

# 음질 향상을 위한 실시간 정합 필터 마이크로폰 어레이 시스템 구현

오 승 수\* · 김 기 만\*

## Implementation of a Real-time Matched Filter Microphone Array System for Speech Enhancement

Seung-Soo Oh\* · Ki-Man Kim\*

### Abstract

We have studied the speech enhancement method in the case of using hands free communication system in noisy environments as automobile. Matched filter microphone array for sound capture in reverberant rooms has been studied. However, the problems of the matched filter array are the system delay and heavy computational loads. This paper presents effective matched filter array method in environments. Simulations and real-time experimental results are included.

### I. 서론

최근 이동 통신의 급속한 발전에 따라 이동하는 차량내에서 통화하는 운전자가 크게 증가하는 실정이다. 그러나 운전중 이동 전화의 사용은 운전자의 주의력을 떨어뜨림으로써 사고의 위험성을 증가시키며, 실제로 이는 음주 운전 상태와 유사한 효과를 갖는 것으로 보고되고 있다. 이에 따라 몇몇 국가에서는 법으로 운전중 이동 전화의 사용을 제한하고 있다. 이러한 문제를 해결하기 위해 하나의 마이크로폰을 핸들이나 운전자 앞 부분에 설치하고 있으나 이 경우 마이크로폰으로 입력되는 신호는 화자의 음성뿐만 아니라 주변 소음이 함께 수신되고 스피커 폰을 사용할 경우 음향 반향 문제

\* 한국해양대학교 공과대학 전파공학과

도 발생할 수 있다. 따라서 핸드즈 프리(hands-free) 통신을 위한 음질 향상 시스템의 구현이 요구되고 있다.

그 동안 실내에서의 음질 향상을 위한 많은 연구가 이루어져 왔다. 그중 대표적인 것이 적응 잡음 제거 기법, 능동 소음 제어 기법, 그리고 빔 형성 기법이다. 적응 잡음 제거 기법은 적응 필터를 이용하여 주 채널에 입력되는 소음 신호를 보조 센서로 수신되는 소음 신호 사이의 상관도를 이용하여 제거한다. 적응 필터 구조로서 transversal, 격자형, Gram-Schmidt 및 혼합형 구조등이 있으며, 적응 알고리즘으로 LMS(Least Mean Square), RLS(Recursive Least Square) 방법등이 연구되었다. 이 방법의 성능은 주 채널로 입력되는 소음과 상관성이 높은 소음 신호를 보조 센서를 통해 획득하는데 있으며, 이는 다양한 차량 소음 환경을 고려할 때 많은 문제점을 갖고 있다. 능동 소음 제어 기법은 소음과 같은 크기를 가지면서 역위상을 갖는 인공 소음을 만들어 스피커를 통해 방사함으로써 서로 상쇄 효과를 얻는 기법이다. 그러나 이 방법은 엔진 소음과 같이 거의 단일 주파수 성분을 가지면서 저주파 대역 신호인 경우에만 적합하고 타이어 잡음과 같이 고주파 신호 성분이나 광대역 잡음인 경우 제거가 어렵다. 빔 형성 기법은 다수의 마이크로폰 입력 신호들을 조합하여 출력단에서 최대 신호 대 잡음비를 갖도록 하는 것이다. 그러나 이 방법은 마이크로폰의 수가 적은 경우 빔 폭이 커져서 원하지 않은 외부 소음도 수신할 수 있다.

본 논문에서는 소음 환경에서 무선 통신 시스템을 사용하는 경우 음질 향상을 위한 방법을 연구하였다. 특히 정합 필터 마이크로폰 어레이를 이용하여 화자의 위치를 추정하고, 마이크로폰으로 수신된 음 가운데 화자의 음성만을 추출한다. 이 경우 정합 필터를 계산하는 동안 실내 조건에 따라 계산량이 크게 증가할 수 있다. 이는 임펄스 응답의 길이가 증가함으로써 발생되며, 실시간 처리 시스템의 구현을 어렵게 한다. 이에 실시간 처리를 고려하여 적은 계산량을 갖는 모델을 제시한다. 제안된 방법은 컴퓨터 시뮬레이션과 함께 범용 디지털 신호 처리용 프로세서를 이용한 실시간 처리 시스템을 구현하고 다양한 환경에서 실험을 수행함으로써 성능을 증명하였다.

