

# 방향성 마이크로폰과 음성 필터링을 이용한 통신 시스템의 음성 인지도 향상

## Performance Enhancement of Speech Intelligibility in Communication System Using Combined Beamforming (directional microphone) and Speech Filtering Method

신민철\*·왕세명\*\*

Mincheol Shin and Semyung Wang

**Key Words** : Directional Microphone(방향성 마이크로폰), Speech Filtering Method (음성 필터링 기술), Speech Intelligibility (음성 명료도)

### ABSTRACT

The speech intelligibility is one of the most important factors in communication system. The speech intelligibility is related with speech to noise ratio. To enhance the speech to noise ratio, background noise reduction techniques are being developed. As a part of solution to noise reduction, this paper introduces directional microphone using beamforming method and speech filtering method. The directional microphone narrows the spatial range of processing signal into the direction of the target speech signal. The noise signal located in the same direction with speech still remains in the processing signal. To sort this mixed signal into speech and noise, as a following step, a speech-filtering method is applied to pick up only the speech signal from the processed signal. The speech filtering method is based on the characteristics of speech signal itself. The combined directional microphone and speech filtering method gives enhanced performance to speech intelligibility in communication system.

### 1. 서론

음성 통신 시스템에 있어서 음성 명료도는 가장 중요한 요소 중 하나이다. 음성 명료도에 큰 영향을 미치는 음성 신호 대 잡음비를 높이기 위해 많은 잡음 제거 기술들이 개발되었다. 잡음 제거 기술 중 자주 사용되고 있는 기법은 방향성 마이크로폰을 이용한 방법이다. 이 방법은 목표 음향 신호가 고정된 방향에서 발생한다고 가정하고 방향성이 없는 마이크로폰의 배열과 신호 처리를 통해 목표로 하는 방향의 소리에 민감하도록 설계한 방법이다. 이 방법은 목표 음원 방향 이외의 방향에서 발생하는 음향 신호를 모두 잡음으로 가정하면 뛰어난 음성 신호 대 잡음비를 얻을 수 있다. 하지만 목표 음원 방향에 잡음이 섞여 있는 경우 좋은 성능을 기대할 수 없다. 본 논문에서는 콤 필터(comb filter)를 이용한 음성 필터링 방법을 방향성 마이크로폰 기술에 결합함으로써 목표 음원이 음성 신호일 경우 음성 신호의 명료도를 높이는 방법을 제안하였다. 음성 필터링 방법은 음성 신호가 기본 주파수 (fundamental frequency)의 정수배가 되는 주파수 성분으로 이루어져 있다는

음성 신호의 특성을 이용하였다. 이 방법의 기본 원리는 앞에서 설명한 방향성 마이크로폰을 이용하여 목적 음성 신호 방향 이외의 잡음을 제거한 뒤 처리된 신호 안의 잡음들을 음성 필터링 방법으로 제거하는 것이다. 결과적으로 본 논문에서 제안된 기법(Combined Beamforming and Speech filtering method, CBS method)의 목적은 앞에서 언급된 두 가지의 잡음 제거 기술을 결합하여 좀 더 나은 잡음 제거 성능을 얻고 음성 명료도를 높이는 것이다. 2 장에서 방향성 마이크로폰을 이용한 잡음 제거 방법에 대해 기술하였고 음성 필터링 방법에 대한 계략적인 설명이 3 장에 설명되었다. 4 장에서는 방향성 마이크로폰과 음성 필터링을 함께 적용한 CBS 기법이 제안되었다. 컴퓨터 시뮬레이션과 CBS 기법에 의해 처리된 결과 신호에 대한 주관적 평가는 5 장과 6 장에 언급하였다.

### 2. 방향성 마이크로폰 잡음 제거 기법

방향성 마이크로폰 잡음 제거 기술은 원래 안테나 분야에서 쓰이던 빔포밍 (beamforming) 기법을 음향 신호에 적용한 것으로 같은 성능의 마이크로폰을 어레이(array)로 만들어 각각의 마이크로폰에 입력되는 음향 신호의 시간차와 위상차를 이용하여 원하는 방향에 대한 마이크로폰의 민감도를 높이는 방법이다.

\* 광주과학기술원 기전공학과  
E-mail : mcshin@gist.ac.kr  
Tel : (062) 970-2429, Fax : (062) 970-2384

\*\* 광주과학기술원 기전공학과

다음 그림 1 은 2 개의 마이크로폰을 사용하여 방향성을 구현하는 1 차 방향성 마이크로폰 모델을 나타낸 것이다.

