

# 차량용 핸즈프리 통신을 위한 음향반향 및 잡음제거기의 실시간 구현

조점균, 박선준, 이충용, 윤대희  
연세대학교 전기·전자공학과

## Real-time Implementation of Acoustic Echo and Noise Canceller for Hands-free Communication in Car Environment

Chom Kun Cho, Seon Joon Park, Chungyong Lee, and Dae Hee Youn  
Dept. of Electrical and Electronic Eng., Yonsei University  
e-mail : [chronos, seonjoon]@lethe.yonsei.ac.kr

### 요 약

최근 이동전화의 사용이 급격히 확산됨에 따라 핸즈프리 단말기를 이용한 전화통신의 필요성이 대두되고 있다. 차량내 핸즈프리 통신상황의 경우 근거리에서 위치한 스피커와 마이크론의 커플링에 의해 발생하는 음향반향과 차량내에 존재하는 배경잡음은 통화 품질을 크게 저하시킨다. 본 논문에서는 이동 통신에 적합한 음향반향제거기와 잡음제거기의 결합시스템을 제안하고, 이를 고정 소수점 DSP를 이용하여 실시간 구현하였다. 실시간 구현을 위하여 음향반향제거기에는 NLMS 알고리즘에 의해 구동되는 제안된 차수의 적응반향제거기법을 사용하였다. 잔여반향 및 배경잡음제거를 위해 CDMA 방식의 셀룰라 이동통신에 사용되는 IS-127 EVRC 음성 부호화기의 표준안에 포함된 잡음제거 방식을 사용하였다. 제안된 시스템을 16 비트 고정소수점 DSP인 OAK DSP Core를 이용하여 약 18.6MIPS의 연산량으로 실시간 구현되었다.

### 1. 서 론

이동통신기술이 사회 전분야에 급격히 확산됨에 따라 안정성과 편의성을 위하여 핸즈프리를 이용한 전화통신의 중요성이 크게 대두되고 있다. 그러나 핸즈프리를 이용할 경우 필연적으로 발생하는 음향반향과 배경잡음은 통화 품질을 저해하는 큰 요소가 된다[1]. 특히 자동차 내부와 같이 잡음이 많은 폐공간에서 전화통화를 할 경우 발생하는 음향반향과 배경잡음은 사용자의 자연스러운 통화를 크게 저해하므로 반드시 제거하여야 한다.

음향반향은 원단화자의 음성이 근단화자의 인접한 핸즈프리 단말기의 스피커와 마이크론 사이의 경로를 통해 귀환됨으로써 발생하며 통화 품질을 저해하는 가장 주요한 요

소이다. 음향반향의 제거에는 일반적으로 적응필터를 이용하여 시간에 따라 변화하는 반향경로를 추정하는 적응반향제거기법이 널리 쓰인다. 배경잡음은 운행 중인 차량 실내와 같이 잡음이 심한 상황에서 이동통신 부호화기의 성능을 크게 저하시킨다. 따라서 핸즈프리 단말기를 사용하여 통화를 할 경우 음향반향과 배경잡음의 제거는 특히 중요한 문제로 대두된다[2][3]. 본 논문에서는 두 가지 문제를 동시에 해결하기 위하여 이동통신용 핸즈프리 통신을 위한 음향반향제거기와 배경잡음제거기를 결합한 시스템을 실시간 시스템을 제안하였다. 제안된 시스템은 16 비트 고정소수점 DSP인 OAK DSP를 이용하여 구현하였다.

### 2. 제안된 시스템

본 논문에서는 음향반향과 배경잡음을 동시에 효과적으로 제거하기 위하여 그림 1과 같이 음향반향제거기(AEC)의 후처리단으로 배경잡음제거기(NR)를 사용하는 실시간 시스템을 제안한다. 제안된 시스템은 크게 아래와 같은 두 부분으로 구성된다.

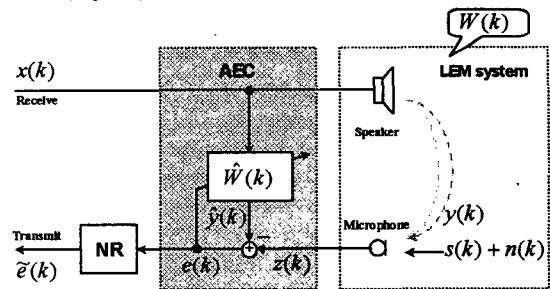


그림 1. 음향반향 및 배경잡음제거 시스템

- 1) 음향반향제거(AEC) : NLMS 알고리즘에 의해 구동되는 음향반향제거기를 사용[4].
- 2) 잔여반향 및 배경잡음제거(NR) : IS-127 잡음제거기 사용[5].

