

신호 대 잡음비 추정 방법을 이용한 다중 주파수 밴드 잡음 억제 시스템

Multi frequency band noise suppression system using signal-to-noise ratio estimation

오인규, 이인성[†]

(In Kyu Oh and In Sung Lee[†])

충북대학교 전자통신공학

(Received September 22, 2015; accepted December 30, 2015)

초 록: 본 논문은 밀접한 간격의 두 개의 마이크 배열 환경에서 SNR(Signal-to-Noise Ratio) 추정을 통한 잡음 억제 방법을 제안한다. 기존의 방법은 전 밴드에서 간섭 함수 기반의 SNR 추정을 통해 이득 함수를 얻는 잡음 억제 방법을 사용한다. 그러나 이 방법은 잡음으로 인한 손상이 모든 특징 벡터 성분에 영향을 미쳐 성능을 저하시키는 문제점을 가지고 있다. 따라서 주파수 영역의 신호를 N개의 다중 주파수 밴드로 구분하고 각 밴드별로 간섭 함수 기반의 SNR 추정을 통한 이득 함수를 얻는 잡음 억제 방법을 제안한다. 제안하는 방법의 성능평가는 ITU-T(International Telecommunications Union Telecommunication)에서 제공되는 객관적인 품질 평가 방법인 PESQ(Perceptual Evaluation of Speech Quality)로 비교하여 나타내었다.

핵심용어: 파워스펙트럼 밀도, 교차 파워스펙트럼 밀도, 간섭 함수, 다중 밴드

ABSTRACT: This paper proposes a noise suppression method through SNR (Signal-to Noise Ratio) estimation in the two microphone array environment of close spacing. The conventional method uses a noise suppression method for a gain function obtained through the SNR estimation based on coherence function from full band. However, this method cause performance decreased by the noise damage that affects all the feature vector component. So, we propose a noise suppression method that allocates a frequency domain signal into N constant multi frequency band and each frequency band gets a gain function through SNR estimation based on coherence function. Performance evaluation of the proposed method is shown by comparison with PESQ (Perceptual Evaluation of Speech Quality) value which is an objective quality evaluation method provided by the ITU-T (International Telecommunications Union Telecommunication).

Keywords: PSD (Power Spectrum Density), CSD (Cross-power Spectrum Density), Coherence function, Multi-band

PACS numbers: 43.72.Kb, 43.50.Ed

1. 서 론

최근 사회는 디지털 멀티미디어 통신 서비스가 국 내외적으로 상용화되고 있다. 그러나 데이터를 디지털 화하고 전송하는 과정에서 수많은 원인에 의해 다

양한 잡음들이 발생하고 있으며, 이러한 잡음을 제 거 또는 억제하는 음성향상 기술은 신호처리 분야에 필수적이다. 최근에 연구되는 음성 향상을 위한 기술로는 MEMS(Micro Electro Mechanical System)마이크를 이용한 다채널 음성 향상^[1] 음성향상을 위한 잡음 추적 알고리즘,^[2] 잡음 데이터를 활용한 음성 기저 행렬과 NMF(Non Negative matrix Factorization) 기반 음성 향상 기법^[3] 등 다양한 방식의 음성 향상

[†]Corresponding author: In Sung Lee (inslee@chungbuk.ac.kr)
Department of Radio Communication Eng Chngbuk National University, 1 Chungdae-ro, Seowon-gu, Cheongju, Chungbuk 28644, Republic of Korea
(Tel: 82-43-261-3223)

기술이 연구 진행 되고 있다. 본 논문에서는 잡음을 억제하기 위한 음성향상 기술로 두 개의 마이크로로부터 입력 신호를 받아 시간영역의 신호를 주파수영역의 신호로 변환한 후에 특정 주파수에서 두 신호들 간의 상관의 정도를 나타내는 간섭함수^[4]를 사용하여 진행하였다.

잡음 억제를 위한 기존의 간섭함수 방법은 전 밴드에서 간섭함수 기반의 SNR(Signal-to Noise Ratio) 추정을 통해 이득 함수를 얻는 잡음 시스템을 사용한다.^[5] 두 개의 마이크로로부터 입력 신호들을 받아 시간영역의 신호를 주파수영역의 신호로 변환한 후 신호들 간의 간섭 함수를 구하게 된다. SNR 추정은 간섭함수의 실수 및 허수 부분을 이용하여 SNR 추정을 하게 된다. 추정되어진 SNR 추정치에 Wiener filter^[6]를 적용하여 이득 함수를 얻고 그 후에 이득함수와 입력신호의 합성과 복원 과정을 통해 잡음이 억제된 시간영역의 출력신호를 얻는 알고리즘 방식이다.