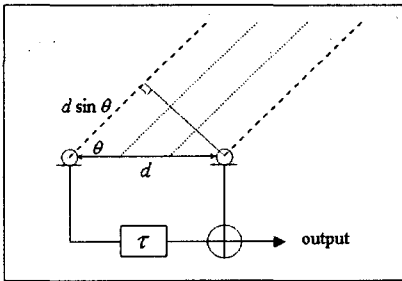


Figure 1 – First order directional microphone model

음향 신호가 평면파이며 이것이 입사각,  $\theta$  로 각각의 마이크로폰에 입력된다는 가정 하에 그림 1 의 모델의 방향성 패턴을 식 (1)과 같이 나타낼 수 있다 [1, 2].

$$E(\omega, \theta) = 1 - e^{-j\omega\left(\frac{d}{c}\cos\theta + \tau\right)} \quad (1)$$

여기서  $\tau$  는 시간 지연,  $d$  는 마이크로폰 사이의 거리 그리고  $\omega$  는 각주파수를 의미한다.  $c$  는 소리의 속도로 340 m/s 이다. 주파수에 따른 방향성의 패턴은 식 (1)에서 시간 지연과 마이크로폰 사이의 거리를 조절함으로써 변화시킬 수 있다.

### 3. 음성 필터링 기법

음성 필터링 기법은 음성 신호가 기본 주파수의 정수배가 되는 주파수 성분들로 이루어져 있는 특성을 이용한다. 아래의 그림 2 는 음성 필터링 방법을 개략적으로 설명하기 위한 그림이다[3, 4].

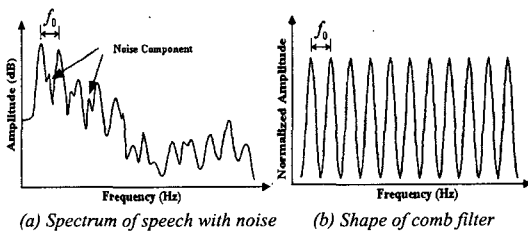


Figure 2 – Brief description of speech filtering method using comb filter ( $f_0$ : fundamental frequency)

그림 2 (a)는 음성과 잡음이 섞인 신호의 스펙트럼을 나타낸다. 이와 같이 음성 신호는 기본 주파

수의 정수배가 되는 주파수 성분들로 이루어져 있고 잡음 성분(noise component)은 원래의 신호를 왜곡 시킨다. 그림 2(b)는 (a)에서 나타낸 음성 신호의 기본 주파수를 밴드 간격으로 하는 콤 필터를 디자인한 것이다. 음성 필터링 방법은 (a)와 같은 잡음이 첨가된 음성 신호에서 음성 신호의 기본 주파수를 검출하여 (b)와 같이 디자인된 콤 필터를 (a)의 신호에 적용하여 음성 신호를 분리하는 방법이다.

이 방법은 목표가 되는 음성 신호의 정확한 기본 주파수 검출과 콤 필터로 인해 손상된 원래 신호의 음질을 향상 시키는 방법이 완전하지 않아 현재 개발 중에 있다.

### 4. CBS 잡음 제거 기법

CBS 잡음 제거 기법은 2 장과 3 장에서 설명한 방향성 마이크로폰 기법과 음성 필터링 기법을 결합한 방법으로 이의 블록다이어그램은 그림 3 과 같다.

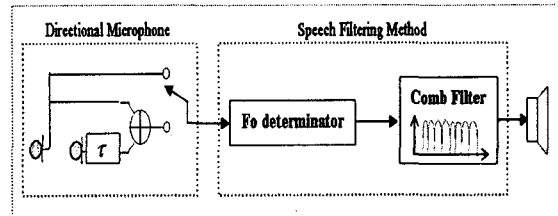


Figure 3 – Block diagram of Combined Beamforming and Speech filtering(CBS) method

그림 3 에 나타낸 CBS 잡음 제거 기법은 앞에서 언급된 두 가지 잡음 제거 기술이 병렬로 사용되었다. 왼쪽 박스에 표시된 방향성 마이크로폰은 목표 신호가 존재하는 방향의 신호들을 민감하게 받아들이고 다른 방향의 잡음들을 감소시킨다. 방향성 마이크로폰에 입력된 목표 음성과 비슷한 위치에 있는 잡음들은 오른쪽 박스에 표시된 음성 필터링을 통해 제거된다. 여기서 콤 필터는 목표 음성 신호의 기본 주파수를 검출한 뒤 이를 밴드 간격으로 하여 디자인된다.

### 5. 컴퓨터 시뮬레이션

본 5 장에서는 컴퓨터 시뮬레이션을 통해 CBS 잡음 제거 기법의 성능을 평가하였다. 시뮬레이션 환경과 입력 음향 신호들을 나타내었고 시뮬레이션에 사용된 CBS 설계에 대해 소개하였다.