잡음제거기를 음향반향제거기의 후처리단으로 사용하는 결합구조에서는 두 프로세싱을 각각 독립적으로 수행하므로 프로세싱 간의 상호작용으로 인해 음향반향제거기가 불안정해지는 현상을 막을 수 있다. 그 밖에도 계산량과 수렴속도 등의 문제로 인하여 제한된 차수의 적응필터를 사용하는 음향반향제거기에서 필연적으로 발생하는 잔여반향을 잡음제거기를 이용하여 줄일 수 있는 장점이 있다[1]. 잡음제거기를 음향반향제거기의 전처리단으로 사용하는 결합구조를 고려할 경우에는 잡음 성분을 미리 제거하여 적응알고리즘의 잡음에 대한 민감도를 줄여줄 수 있는 장점이 있다. 하지만 음향반향제거기의 선형성을 보장하지 못하므로 올바른 반향경로를 추정하기가 어려운 단점이 있다[2].

따라서 본 논문에서는 그림 1과 같이 음향반향제거기의 후처리단으로 잡음제거기를 사용하는 구조를 사용하였다. 또한, 표준화된 IS-127 음성부호화기의 잡음제거기를 사용함으로써 음향반향제거기와 음성부호화기의 결합된 시스템의 구현을 용이하도록 하였다. 그림 1에서  $x(k)$ 는 원단화자신호,  $s(k)$ 는 근단화자신호,  $y(k)$ 는 음향반향신호,  $n(k)$ 는 잡음신호를 의미한다.  $\hat{y}(k)$ 는 적응필터를 이용해 추정된 반향신호,  $z(k)$ 는 마이크로폰의 입력신호,  $e(k)$ 는 음향반향제거기의 출력인 오차신호이고,  $\tilde{e}(k)$ 는 잡음제거기의 출력신호이다.

## 2.1 음향반향제거기

음향반향제거를 위한 적응알고리즘은 음성신호의 전력 특성을 반영하여 효과적으로 반향신호를 제거할 수 있으면서 계산량이 적은 NLMS 알고리즘을 사용하였다[4]. NLMS 알고리즘은 입력신호가 음성신호와 같이 시간에 따라 전력이 크게 변화하는 신호일 때, 지속적이고 안정적인 수렴을 보장하기 위해서 입력신호의 전력에 따라 가변하는 수렴상수를 사용한다. 음향반향제거기는 음향공간의 특성뿐만 아니라 마이크로폰과 스피커의 특성을 포함하는 폐공간 (loudspeaker-enclose-microphone; LEM) 시스템의 충격응답을 모델링한다.

## 2.2 IS-127 잡음제거기

잔여반향 및 배경잡음을 제거하기 위하여 일반적인 배경잡음제거기법을 사용할 수 있으나 본 논문에서는 이동통신용 음성부호화기와 결합을 고려하여 표준 규격 중 하나인 IS-127 (EVRC)의 잡음제거기를 사용하였다[5]. IS-127의 잡음제거기는 블록단위로 데이터를 처리하며, 이 때 블록크기는 80 샘플이다. 간략화한 IS-127의 잡음제거방법은 그림 2와 같다. 주파수 대역을 총 16개 비균일한(non-uniform) 대역으로 나누어 대역별로 SNR을 구하고 전대역에 대하여 입력신호의 스펙트럼 편차를 구한다. 이를 이용하여 구한 대역별 잡음제거 이득을 입력신호에 가함으로써 잡음을 제거한다.

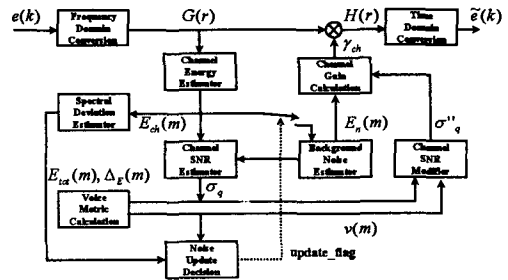


그림 2. IS-127의 잡음제거방법

## 3. 실시간 구현

제안된 시스템의 실시간 구현에 사용된 OAK DSP Core는 40 MIPS의 연산 속도를 갖는 16 비트 고정 소수점 DSP로써 저전력과 휴대성이 필수적인 디지털 셀룰라 단말기, 고속 모뎀 등의 통신 장비와 전자 응용기에 적합하게 설계되어 있다. 또한 일반 범용 DSP와 달리 시스템에 필요한 프로그램 메모리, 데이터 메모리, 외부 연결 회로 등을 ASIC 공정에서 한 개의 칩으로 집적할 수 있으므로 시스템 구현이 용이하다.