이러한 전 밴드 방식에서는 부분적인 제한된 대역에서 잡음으로 인한 손상이 모든 특징 벡터 성분에 영향을 미친다는 문제점을 가지고 있다. 반면 다중 밴드 방식에서는 각각의 밴드마다 독립적으로 음향 특징 분석을 수행함으로써 전 밴드 방식의 문제점을 해결 할 수 있다.

기존 전 밴드 방식의 문제점을 해결하기 위해 본 논문에서는 간섭 함수 기반의 SNR 추정 다중 주파수 밴드 잡음 억제 시스템을 제안한다. 주파수 영역의 신호를 일정한 간격으로 밴드를 분할하여 N개의 밴드로 만들어 다중 주파수 밴드를 만든다. 그 후에 각 밴드별로 간섭 함수를 구하고 이 간섭 함수를 통해 SNR추정 후 Wiener filter에 적용하여 이득함수를 얻고 다시 전 밴드로 만들어주기 위해 이득함수를 원래의 각 밴드별로 배치시키고 이득함수와 입력신호의 합성과 복원 과정을 통해 시간 영역의 잡음이 억제된 출력신호를 얻는 알고리즘을 제안한다.

본 논문은 II장에서 제안하는 SNR 추정 방법을 이용한 다중 주파수 밴드 잡음 억제 시스템에 대해 설명하고, III장에서 제안하는 방법의 성능평가, IV장에서는 결론에 대하여 서술하였다.

II. SNR 추정 방법을 이용한 다중 주파수 밴드 잡음 억제 시스템

먼저 Fig. 1과 같이 소음과 음성 신호가 공간적으로 분리 된 두 개의 밀접한 간격의 무지향성 마이크를 방에 배치한다. 이 경우에 마이크에 의해 기록 된 신호는 다음과 같이 정의 될 수 있다.

$$y_i(m) = x_i(m) + n_i(m) \quad (i = 1, 2). \quad (1)$$

i 는 마이크의 지수, m 은 샘플 지수, x 와 n 은 각각의 마이크의 음성과 잡음을 나타낸다. 시간영역의 입력 신호를 주파수 영역의 신호로 변환하기 위해 Eq.(1)의 양변에 FFT(Fast Fourier Transform)를 적용하면 Eq.(2)와 같이 주파수 도메인에서 표현 될 수 있다.

$$Y_i(w, k) = X_i(w, k) + N_i(w, k) \quad (i = 1, 2). \quad (2)$$

k 는 프레임 지수, w 는 각주파수를 의미한다.

기존의 전 밴드 방식의 문제점을 개선하기 위해 주파수 영역의 신호로 표현된 Eq.(2)를 Fig. 2와 같이 다중 주파수 밴드로 분할하여 2밴드, 4밴드, 8밴드, 16밴드와 같이 N개의 다중 주파수 밴드로 분할한 후 각 밴드별로 간섭함수를 구한 후 SNR 추정 잡음 억제 알고리즘을 수행하게 된다.

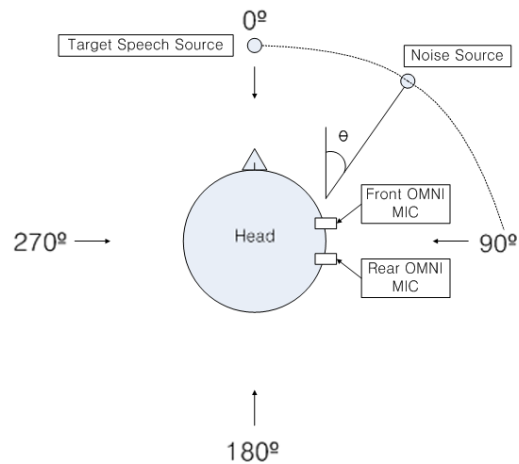


Fig. 1. Placement of two omnidirectional microphones and sound sources.

