

# Matched Filter Array를 이용한 음질 향상 시스템 구현

오승수 · 김기만  
한국해양대학교 전파공학과

## Implementation of Speech Enhancement System using Matched Filter Array

Seung-soo Oh · Ki-man Kim  
Dept. of Radio Science & Eng., Korea Maritime University  
E-mail : sunnysid@chollian.net

### 요 약

최근 화상 회의 시스템에서 화자 위치 추정 및 음질 향상 기술이 연구되고 있다. 이 시스템에서는 마이크로폰 어레이를 이용, 화자의 위치를 파악하여 화자의 방향으로 카메라를 자동으로 조정해 주게 된다. 본 연구에서는 마이크로폰 어레이를 통해 수신된 신호를 이용하여 Matched Filter Array를 구성하고 음질을 향상시켰다. 이때 역변환 필터로써 IIR필터를 사용하여 계산량을 줄였으며, 범용 DSP 프로세서를 이용한 하드웨어를 제작하여 그 성능을 확인하였다.

### ABSTRACT

Recently, speaker localizing estimation technique has been rising in teleconference systems. In this system, it is recognized speaker location using microphone array and camera is directed to speaker location automatically. In this paper, it was described to be able to enhance the speech quality through microphone array, decrease computational loads using IIR filter as inverse filter, and confirmed to implement hardware using DSP processor.

### 1. 서 론

최근 화상 회의 시스템에서 화자 위치 추정과 음질 향상 기술이 부각되고 있다. 이 시스템에서는 마이크로폰 어레이를 이용, 화자의 위치를 파악하여 화자의 방향으로 카메라를 자동으로 조정하면서 음원을 선택하여 음질을 향상시킨다. 음질 향상 위한 기존의 방법으로써 적응 잡음 제거 기법, 능동 소음 제어 기법, 그리고 빔 형성 기법 등이 있다. 적응 잡음 제거 기법은 적응 필터를 이용하여 주 채널에 입력되는 소음 신호를 보조 센서로 수신되는 소음 신호 사이의 상관도를 이용하여 제거한다. 적응 필터 구조로써 transversal, 격자형, Gram-Schmidt 및 혼합형 구조 등이 있으며, 적응 알고리즘으로 LMS(Least Mean Square), RLS(Recursive Least Square) 방법 등이 연구되었다. 이 방법의 성능은 주 채널로 입력되

는 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 거의 단일 주파수 성분을 가지면서 저주파 대역 신호인 경우에만 적합하고 고주파 신호 성분이나 광대역 잡음인 경우 제어가 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 정합 필터 마이크로폰 어레이를 이용하여 음질 향상을 위한 방법을 연구하였다. 정합 필터를 계산하는 동안 실내 조건에 따라 계산량이 크게 증가할 수 있으며, 이는 임펄스 응답

의 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 실험을 수행함으로써 성능을 증명하였다.

### II. Matched Filter Array 시스템

정합 필터 어레이는 원래 영상 회외와 같은 일반 실내에서 잔향의 영향을 감소시키기 위한 방법으로 연구되었다. 그림 1과 같이 화자와 i번째 마이크로폰 사이의 임펄스 응답을  $h_i(t)$ 라 할 때 원래의 음성 신호  $s(t)$ 는  $h_i(t)$ 와 컨볼루션 되어 마이크로폰에 입력된다. 마이크로폰에 입력된 신호는  $h_i(t)$ 의 역필터  $h_i(-t)$ 를 통과하여 원래 음성 신호로 재현된다. 만약 마이크로폰의 수가 N개 라면 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 N이 충분히 크다면 임펄스 응답과 그 역필터 사이의 컨볼루션은 임펄스 형태가 되며 결국 정합 필터 어레이 출력은 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.

### III. 실시간 처리를 위한 Matched Filter Array

본 논문에서는 범용 디지털 신호 처리용 프로세서를 이용하여 정합 필터 어레이를 구현할 목적으로 실시간 처리에 적합한 축소 모델을 연구하였다.