### 3.1 알고리즘 구현의 최적화

DSP는 한정된 저장공간과 제한된 연산속도를 갖으므로 사용하는 알고리즘의 최적화과정이 필요하다. 잡음제거기는 주파수 축에서 프로그램이 수행되므로 FFT/IFFT가 필수적으로 요구된다. 음성신호와 같이 실수 값만을 갖는 신호인 경우, DFT의 대칭성을 이용하여 2N 포인트 실수 DFT를 N 포인트 복소수 FFT로 계산할 수 있으며, 이로 인해 계산량을 절반으로 줄일 수 있다[6].

본 논문에서 사용한 최적화된 실수 FFT 알고리즘은 패킹(packaging) 알고리즘으로, 2N 포인트 실수 입력 데이터를 N 포인트 복소수 데이터로 패킷화하여 N 포인트 복소수 FFT를 수행하고 N 포인트 복소수 출력을 N 포인트 복소수 데이터로 언패킹(unpacking)한다[6]. 이러한 결과는 원래 2N 포인트 실수 입력 DFT 결과와 동일하므로 실수 FFT의 크기는 반으로 줄어들어서 일반적인 FFT보다 계산량을 절반으로 줄일 수 있다. 또한 이를 시간 간축 방식 FFT에 적용할 경우 고정소수점 오차에 의한 평균잡음 전력이 작고, 오버플로우(overflow)가 일어날 확률이 적어 고정 소수점 DSP를 이용한 실시간 구현에 용이하다.

### 3.2 연산량

OAK DSP의 클럭 주파수를  $f_{CLK}$ 이라 하고, 표준화 주파수를  $F_s$ 라고 하면, 하나의 음성 샘플을 처리하는데 최대한 사용할 수 있는 사이클 수는 식(1)과 같다.

$$f_{CLK} \times \frac{1}{F_s} = 400,000 \times \frac{1}{8,000} = 5,000 \text{ (cycles / sample)} \quad (1)$$

IS-127 잡음제거기는 80 샘플의 블록단위로 처리하므로, 제안된 시스템의 실시간 구현을 위해서는 하나의 데이터 블록이  $5,000 \times 80 = 400,000$  사이클 안에 처리해야 한다.

실제 차내에서 측정된 충격응답을 기반으로(4.1 절 참조) 음향반향제거기에 사용된 적응필터의 차수는 256 으로 결정하였으며, 동시통화검출을 위해서는 상호상관도를 이용한 방법을 사용하였다[7]. 제안된 시스템이 한 블록의 데이터를 처리하는데 각 과정에서 소요되는 연산량은 표 1과 같다. 적응필터의 계수적용 과정이 58.0%로 가장 많은 연산량을 차지하였으며, FFT/IFFT, 입력신호의 필터링, 대역별 신호대잡음비 계산이 각각 15.8%, 12.5%, 3.0%를 차지하였다. OAK DSP 를 이용한 실시간 구현시 음향반향제거기와 잡음제거기에 필요한 연산량을 표 2에 나타냈다. 음향반향제거기와 잡음제거기를 위해서 각각 약 13.7, 4.6MIPS 가 필요하였으며, 입출력과정을 포함한 전체 시스템을 구현하는데 약 18.6 MIPS (사용가능 연산량의 약 46.5%)가 필요하였다.