## II. 화자 위치 추정 및 정합 필터 어레이 시스템

마이크로폰 어레이를 이용한 화자 위치 추정 기법으로 입사각 추정 기법과 시간 지연 추정 기법이 있다. 입사각 추정 기법은 주로 통신 시스템에서 적용되는 방법으로 수신된 신호 벡터와 가정된 신호 벡터 사이의 상관도를 측정함으로써 신호의 입사 방향을 추정한다. 이에는 Bartlett 방법, 최소 분산 방법, MUSIC 방법등이 있으나 그 계산량이 많다는 것과 송신된 신호는 단일 경로로 전달되어야 하는 문제점이 있다.

시간 지연 추정 기법은 각 마이크론에 수신된 신호 사이의 상관 계수를 구하여 위치를 추정하는 방법으로 CPSP와 같은 파라미터로 시간 지연을 추정하고 이의 여러 쌍을 사용하여 화자의 위치를 추정한다. 시간 지연 추정 기법의 문제점은 잔향 성분이 큰 실내 환경에서는 성능이 크게 저하된다는 것이다.

정합 필터 어레이는 원래 영상 회의와 같은 일반 실내에서 잔향의 영향을 감소시키기 위한 방법으로 연구되었다. 그림 1과 같이 화자와  $i$ 번째 마이크론 사이의 임펄스 응답을  $h_i(t)$ 라 할 때 원래의 음성 신호  $s(t)$ 는  $h_i(t)$ 와 컨볼루션 되어 마이크론에 입력된다. 마이크론에 입력된 신호는  $h_i(t)$ 의 역필터  $h_i(-t)$ 를 통과하여 원래 음성 신호로 재현된다. 만약 마이크론의 수가  $N$ 개라면 정합 필터 어레이의 출력  $y(t)$ 는 다음과 같다.

$$y(t) = s(t) * \sum_{i=1}^N h_i(t) * h_i(-t) \quad (1)$$

만약  $N$ 이 충분히 크다면 임펄스 응답과 그 역필터 사이의 컨볼루션은 임펄스 형태가 되며 결국 정합 필터 어레이 출력은 원래의 음성 신호와 같아진다. 이러한 정합 필터 어레이는 빔 형성 기법에 비해 잔향에 강한 장점을 갖고 있으나 임펄스 응답의 길이가 긴 경우 역필터의 길이가 길어져 계산량이 크게 증가되는 단점을 갖고 있다.

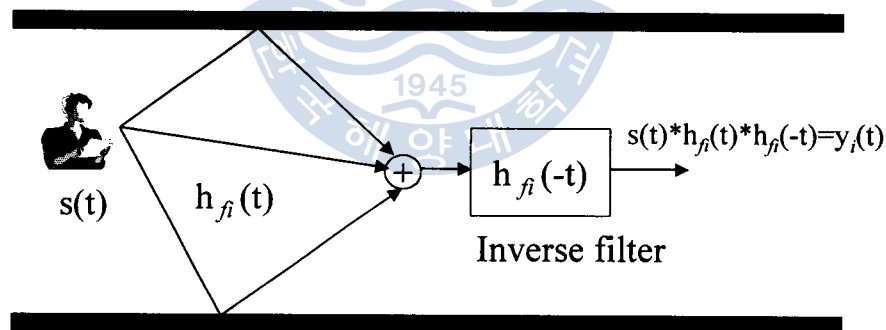


그림 1. 정합 필터 마이크론 어레이의 개념도.

### III. 실시간 처리를 위한 정합 필터 마이크론 어레이 및 구현

본 논문에서는 범용 디지털 신호 처리용 프로세서를 이용하여 정합 필터 어레이를

를 위하여 Texas Instrument사의 부동 소수점 연산 프로세서인 TMS320C31를 사용하였으며 사용된 프로세서는 50MHz의 클럭으로 동작한다. 또한 프로세서내에 내장된 메모리가 2Kwords 밖에 되지 않기 때문에 128Kwords의 외부 메모리를 추가하였으며, 추가된 메모리로써 15ns의 access time을 갖는 고속 SRAM을 사용하여 zero-wait로 동작하도록 하였다. 어드레스 디코딩을 위한 주변 회로들은 모두 7ns 지연 시간을 갖는 고속 GAL로 대체되었다. 프로그램의 최적화를 위하여 모든 프로그램은 어셈블러

구현할 목적으로 실시간 처리에 적합한 축소 모델을 연구하였다. 아울러 이 모델을 이용하여 화자의 위치도 추정하도록 하였다.