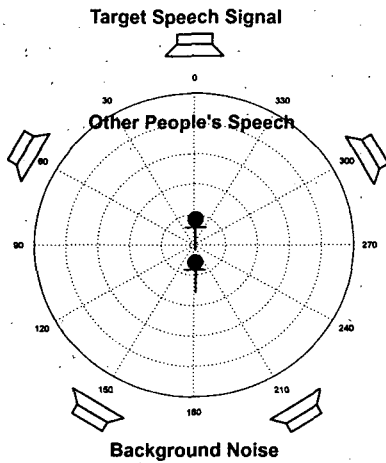


Figure 4 – Simulation setting for the evaluation of the CBS method

그림 4 에서 나타낸 것과 같이 목표 음성 신호를 정면(0 도), 여러 사람들의 목소리가 음성거리는 신호를 60, 300 도, 그리고 주변 잡음 신호를 150, 210 도에 위치 시켰다. 이와 같은 시뮬레이션 환경은 뒤 쪽에 주변 잡음이 있고 듣고자 하는 음성 주위에 다른 사람들의 음성이 있는 환경을 구성한 것이다. 이 때 방향성 마이크로폰 시스템을 구성하고 있는 두 개의 마이크로폰 사이의 거리를 0.01m 로 한다.

그림 5 는 본 시뮬레이션에 사용된 목표 음성 신호와 잡음 신호들을 나타내었다. 목표 음성 신호는 남성의 “아[a]” 발음을 선택하였고 이의 시간 영역 신호와 주파수 영역의 스펙트럼을 각각 그림 5 의 (a)와 (b)에 나타내었다. (c)는 여러 사람들의 음성거리는 소리이며 (d)는 배경 잡음이다.

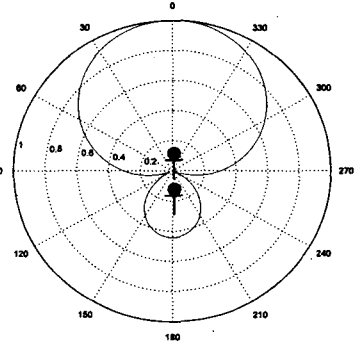
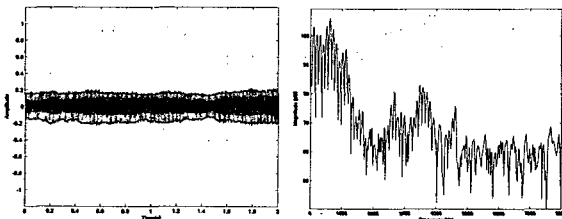
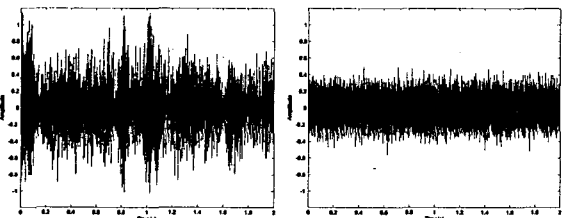


Figure 6 – First order directional polar response

그림 6 은 그림 3 의 왼쪽 박스에 표시된 1 차 방향성 마이크로폰 시스템의 방향성 패턴(민감도)이다. 이것은 마이크로폰 사이의 거리를 0.01m 로 고정한 상태에서 0Hz 에서 8000Hz 까지 500Hz 간격으로 구한 방향성 패턴을 평균한 것이다. 일반적으로 1 차 방향성 마이크로폰 시스템에서 마이크로폰 사이의 거리가 가까우면 전 주파수 영역의 방향성 패턴이 거의 일정하므로 이와 같이 평균된 방향성 패턴을 전 주파수 대역의 패턴이라고 가정하고 모의 실험을 진행하였다.



(a) Target speech signal (b) Spectrum of target speech signal



(c) Other people's speech signal (d) Background noise signal

Figure 5 – Target speech signal and noises used in simulation setting

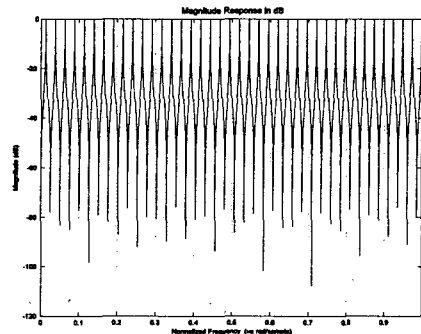


Figure 7 – Comb filter shape

그림 7 은 그림 3 의 오른쪽 박스로 표현된 음성 필터링 방법에 사용된 콤 필터를 나타낸 것이다. 이 필터는 밴드폭이 2.7Hz 이며 각각의 밴드 간격이 101.18Hz 로 그림 5 의 (b)에 나타난 목표 음성 신호의 기본 주파수와 동일하다. 이때 목표

음성 신호의 기본 주파수는 캡스트럼(Cepstrum) 방법을 이용하여 검출하였다 [5].