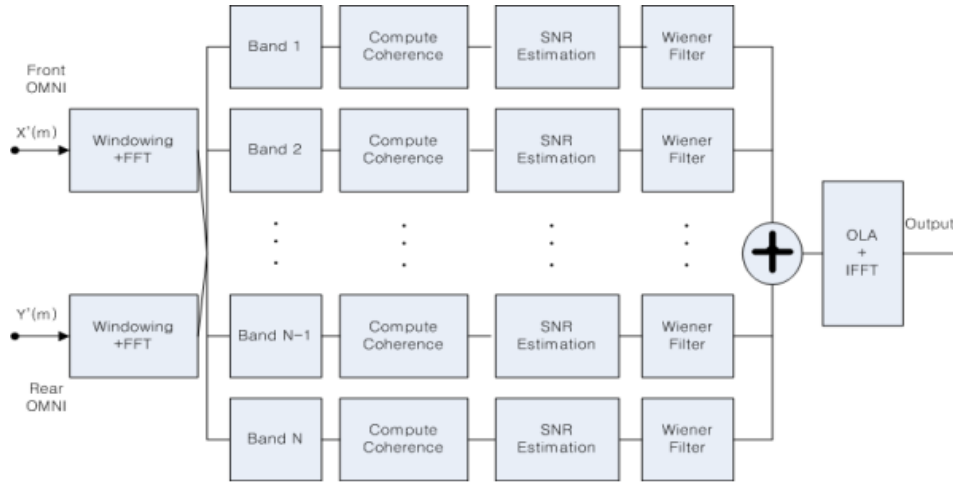


Fig. 2. SNR estimation multi-frequency band noise suppression algorithm of the N-order.

간섭 함수는 두 임의의 프로세스 사이에 선형 관계의 척도이다. 이것은 특정 주파수에서 두 신호의 상관의 정도를 나타낸다. 두 개의 입력 신호 간의 간섭함수는 다음과 같이 정의된다.

$$\Gamma_{y_1 y_2}(j) = \frac{\Phi_{y_1 y_2}(j)}{\sqrt{\Phi_{y_1 y_1}(j)\Phi_{y_2 y_2}(j)}} \quad (j = 1, 2, \dots, N). \quad (3)$$

Φ_{uv} 는 CSD(Cross-Power Spectral Density), Φ_{uu} 는 PSD(Power Spectral Density), j 는 밴드의 지수를 의미한다. 간섭 함수의 크기는 목표 음성 신호가 특정 주파수에서 존재 혹은 부재 여부를 확인하는 객관적인 자료로써 여러 연구에 사용되었다.^[7,8] 간섭함수를 수치화 하였을 때 간섭 함수는 두 신호 간에 상관관계가 있는 경우 1에 가까운 값을 비 상관관계가 있는 경우 0에 가까운 값을 나타낸다.

두 개의 마이크에서 잡음 신호 간 간섭 함수는 음성 신호와 잡음 신호들로부터 계산될 수 있다. 음성 신호와 잡음신호의 구성요소들이 비 상관관계에 있다고 가정하면 마이크로 입력되어진 입력신호의 CSD는 다음과 같이 나타낼 수 있다.

$$\Phi_{y_1 y_2}(j) = \Phi_{x_1 x_2}(j) + \Phi_{n_1 n_2}(j). \quad (4)$$

Eq.(4)를 이용하여 Eq.(3)에 대입을 하면 다음과 같이 표현할 수 있다.

$$\Gamma_{y_1 y_2}(j) = \frac{\Phi_{x_1 x_2}(j)}{\sqrt{\Phi_{y_1 y_2}(j)\Phi_{y_1 y_2}(j)}} + \frac{\Phi_{n_1 n_2}(j)}{\sqrt{\Phi_{y_1 y_1}(j)\Phi_{y_2 y_2}(j)}}. \quad (5)$$

각 채널에서 입력신호의 PSD는 채널에서 음성신호와 잡음신호의 PSD 합이 동일하다는 사실을 이용하면, 방정식을 다음과 같이 표현할 수 있다.

$$\Gamma_{y_1 y_2}(j) = \Gamma_{x_1 x_2}(j) \sqrt{\frac{\Phi_{x_1 x_1}(j)}{\Phi_{x_1 x_1}(j) + \Phi_{n_1 n_1}(j)}} \sqrt{\frac{\Phi_{x_2 x_2}(j)}{\Phi_{x_2 x_2}(j) + \Phi_{n_2 n_2}(j)}} + \Gamma_{n_1 n_2}(j) \sqrt{\frac{\Phi_{n_1 n_1}(j)}{\Phi_{x_1 x_1}(j) + \Phi_{n_1 n_1}(j)}} \sqrt{\frac{\Phi_{n_2 n_2}(j)}{\Phi_{x_2 x_2}(j) + \Phi_{n_2 n_2}(j)}}. \quad (6)$$

SNR은 음성신호의 PSD와 잡음 신호의 PSD를 이용하여 다음과 같이 나타낼 수 있다.

$$SNR_i(j) = \frac{\Phi_{x_i x_i}(j)}{\Phi_{n_i n_i}(j)} \quad (i = 1, 2). \quad (7)$$