일반적으로 역필터는 임펄스 응답을 주파수 영역으로 변환하여 각 주파수 성분의 역수를 계산한 뒤 다시 역변환하여 구한다. 이렇게 하여 구해진 역필터는 FIR 형태를 갖는다. 그러나 잔향 시간이 길거나 실시간 처리를 원활히 하고자 한다면 보다 짧은 탭 수로 줄여야 한다. 이를 위한 대표적인 방법으로 resampling 방법이 있으나 이 방법은 샘플링 주파수가 변하는 문제점이 있다. 본 논문에서는 Pole-Zero 모델, 즉 IIR 형태로 변환한다. FIR로부터 IIR로 변환에는 다양한 방법이 있으나 효율적인 방법으로 다음과 같은 Balanced Model Reduction 방법을 사용토록 한다. 먼저 계산된 역필터의 계수를 M개라 한다면 그의 z 변환은 다음과 같다.

$$F(z) = f_0 + f_1 z^{-1} + \dots + f_{M-1} z^{-(M-1)} \quad (2)$$

역필터  $F(z)$ 은 다음과 같은 차분 방정식으로 표현할 수 있다.

$$x(n+1) = Ax(n) + Bu(n) \quad (3)$$

$$y(n) = Cx(n) + Du(n) \quad (4)$$

여기서 A, B, C, D는 다음과 같다.

$$A = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$C = [f_1 \ f_2 \ \dots \ f_{M-1}] \quad D = f_0 \quad (5)$$

이제 Hankel 행렬 H를 구성한 뒤 Singular Value Decomposition을 수행하면 다음과 같다.

$$H = V \Lambda V^T \quad (6)$$

여기서  $\Lambda$ 내의 값은 내림차순으로 정렬되어 있다고 가정하면 이 Singular Value 분포에 의해 적절한 차수를 결정한다. 결정된 차수를 k라 하면 축소 모델을 위한 State-space 표현은 다음과 같다.

$$A_k = V_k^T A V_k \quad (7)$$

$$B_k = V_k^T B \quad (8)$$

$$C_k = C V_k \quad (9)$$

여기서  $V_k$ 는 다음을 만족하는  $M \times k$  행렬이다.

$$V = [V_k \ V_{M-k}] \quad (10)$$

마지막으로  $A_k, B_k, C_k, D$ 를 이용하여 저차 역필터의 계수들을 계산하며 이는 항상 IIR 형태가 된다.

### IV. 실시간 처리 시스템 구현

제안된 방법의 성능을 입증하기 위하여 범용 디지털 신호 처리 프로세서를 이용한 실시간 처리 시스템을 구현하였다. 구현된 하드웨어는 아날로그 부분과 디지털 부분으로 구성되었다. 그림 2에 제작된 시스템의 구성을 나타내었다.

(1) 아날로그 부분

아날로그 부분은 크게 증폭기, 고역 통과 필터, 저역 통과 필터로 구성된다. 증폭기는 OP AMP를 사용하여 2단으로 구현하였으며, 가변 증폭할 수 있도록 하였다. 고역 통과 필터는 직류 바이어스 및 전원 잡음을 제거하기 위해 350Hz에서 차단 주파수를 가지며 4차 Chebyshev 필터를 구현하였다. 또한 저역 통과 필터는 A/D 변환시 에러이러징 현상을 방지하기 위해 부차되었다. 음성의 주파수 대역을 고려하여 3400Hz에서 차단 주파수를 갖는 4차 Chebyshev 필터이다.

## (2) 디지털 부분

디지털 부분은 A/D 및 D/A 변환부, 디지털 신호 처리부 및 기타 주변 회로로 구성된다. A/D 변환기는 최대 8채널 입력, 12-bit 분해능을 갖고 최대 100kHz의 샘플링 속도를 갖는 칩을 사용하였다. D/A 변환기 역시 12-bit를 갖도록 하였다. 신호 처리를 위하여 Texas Instrument사의 부동 소수점 연산 프로세서인 TMS320C31를 사용하였으며 사용된 프로세서는 50MHz의 클럭으로 동작한다. 또한 프로세서 내에 내장된 메모리가 2Kwords 밖에 되지 않기 때문에 128Kwords의 외부 메모리를 추가하였으며, 추가된 메모리로서 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 어셈블러로 직접 작성되었다.

## V. 실험 및 결과 고찰

실험에 사용된 마이크로폰은 오디오 테크니카사의 AT9500 모델로 핀 형태로써 무지향성을 갖고 그 수는 4개이며, 8kHz의 샘플링 주파수를 갖도록 하였다. 정합 필터를 구성하기 위해 각 마이크로폰의 위치에서 임펄스 응답을 측정하였으며, 이때 의사 잡음을 이용하였다. 그림 3에는 원음과 마이크로폰으로 수신된 신호의 파형을 나타내었다. 원음은 음의 전달과정에서 왜곡되어 수신되고 있음을 볼 수 있다. 그림 4에는 기존의 FIR 모델로 Matched Filter Array를 구성한 경우와 본 논문에서 제안된 IIR 모델로 처리한 경우를 보여준다. 두 그림으로부터 제안된 방법을 사용하더라도 성능에는 큰 차이가 없음을 알 수 있고, 이에 반해 제안된 방법은 약 1/4의 계산량으로 처리되었다.