## 6. 결론

5 장에 설명된 시뮬레이션을 통해 만들어진 결과 신호들을 평가하기 위해 주관적 평가 방법이 사용되었다. 잡음의 양(noise quantity), 목표 음성의 명료도(speech intelligibility) 그리고 목표 음성의 음질(target quality), 이 세가지 항목을 평가하였다.

	Noise quantity	Speech intelligibility	Target Quality
5	Very much	Very good	Very good
4	Much	Good	Good
3	Middle	Middle	Middle
2	A little	Poor	Poor
1	none	Very poor	Very poor

Table 1 - Scoring table

표 1 에서 나타낸 것과 같이 각각의 평가 항목에 1~5 까지의 점수를 배정하고 정상적인 사람 8 명을 대상으로 각 개인이 느끼는 잡음의 양, 음성 명료도, 음질에 배점을 하도록 하였다. 본 주관적 평가는 아무 처리가 되지 않은 신호, 방향성 마이크로폰만을 거친 신호, 음성 필터만을 거친 신호, 그리고 CBS 기법을 사용한 신호를 가지고 평가된 결과를 비교하기 쉽게 그림 8 과 같은 그래프로 나타내었다.

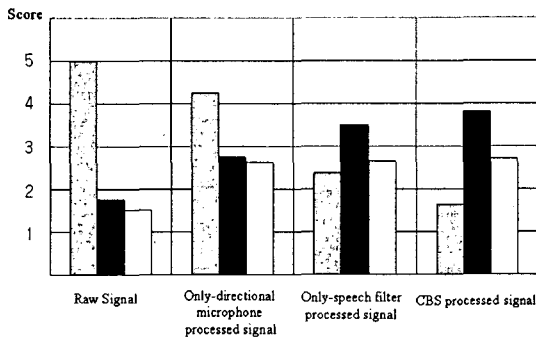


Figure 8 - Results of subjective test with four signals (Gray bar: noise quantity, black bar: speech intelligibility, white bar: quality)

그림 8 은 주관적 평가의 결과를 나타내는 그래프이다. 이 평가에서 쓰인 평가 항목인 잡음의 양은 회색 막대로, 음성 명료도는 검정 막대로, 그리고 음성 신호의 음질은 흰색 막대로 표시하였다.

그림 8 에 나타난 바와 같이, 잡음의 양은 방향성 마이크로폰만을 적용한 신호보다 음성 필터만을 거친 신호에서 더 적었고 CBS 방법을 적용한 신호에서 가장 적었다. 이는 CBS 방법이 잡음 제거에 있어서 가장 좋은 성능을 가지고 있다는 것을 보여준다. 음성 명료도에 있어서도 CBS 방법을 적용한 신호가 가장 좋은 점수를 얻었다. 하지만 음질에 있어서 그리 좋은 성능을 얻지 못하였는데 이는 앞 장에서 언급한 것과 같이 음성 필터링 방법을 적용할 때에 목표가 되는 음성 신호의 기본 주파수 검출이 정확도가 떨어지고 필터링 후 손상된 음질을 보충하는 과정이 생략되어 있기 때문이다. 이를 개선하기 위해서 향상된 성능의 실시간 음성 기본 주파수 검출기와 콤 필터를 통과한 음성 신호의 보상 방법의 개발이 필요하다.

결론적으로 본 논문에서 제안된 CBS 방법을 이용하여 방향성 마이크로폰과 음성 필터링 방법만을 사용하였을 때보다 잡음 제거와 음성 명료도 측면에서 더 향상된 성능을 얻을 수 있었다.

이로써 각각의 기술의 향상도 중요하지만 두 가지 이상의 방법이 결합되어 상승 효과를 기대할 수 있는 방법들은 두 가지를 모두 고려한 최적의 알고리즘을 찾는 것이 더 중요하다고 할 수 있다.

## 참고문헌

- [1] Peter V. Loeppert, Timothy K. Wickstrom, "Advanced microphone technology", Emkay Innovative Products
- [2] Heinz Teutsch, Gary W. Elko, "First- and second- order adaptive differential microphone arrays", Agere System, Murray Hill, NJ, USA
- [3] A. Nehorai, B. Porat, "Adaptive comb filtering for harmonic signal enhancement", IEEE Trans. on Acou. Speech Sig Process. ASSP, vol. ASSP-34, pp.1124-1138.
- [4] Koichi Yanagisawa, Kyoko Tanaka, 1999 "Applying comb filter to noise reduction of hearing aid", IEEE SMC '99 Conference Proceedings, vol.6, pp. 352 - 357
- [5] John R. Deller, John H. L. Hansen, John G. Proakis, 2000, Discrete-Time Processing of Speech Signals, IEEE Press, New York