Eq.(7)을 이용하여 Eq.(6)에 대입을 하면 다음과 같은 방정식을 얻는다.

$$\Gamma_{y_1y_2}(j) = \Gamma_{x_1x_2}(j) \left[\sqrt{\frac{SNR_1(j)}{1+SNR_1(j)} \frac{SNR_2(j)}{1+SNR_2(j)}} \right] + \Gamma_{n_1n_2}(j) \left[\sqrt{\frac{1}{1+SNR_1(j)} \frac{1}{1+SNR_2(j)}} \right]. \quad (8)$$

밀접한 마이크의 간격에서 SNR_1 과 SNR_2 의 값은 거의 동일하다. 따라서 다음과 같이 SNR_1 과 SNR_2 의 근사치인 SNR 로 나타낼 수 있다.

$$\Gamma_{y_1y_2}(j) \simeq \Gamma_{x_1x_2}(j) \frac{SNR(j)}{1+SNR(j)} + \Gamma_{n_1n_2}(j) \frac{1}{1+SNR(j)}. \quad (9)$$

간섭 함수는 잡음 영역을 기반으로 하여 분석적으로 모델링 될 수 있다. 확산 잡음 영역에서 간섭 함수는 실수이고, 두 마이크 사이의 거리가 감소함으로써 그 값이 증가한다. 확산 잡음 영역에서 두 개의 입력 신호들의 간섭 함수에서 마이크 간격의 영향은 Reference [9]에 보고되었다. 두 개의 입력 신호들의 간섭 함수는 Reference [9]에 의하여 다음과 같이 주어진다.

$$\Gamma_{y_1y_2}(j) \simeq [\cos(\omega\tau) + j \sin(\omega\tau)] \frac{SNR(j)}{1+SNR(j)} + [\cos(\omega\tau\cos\theta) + j \sin(\omega\tau\cos\theta)] \frac{1}{1+SNR(j)}. \quad (10)$$

Eq.(10)에서 $\tau = f_s(d/c)$, $c = 340\text{m/s}$ 소리의 속도, f_s 는 샘플링 주파수, d 는 마이크의 간격, θ 는 잡음 신호의 각도를 나타낸다. N 개의 밴드에서 각 밴드별로 구한 간섭 함수를 이용하여 SNR 추정을 하게 된다. 먼저 각 밴드별로 Eq.(10)에서 실수 부분만 가져온다.

$$R(j) = \frac{SNR(j)}{1+SNR(j)} \cos\alpha + \frac{1}{1+SNR(j)} \cos\beta, \quad (11)$$

$R(j)$ 는 간섭함수 $\Gamma_{y_1y_2}(j)$ 의 실수 부분이고, $\alpha = \omega\tau$, $\beta = \alpha\cos\theta$ 이다. Eq.(11)는 SNR 로 정리하여 다음과 같이 표현 할 수 있다.

$$SNR(j) = \frac{\cos\beta - R(j)}{R(j) - \cos\alpha}. \quad (12)$$

다시 Eq.(10)을 이용하여 허수 부분만 가져온다.

$$I(j) = \frac{SNR}{1+SNR} \sin\alpha + \frac{1}{1+SNR} \sin\beta. \quad (13)$$

$I(j)$ 는 간섭함수 $\Gamma_{y_1y_2}(j)$ 의 허수부분이다. 앞에서와 마찬가지로 SNR 을 구하기 위해 Eq.(13)을 SNR 로 정리하여 다음과 같이 표현 할 수 있다.

$$SNR(j) = \frac{\sin\beta - I(j)}{I(j) - \sin\alpha}. \quad (14)$$

$SNR(j)$ 를 구하는 식인 Eqs.(12)와 (14)를 동일하게 놓고 다음과 같이 방정식을 구한다.

$$[I(j) - \sin\alpha] \cos\beta + [\cos\alpha - R(j)] \sin\beta + R(j) \sin\alpha - I(j) \cos\alpha = 0. \quad (15)$$

Eq.(15)에서, 미지수는 β 뿐이다. 따라서 β 를 구하기 위하여 다음과 같이 치환한다.

$$\begin{cases} A(j) = I(j) - \sin\alpha, \\ B(j) = \cos\alpha - R(j), \\ C(j) = R(j) \sin\alpha - I(j) \cos\alpha. \end{cases} \quad (16)$$

Eq.(16)에 양변을 제곱하고, $\cos^2\beta = 1 - \sin^2\beta$ 라는 식을 이용하여 다음과 같은 방정식을 얻는다.