표 1. 결합구조에서 각 부분별 연산량 (1 블록 당)

항수	연산량(cycles)	비율(%)
필터링	23,200	12.5
동시통화검출기	3,200	1.7
계수적용	108,000	58
전력 예측	2,400	1.3
FFT/IFFT	29,405	15.8
대역별 전력	3,700	2.0
대역별 신호대잡음비	5,499	3.0
스펙트럼 편차	2,000	1.1
잡음구간판별과 갱신	2,996	1.6
대역별 이득치	1,594	0.9
기타	4,068	2.1
합계	186,062	100.0

표 2. 전체 연산량

	연산량 (MIPS)	비율(%)
입출력 버퍼링	0.35	0.9
음향반향제거기	13.68	34.2
잡음제거기	4.57	11.4
합계	18.6	46.5

## 4. 시스템 평가

### 4.1 차내 충격응답의 측정

마이크로폰을 운전석 햇빛 가리개의 중앙에, 스피커를 시가용 핸즈프리가 놓이는 곳에 위치시킨 후 차내 충격응답을 측정하였다. 그림 3은 차의 엔진과 에어컨을 모두 켜진 경우에

측정한 차내 충격응답으로, 8kHz 표본화율에서 512 차의 FIR 필터로 모델링한 후 정규화한 것이다.

본 논문에서 사용한 NLMS 알고리즘의 연산량은 적응차수에 비례하기 때문에 실시간 구현을 위해서는 적응필터 차수의 결정이 매우 중요하다. 그림 3의 경우 8kHz 표본화율에서 충격응답의 길이가 256 이상이던 반향신호의 전력에 99% 이상이 포함된다. 즉, 256 차의 적응필터를 사용할 경우 차내 음향반향 경로의 대부분을 모델링할 수 있음을 알 수 있다. 차내에서 측정된 대부분의 충격응답에서 99% 이상의 에너지가 256 탭 이내에 있음을 확인하였다. 따라서 실시간 시스템에 사용되는 적응필터의 차수를 256 으로 결정하였다. 적응필터의 수렴상수  $\mu$ 는 실험을 통하여 0.3 으로 결정하였다. 실제 음성 신호를 사용할 경우,  $0.2 < \mu < 0.5$  일 경우에 ERLE 성능에 큰 차이를 보이지 않았다.

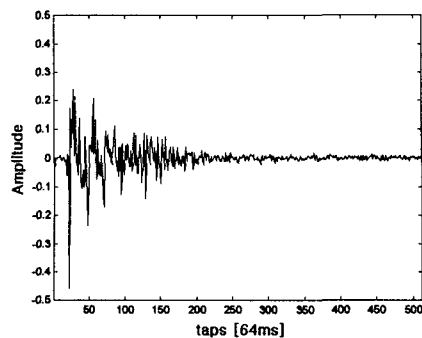


그림 3. 정규화한 차내 충격응답( $\|h_{norm}(\cdot)\|^2 = 1$ )

### 4.2 단일통화 상황

그림 4는 차내에서 실측된 단일통화 상황에서의 음성 신호이다. 그림 4 (a)는 스피커를 통해서 차안으로 방사된 원단파신호이며, 그림 4 (b)는 원단파자의 음향반향성분과 차내 배경잡음이 섞여 마이크로폰으로 입력된 신호이다. 제안된 시스템의 성능 평가를 위하여 반향제거 성능의 객관적 지표로 사용되는 ERLE(Echo Return Loss Enhancement)를 측정하였다. 그림 4의 신호에 대해 그림 5와 같은 ERLE 곡선을 얻을 수 있었다. 그림 5 (a)는 음향반향제거기만을 통과시킨 경우이며, 그림 5 (b)는 후처리단을 통과시킨 후의 결과이다. 결과로부터 음향반향제거기만 수행했을 때보다 결합시스템을 수행할 경우에 평균 10dB 이상의 성능향상을 얻을 수 있음을 알 수 있다.

### 4.3 동시통화 상황

그림 6은 동시통화 상황의 음성 데이터를 보여주고 있다. 그림 4 (b)의 측정된 마이크로폰 신호에 그림 6 (a)에서 동시통화로 표시된 구간에서 실측한 근단파자 신호를 첨가하여 동시통화상황을 발생시켰다. 그림 6 (b)는 결합시스템을 통과한 오차신호이다. 근단파자의 신호에 손상이 없이 반향신호만을 잘 제거한 것을 볼 수 있으며, 그림 7에 보인

ERLE 성능과 같이 동시통화 구간에서도 단일통화 상황과 마찬가지로 평균 10dB 정도 ERLE 성능향상을 얻을 수 있음을 알 수 있다.