일반적으로 역필터는 임펄스 응답을 주파수 영역으로 변환하여 각 주파수 성분의 역수를 계산한 뒤 다시 역변환하여 구한다. 이렇게 하여 구해진 역필터는 FIR 형태를

로 직접 제작되었다.

#### IV. 실험 및 결과 고찰

실험은 실내에서 마이크로폰 어레이를 사용하여 음질을 향상시켰다. 실험에 사용된 마이크로폰은 오디오 테크니카사의 AT9500 모델로 핀 형태로써 무지향성을 가지며, 8kHz의 샘플링 주파수를 갖도록 하였다. 실험은 FIR과 제안된 모델에 따른 성능, 잡음 환경에서 성능을 고찰하였다. 정합 필터를 구성하기 위해 각 마이크로폰의 위치에서 임펄스 응답을 측정하였으며, 이때 의사 잡음과 시스템 모델링 기법을 이용하여 구하였다.

그림 2에 FIR 모델로 처리된 결과와 제안된 모델로 처리된 결과를 나타내었다. 그림에서 두 결과는 파형상에서 차이가 없음을 알 수 있다. 그러나 계산 과정에서 제안된 방법은 FIR 모델에 비해 약 1/4 가량의 계산량을 갖는다. 그림 3에는 잡음 환경에서 처리된 결과를 보여준다. 결과는 마이크로폰 수가 증가할수록 원음에 가까워진다는 것을 알 수 있다.

#### V. 결론

본 논문에서는 실내에서 음질을 향상하기 위한 시스템을 구현하였다. 구현된 방법은 정합 필터 마이크로폰 어레이를 이용하였으며, 실시간 처리에 적합하도록 계산량을 감소시킨 방법을 제안하고 적용하였다. 계산량 감소를 위해 Balanced model reduction 방법을 이용하여 pole-zero 모델로 변형하였다. 그 결과 약 1/4의 계산량만으로도 만족할 결과를 얻었다. 또한 범용 디지털 신호 처리용 프로세서를 사용한 시스템을 구현하고 실시간 처리 실험들을 다양한 환경하에서 수행하였다. 앞으로 본 연구를 발전시키기 위해 화자의 머리 움직임등을 보상할 수 있는 기법들에 대해 연구가 이루어져야 할 것이다.

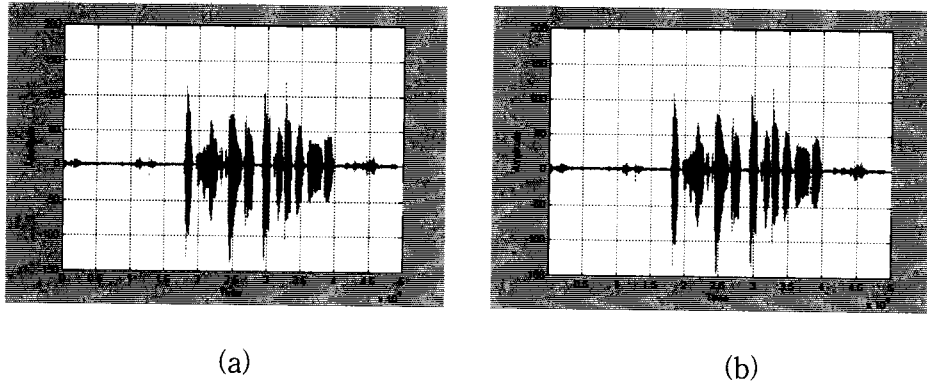


그림 2. (a) FIR 모델 처리 결과, (b) 제안된 모델 처리 결과 파형.