$$[A(j)^2 + B(j)^2] \sin^2\beta + 2B(j)C(j) \sin\beta + [C(j)^2 - A(j)^2] = 0. \quad (17)$$

근의 공식을 활용하여 Eq.(17)의 근을 구한다.

$$\sin\beta(j) = \frac{-B(j)C(j) \pm |A(j)| \sqrt{A(j)^2 + B(j)^2 - C(j)^2}}{A(j)^2 + B(j)^2}. \quad (18)$$

Eq.(18)에서 $\sin\beta(j)$ 를 구하여 Eq.(14)에 대입하면 $SNR(j)$ 를 구할 수 있다. 그 다음 각 밴드에서 구한 $SNR(j)$ 를 이용하여 Wiener filter에 적용하여 이득함수를 얻을 수 있다.

$$G(j) = \sqrt{\frac{SNR(j)}{SNR(j)+1}} \quad (19)$$

Wiener filter는 신호와 잡음이 섞여 있는 정상 입력에 대한 필터출력과 예측된 희망 출력과의 평균 제곱 오차를 최소화 하는 기준으로 설계되어진 최적의 필터이다.

Fig. 2의 알고리즘과 같이 각각의 밴드에서 Wiener filter를 통해 나온 이득함수를 원래의 주파수 밴드별로 배치시켜 전 밴드의 이득함수로 만들어 준다. 그 후에 전 밴드의 이득함수와 입력신호를 이용하여 OLA(OverLap-Add)를 통해 신호를 합성시킨 후 합성된 주파수 영역의 신호를 다시 입력신호와 같이 시간영역의 신호로 변환하기 위하여 IFFT(Inverse Fast Fourier Transform)를 적용한다. 이 과정을 통해 음성 향상을 위한 잡음을 억제시킨 출력신호가 나오게 된다.

III. 성능평가

제안하는 방법의 성능 평가는 ITU-T(International Telecommunications Union Telecommunication)에서 제공하는 음질 평가 방법인 PESQ(Power Spectrum Density)를 사용한다.^[10] 제안하는 Fig. 2의 알고리즘을 테스트하기 위해 두 마이크의 거리를 15 cm 간격으로 배열시키고, 원음의 위치는 Fig. 1의 0° 위치, 잡음의 위치는 잡음의 개수에 따라서 1개일 경우 Fig. 1의 90° 위치, 2개일 경우 Fig. 1의 90°와 180° 위치로 각각 설정하였다. 그리고 샘플링 주파수는 16 kHz, 주파수는 N개의 band 분할 차수에 따라서 Table 1과 같이 분할하였다. 그리고 SNR을 각각 0 dB로 설정하여 잡음 신호로 CT(Competing Talker), SWN(Speech-Weighted Noise), BAB(Babble noise), WGN(White Gaussian Noise), CN(Car Noise)같은 다양한 환경의 음성을 남성 음성의 원음과 혼합하여 기존의 전 밴드 방법과 제안하는 다중 주파수 밴드 방법을 비교 테스트 하였다.

Fig. 3은 잡음 신호로 CT를 섞어 비교 테스트한 PESQ 결과이다. CT1은 경쟁 화자가 1명, CT2는 경쟁 화자가 2명일 경우로써 기존의 전 밴드와 제안하는 다중 주파수 밴드 방법을 이용하여 2밴드, 4밴드, 8밴드, 16밴드로 나눈 방법을 각 밴드 별로 PESQ 값으로 각

Table 1. Each band frequency of divided N band.

band division order	band number	frequency
1	band 1	1 Hz ~ 16 kHz
2	band 1	1 Hz ~ 8 kHz
	band 2	8001 Hz ~ 16 kHz
4	band 1	1 Hz ~ 4 kHz
	band 2	4001 Hz ~ 8 kHz
	band 3	8001 Hz ~ 12 kHz
	band 4	12001 Hz ~ 16 kHz
8	band 1	1 Hz ~ 2 kHz
	band 2	2001 Hz ~ 4 kHz
	band 3	4001 Hz ~ 6 kHz
	band 4	6001 Hz ~ 8 kHz
	band 5	8001 Hz ~ 10 kHz
	band 6	10001 Hz ~ 12 kHz
	band 7	12001 Hz ~ 14 kHz
	band 8	14001 Hz ~ 16 kHz
16	band 1	1 Hz ~ 1 kHz
	band 2	1001 Hz ~ 2 kHz
	band 3	2001 Hz ~ 3 kHz
	band 4	3001 Hz ~ 4 kHz
	band 5	4001 Hz ~ 5 kHz
	band 6	5001 Hz ~ 6 kHz
	band 7	6001 Hz ~ 7 kHz
	band 8	7001 Hz ~ 8 kHz
	band 9	8001 Hz ~ 9 kHz
	band 10	9001 Hz ~ 10 kHz
	band 11	10001 Hz ~ 11 kHz
	band 12	11001 Hz ~ 12 kHz
	band 13	12001 Hz ~ 13 kHz
	band 14	13001 Hz ~ 14 kHz
	band 15	14001 Hz ~ 15 kHz
	band 16	15001 Hz ~ 16 kHz