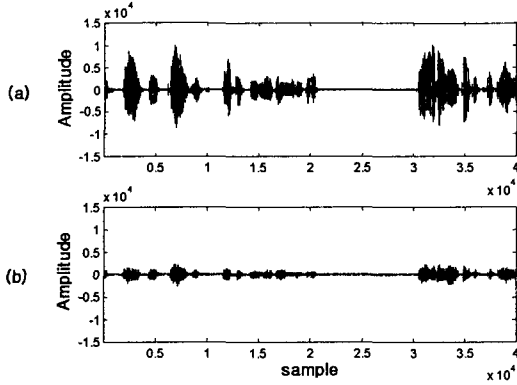


그림 4. (a) 원단화자신호, (b) 마이크론신호

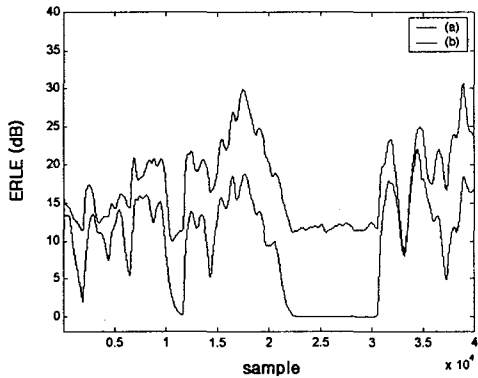


그림 5. ERLE 곡선 ( $N=256, \mu=0.3$ )  
(a) 음향반향제거기만 사용한 경우,  
(b) 결합 구조를 사용한 경우

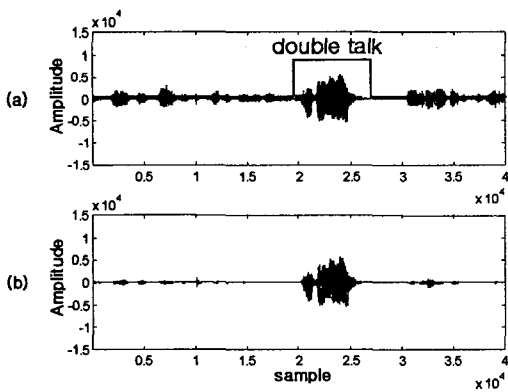


그림 6. (a) 마이크론신호, (b) 이중통화 구간 마이크론신호

(b) 결합시스템을 통과한 오차 신호

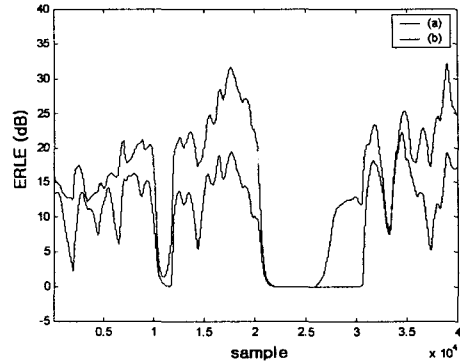


그림 7. ERLE 곡선 ( $N=256, \mu=0.3$ )  
(a) 음향반향제거기만 사용한 경우,  
(b) 결합 구조를 사용한 경우

## 5. 결론

본 논문에서는 차량환경에 적합한 음향반향제거기와 잡음제거기의 결합시스템을 제안하였고, 이를 16 비트 고정소수점 연산을 하는 OAK DSP 를 이용하여 실시간 구현하였다. 음향반향제거기에는 NLMS 알고리즘에 의해 구동되는 적응반향제거기법을 사용하였으며, 실측한 차내 충격응답을 기반으로 적응필터 차수는 256 으로 결정하였다. 배경잡음을 제거하기 위해서 IS-127 음성 부호화기의 표준안에 포함된 잡음제거방식을 사용하였다. 결합시스템을 이용하면 음향반향제거기만 사용하였을 때보다 평균 10dB 이상의 ERLE 성능 향상을 얻을 수 있었다.

## 참고문헌

- [1] S. Gustafsson, R. Martin, and P. Vary, "Combined acoustic echo control and noise reduction for hands-free telephony," *Signal Processing*, vol. 64, no. 1, pp. 21-32, 1998.
- [2] C. Beaugeant, V. Turbin, P. Scalart, and A. Gilloire, "New optimal filtering approaches for hands-free telecommunication terminals," *Signal Processing*, vol. 64, no. 1, pp. 33-47, 1998.
- [3] 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.
- [4] S. Haykin, *Adaptive Filter Theory*. 3rd Edition. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [5] TIA/EIA/IS-127, *Enhanced variable rate codec, speech service option 3 for wideband spread spectrum digital systems*, 1997.
- [6] H. V. Sorensen, D. L. Jones, M. T. Heideman and C. S. Burrus, "Real-valued fast fourier transform algorithms,"