사용한 경우, 그림 3의 제안된 구조에 512탭의 적응필터를 사용한 경우를 비교하였다.

배경잡음이 없는 그림 5의 결과를 보면, 기존의 결합구조보다 제안된 구조가 평균적으로 8.9dB의 ERLE가 향상된 것을 볼 수 있다. 또한 제안된 구조를 사용할 경우 반향경로길이 전체를 적응필터가 모델링한 경우보다도 평균 1.9dB의 ERLE가 향상되었다. 특히, 제안된 방법을 사용한 경우 낮은 차수의 적응필터를 사용함으로써 적응필터의 차수가 높을수록 초기부분에서 수렴속도가 저하되는 문제를 어느 정도 극복할 수 있었다.

ENR이 10dB인 경우에는 배경잡음이 적응필터의 추정오차를 증가시키므로 반향경로전체를 모델링한 것보다 512탭의 적응필터와 잡음제거기를 결합한 구조가 좀더 향상된 성능을 가짐을 볼 수 있다. 또한 제안된 구조를 사용할 경우에는 평균 7.4dB의 ERLE 성능향상을 얻을 수 있었으며, 잡음환경에서도 평균 20dB 이상의 ERLE 성능을 얻을 수 있었다. 따라서 제안된 잔여반향

제거기법을 이동통신에서 핸드프리 상황에 적용할 경우 배경잡음이 존재하는 상황에서도 반향제거기의 성능을 크게 향상시킬 수 있음을 알 수 있다.

## 5. 결론

본 논문에서는 반향제거기의 추정오차로 인한 잔여반향과 배경잡음을 억제하기 위하여 반향제거기와 잡음제거기의 결합구조를 사용하였다. 핸드프리 음성통신환경에서 발생하는 반향신호가 음성신호의 특성을 갖는다는 사실에 기반하여 음성 분석을 통한 역필터링을 통해 잔여반향을 크게 줄일 수 있다. 간단한 주관적 음질평가 결과 제안한 방법은 기존의 결합구조만을 사용하는 방법에 비하여 전송되는 잔여반향신호로 인한 원단화자의 불편함을 크게 감소시킬 수 있었다. 제안한 방법은 이동통신에 사용되는 저전송율 음성부화기의 부가기능으로 포함된 배경잡음기법을 이용할 경우 매우 적은 부가 계산량만으로 구현할 수 있다.

## 참고 문헌

- [1] P. Dreiseitel, E. Hänslar, and H. Puder, "Acoustic echo and noise control - a long lasting challenge," in *Proc. EUSIPCO*, pp.945-952, 1998.
- [2] S. Gustafsson, R. Martin, and P. Vary, "Combined acoustic echo control and noise reduction for hands-free telephony," *Signal Processing*, 64, pp.21-32, 1998.
- [3] R. L. Bouquin and G. Faucon, "How to reduce the noise influence in a joint system developed for echo and noise cancellation," *IEEE Signal Processing Letters*, vol.4, no.10, pp.280-282, Nov. 1997.
- [4] A. Gilloire and J-F Zürcher, "Achieving the control of the acoustic echo in audio terminal," in *Proc. EUSIPCO*, pp.491-494, 1988.
- [5] V. Turbin, A. Gilloire, and P. Scalart, "Comparison of three post-filtering algorithms for residual acoustic echo reduction," in *Proc. ICASSP*, pp. 307-310, 1997.
- [6] F. Basbug, K. Swaminathan, and S. Nandkumar, "Integrated noise reduction and echo cancellation for IS-136 systems," in *Proc. ICASSP*, pp. 1863-1866, 2000.

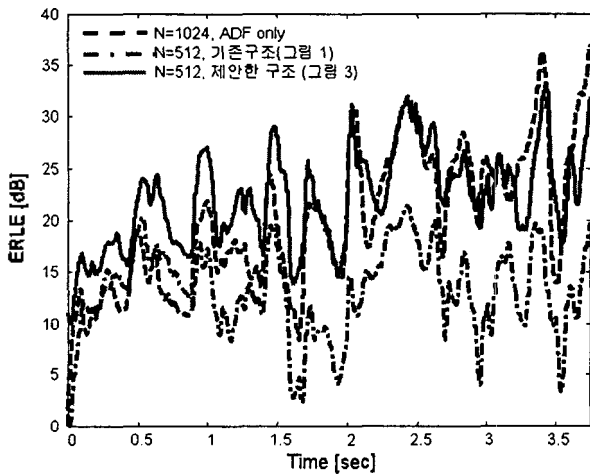


그림 5. ENR= $\infty$ 인 경우의 ERLE 곡선

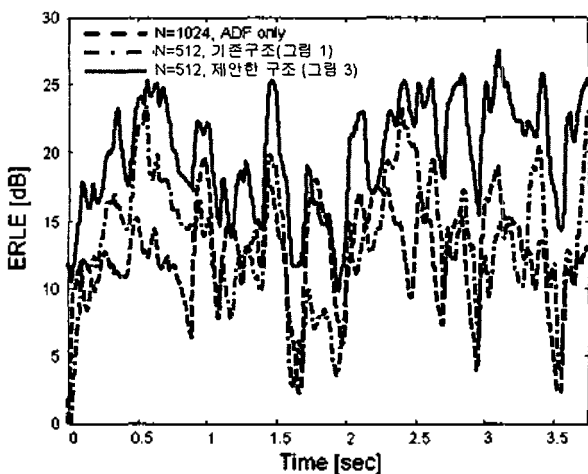


그림 6. ENR=10dB인 경우의 ERLE 곡선