각 비교하였다. CT가 1명일 경우 전 밴드방식과 비교하였을 때 제안하는 다중 주파수 밴드방식에서 최고 0.1619, CT가 2명일 경우 전 밴드방식에 비교하였을 때 제안하는 다중 주파수 밴드 방식에서 최고 0.2662 증가한 결과를 볼 수 있었다.

Fig. 4는 잡음신호로 BAB를 혼합한 경우와 SWN을 혼합한 경우 그리고 BAB와 SWN을 2개의 잡음신호를 함께 혼합한 경우로 각 밴드별로 PESQ 값을 비교 하였다. BAB를 혼합한 경우 제안하는 다중 주파수

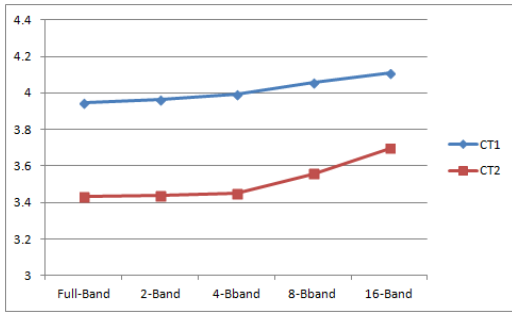


Fig. 3 PESQ Performance evaluation of the proposed algorithm using a CT.

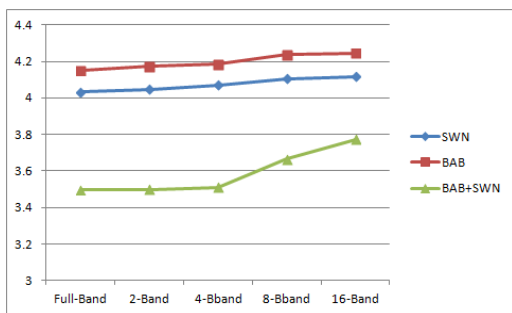


Fig. 4. PESQ Performance evaluation of the proposed algorithm using a BAB, SWN.

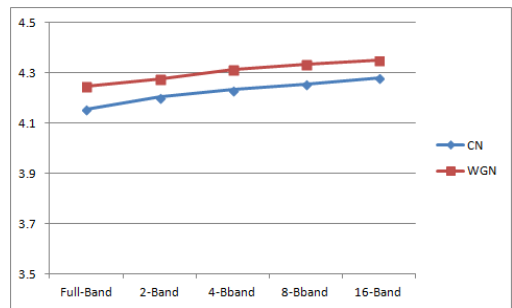


Fig. 5. PESQ Performance evaluation of the proposed algorithm using a CN, WGN.

밴드방식에서 최고 0.0939 증가한 결과를 볼 수 있었고, SWN를 혼합한 경우 제안하는 다중 주파수 밴드 방식에서 최고 0.0822 증가한 결과를 볼 수 있었고, 마지막으로 BAB와 SWN을 함께 혼합한 경우에도 제안하는 다중 주파수 밴드방식에서 최고 0.2745 증가한 결과를 볼 수 있었다.

Fig. 5는 잡음신호로 CN을 혼합한 경우와 WGN을 혼합한 경우로 각 밴드별로 PESQ 값을 비교하였다. CN을 혼합한 경우 제안하는 다중 주파수 밴드방식에서 최고 0.0876 증가한 결과를 볼 수 있었고, WGN

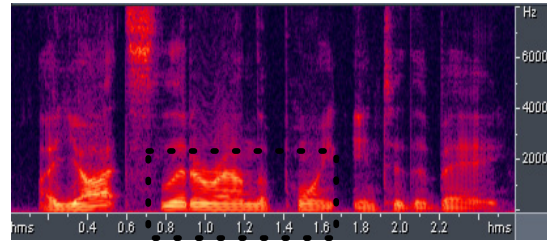


Fig. 6. The spectrogram of the original sound source.