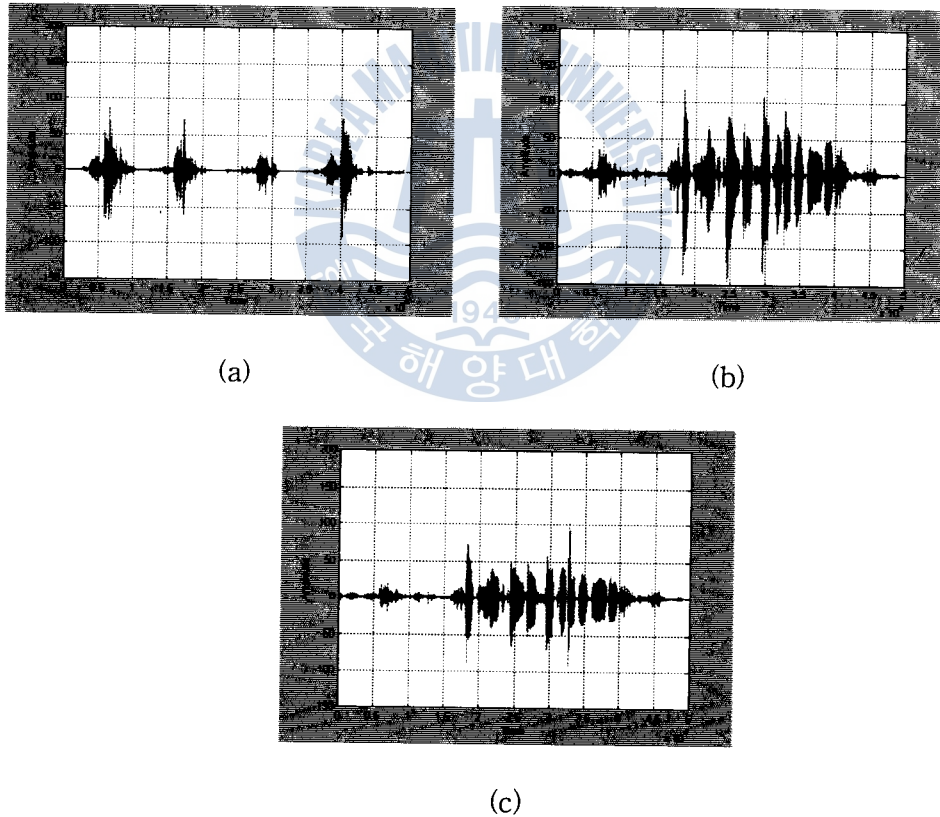


그림 3. (a) 잡음 신호, (b) 한 개의 정합 필터 마이크론 출력, (c) 8개의 정합 필터 마이크론 어레이 출력.

## 참 고 문 헌

- [1] R.J. Renomeron, D.V. Rabinkin, J.C. French and J.L. Flanagan, "Small-scale matched filter array processing for spatially selective sound capture," *134th Meeting of the Acoustical Society of America*, Dec. 1997.
- [2] E.E. Jan and J. Flanagan, "Sound capture from spatial volumes: matched filter processing of microphone arrays having randomly-distributed sensors," *Proc ICASSP*, pp.917-920, 1996.
- [3] P. Wauters, K. Eneman, K. Delaet and R. Lauwereins, "Adaptive speech beamforming using the TMS320C40 multi-DSP," Texas Instrument Technical Note SPRA305, Dec. 1995.
- [4] Y. Kaneda and J. Ohga, "Adaptive microphone array system for noise reduction," *IEEE Trans. Acoust. Speech Signal Proc.*, vol.34, no.6, pp.1391-1400, Dec. 1986.
- [5] Y. Grenier, "A microphone array for car environments," *Speech Communication*, vol.12, no.1, pp.25-39, 1993.
- [6] S. Nordholm, I. Claesson, B. Bengtsson, and P. Eriksson, "A multi-DSP implementation of a broadband adaptive beamformer for use in a handsfree mobile radio telephone," *IEEE Trans. Vehicular Technology*, vol.40, no.1, Feb. 1991.
- [7] B. Beliczynski, I. Kale and G.D. Cain, "Approximation of FIR by IIR digital filters: an algorithm based on balanced model reduction," *IEEE Trans. Signal Proc.*, vol.40, no.3, pp.532-541, March 1992.

## 요 약

자동차 실내와 같은 소음 환경에서 무선 통신 시스템을 사용하는 경우 음질 향상을 위한 방법을 연구하였다. 그 동안 많은 연구가 이루어졌던 적응 잡음 제거, 능동 소음 제어, 빔 형성 기법과는 달리 정합 필터 마이크로폰 어레이를 이용하여 화자의 위치를 추정하고, 마이크로폰으로 수신된 음 가운데 화자의 음성만을 추출한다. 또한 실시간 처리를 고려하여 적은 계산량을 갖는 모델을 제시한다. 제안된 방법은 컴퓨터 시뮬레이션과 함께 범용 디지털 신호 처리용 프로세서를 이용한 실시간 처리 시스템을 구현하고 다양한 환경에서 실험을 수행함으로써 성능을 증명하였다.