## V. 결론

본 논문에서는 정합 필터 마이크로폰 어레이를 이용한 음질 향상 시스템을 구현하였으며, 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 움직임 등을 보상할 수 있는 기법들에 대해 연구가 이루어져야 할 것이다.

## 참고문헌

- [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, "Aadaptive 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.

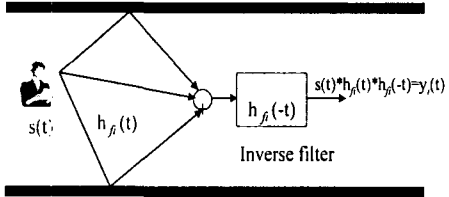
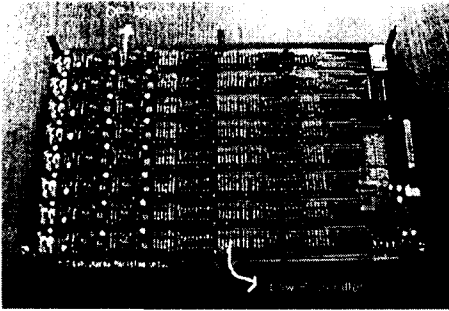
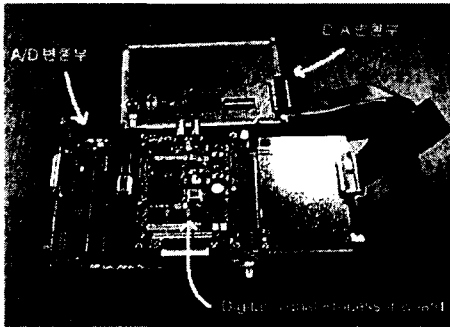


그림 1. Matched Filter Array의 개념도

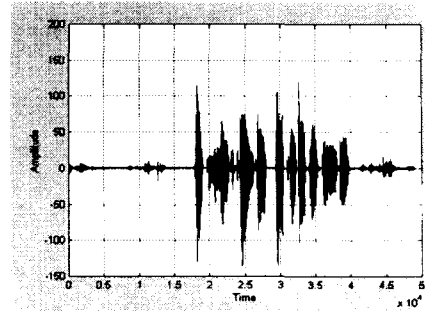


(a) 아날로그 부분



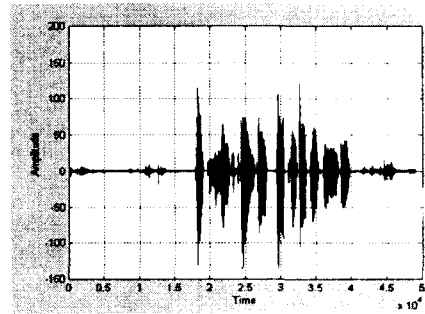
(b) 디지털 부분

그림 2. 제작된 하드웨어 구성도

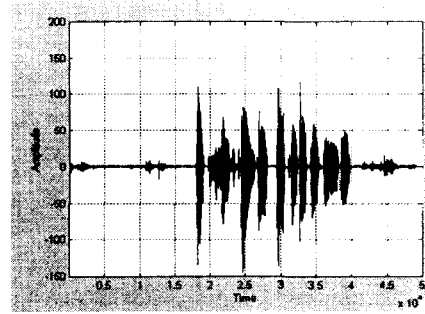


(b)

그림 3. (a) 원음의 파형, (b) 수신된 음의 파형

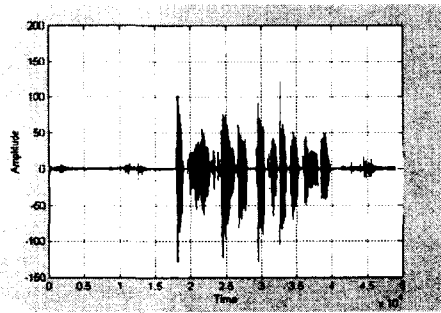


(a)



(b)

그림 4. (a) FIR 모델로 처리한 결과, (b) 제안된 모델로 처리한 결과



(a)