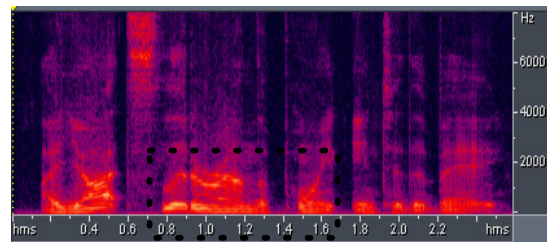


Fig. 7. The spectrogram of conventional full band algorithm using sound mixed a CT 1 person and original sound source.

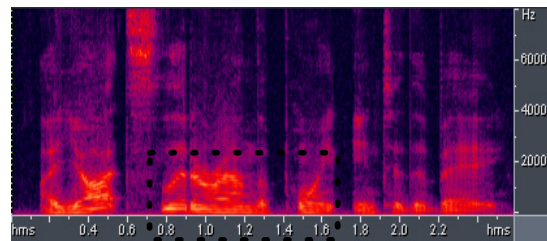


Fig. 8. The spectrogram of proposed multi band algorithm using sound mixed a CT 1 person and original sound source.

을 혼합한 경우에도 제안하는 다중 주파수 밴드방식에서 최고 0.1157 증가한 결과를 볼 수 있었다.

이러한 PESQ 결과에서 볼 수 있듯이 원음과 다양한 잡음을 혼합한 음성 환경에서 제안하는 다중 주파수 밴드 방식을 이용하여 성능 테스트를 하였을 경우에 밴드의 수가 증가함에 따라서 기존의 전 밴드 방식보다 PESQ 값이 더 높은 결과가 나와 잡음 억제 성능이 향상되어 음성의 품질이 향상된 것을 볼 수 있었다.

Fig 6은 남성 음성을 녹음한 원음의 스펙트로그램 결과이고, Fig. 7은 남성 음성의 원음에 CT가 1명인 혼합된 음성을 이용하여 기존의 전 밴드 알고리즘에 적용한 스펙트로그램 결과이며, Fig. 8은 남성 음성의 원음에 CT가 1명인 혼합된 음성을 이용하여 제안

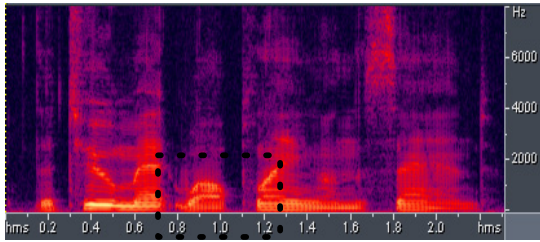


Fig. 9. The spectrogram of the original sound source.

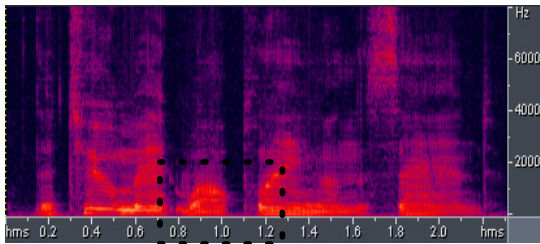


Fig. 10. The spectrogram of conventional full band algorithm using sound mixed a CT 2 person and original sound source.

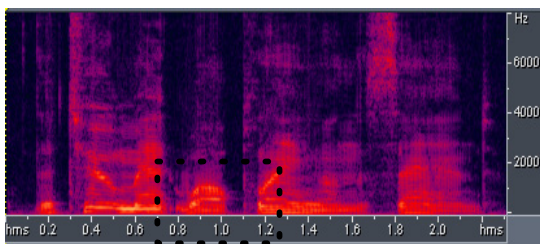


Fig. 11. The spectrogram of proposed multi band algorithm using sound mixed a CT 2 person and original sound source.

하는 다중 주파수 밴드 알고리즘을 적용한 16밴드의 스펙트로그램 결과이다.

스펙트로그램 결과에서 Figs. 6~8의 검은색 실선으로 표시된 0.8s~1.6s 구간을 비교해 보았을 때 기존의 전 밴드 방법보다 제안하는 방법인 다중 주파수 밴드 방법이 잡음을 더 억제하여 더 원음에 가까운 신호의 음성이 복원된 것을 확인할 수 있었다.

Fig. 9는 다른 남성 음성을 녹음한 원음의 스펙트로그램 결과이고, Fig. 10은 남성 음성의 원음에 CT가 2명인 혼합된 음성을 이용하여 기존의 전 밴드 알고리즘에 적용한 스펙트로그램 결과이며, Fig. 11은 남성 음성의 원음에 CT가 2명인 혼합된 음성을 이용하여 제안하는 다중 주파수 밴드 알고리즘을 적용한 16밴드의 스펙트로그램 결과이다.

스펙트로그램 결과 Figs. 9~11의 검은색 실선으로 표시된 0.8s~1.2s 구간을 비교하였을 때 제안하는 방법인 다중 주파수 밴드 방법이 잡음을 더 억제하여 원음에 가깝게 신호의 음성이 복원된 것을 확인할 수 있었다.

IV. 결 론

본 논문에서는 밀접한 간격의 두 개 마이크의 배열 환경에서 주파수 영역의 신호를 N개의 다중 주파수 밴드로 할당하여 각 밴드에서 간섭 함수 기반의 신호 대 잡음비 추정을 통해 이득함수를 얻고 전 밴드로 만들어 주기 위해 이득함수를 원래의 주파수 밴드별로 배치시켜 전 밴드로 만든 후 전 밴드의 이득함수와 입력신호를 이용하여 합성과 복원 과정을 통해 잡음을 억제한 출력 신호가 나오는 알고리즘을 제안하였다.

성능평가에서 보듯이 스펙트로그램 결과에서 제안하는 방법이 기존의 방법보다 잡음을 더 억제하여 원음과 유사한 신호로 복원된 것을 확인하였고, 또한 PESQ 품질 평가에서도 다양한 음성을 실험한 결과 최저 0.0876, 최고 0.2662의 성능이 향상된 것을 확인하여 제안하는 방법에 따라서 음성의 품질이 향상된 것을 확인할 수 있었다.

따라서 향후 두 개의 마이크로폰이 아닌 다수의 마이크로폰 배열 환경에서 잡음 억제를 위한 방법으로 제안한 방법을 통해 연구를 진행하고자 한다.

감사의 글

이 논문은 2014년도 충북대학교 학술연구지원 사업의 연구비 지원에 의하여 연구되었음.

References

1. Z. I. Skordilis, A. Tsiami, P. P. Maragos, G. Potamianos, L. Spelgattil, and R. Sannino, "Multichannel speech enhancement using MEMS microphones," in IEEE ICASSP, 2729-2733 (2015).
2. M. Kalamani, S. Valarmathy, and M. Krishnamoorthi, "Noise tracking algorithm for speech enhancement" NSP

Appl. Math. Inf. Sci. **9**, 691-698 (2015).

3. K. S. Kwon, H. Y Kim, and N. S. Kim, "Speech basis matrix using noise data and NMF-based speech enhancement scheme" (in Korean), J. Soc. KICS, **4**, 04-02 (2015).
4. N. Yousefian, K. Kokkinakis, and P. C. Loizou, "A coherence-based algorithm for noise reduction in dual-microphone application," in Proc, Signal Process, 1904-1908 (2010).
5. N. Yousefian and Philipos C. Loizou, "A dual-microphone algorithm that can cope with competing-talker scenarios," IEEE Trans. Audio, Speech, Lang. Process. **21**, 145-155 (2013).
6. T. Van den Bogaert, S. Doclo, J. Wouters, and M. Moonen, "Speech enhancement with multichannel Wiener filter techniques in multi-microphone binaural hearing aids," J. Acoust Soc. Am. **125**, 360-371, (Jan. 2009).
7. R. L. Bouquin-Jeannes, A. A. Azirani, and G. Faucon, "Enhancement of speech degraded by coherent and incoherent noise using a cross-spectral estimator," IEEE Trans. Speech Audio process. **5**, 484-487 (Sep. 1997).
8. R. L. Bouquin-jeannes and G. Faucon, "Using the coherence function for noise reduction," Proceedings of the IEEE. **139**, 276-280 (1992).
9. M. Brandstein and D. Ward, "Microphone arrays: signal processing techniques and applications." in *handbook of Microphone Arrays*, edited by Brandstein (Springer-Verlog, Berlin, 2011).
10. ITU-T Recommendation, Perceptual evaluation of speech quality(PESQ), an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs, 2011.

저자 약력

▶ 오 인 규 (In Kyu Oh)



2014년: 충북대학교 정보통신과 학사
 2014년 ~ 현재: 충북대학교 전자통신공학과 석사
 <관심분야> 음성/오디오 신호처리

▶ 이 인 성 (In Sung Lee)



1983년: 연세대학교 전자공학과 학사
 1985년: 연세대학교 전자공학과 석사
 1992년: Texas A&M University
 전기공학과 박사
 1993년 ~ 1995년: 한국전자 통신연구원
 이동통신 기술연구단 선임연구원
 1995년 ~ 현재: 충북대학교 전자통신공학과 교수
 <관심분야> 음성영상 신호압축, 이동통신, 적응필터