

논문 2005-42SP-4-13

변조 기법을 이용한 마이크로폰 어레이의 저주파 대역 특성 개선

(Improvement of Microphone Array Performance in the Low Frequencies Using Modulation Technique)

김 기 백*, 조 남 익**

(Gibak Kim and Nam ik Cho)

요 약

본 논문에서는 진폭 변조 기법을 이용하여 빔포머의 저주파 대역 특성을 개선함으로써 전체적인 잡음 제거 성능을 향상하고자 한다. 일차원 등간격 선형 마이크로폰 어레이를 사용하여 잡음 제거를 할 때, 협대역의 잡음 신호에 대해서는 delay-and-sum 빔포밍으로도 적절한 빔폭을 갖는 빔패턴이 형성되므로 효과적으로 잡음을 제거할 수 있다. 그러나 광대역 잡음 신호에서는 aliasing을 피하기 위해 고주파 신호에 맞게 마이크로폰들을 등간격으로 조정하면 저주파로 갈수록 빔폭이 넓어져서 저주파 대역의 잡음은 잘 제거되지 않는다. 광대역 신호에 대해 일정한 빔폭을 갖게 하는 방법들로서 서브 어레이들을 이용한 빔포밍^{[1][2][3][4]}이나 멀티 빔포밍^[5] 등이 제안되어 왔다. 하지만, 이러한 방법들은 주파수에 따라 마이크로폰 간격을 다르게 해야 한다는 원리를 기반으로 하는 것이므로 저주파 대역에서도 고주파 대역에서와 같은 좁은 빔폭을 얻기 위해서는 어레이의 크기가 커져야 하는 단점이 있다. 제안하는 알고리즘은 통신에서 주로 사용하는 진폭 변조 기법을 빔포밍에 이용한 것으로서 저주파 대역 신호를 고주파 대역으로 옮김으로써 작은 크기의 마이크로폰 어레이에서도 저주파 대역 잡음을 줄일 수 있는 장점이 있다. 등간격 선형 마이크로폰 어레이에 제안된 방법을 적용한 실험 결과에서 기존의 방법들에 비해 잡음의 저주파 성분은 잘 제거됨을 알 수 있다.

Abstract

In this paper, we employ the modulation technique for improving the characteristics of beamformer in the low frequencies and thus improving the overall noise reduction performance. In the 1-dimensional uniform linear microphone arrays, we can suppress the narrowband noise component using the delay-and-sum beamforming. But, for the wideband noise signal, the delay-and-sum beamformer does not work well for the reduction of low frequency component because the inter-element spacing is usually set to avoid spatial aliasing at high frequencies. Hence, the beamwidth is not uniform with respect to each frequency and it is usually wider at the low frequencies. In order to obtain the beamwidth independent of frequencies, subarray systems^{[1][2][3][4]} and multi-beamforming^[5] have been proposed. However these algorithms need large space and more microphones since they are based on the theory that the size of the array is proportional to the wavelength of the input signal. In the proposed beamformer, we reduce the low frequency noise by using modulation technique that does not need additional sensors or non-uniform spacing. More precisely, the array signals are split into subbands, and the low frequency components are shifted to high frequencies by modulation and reduced by the delay-and-sum beamforming techniques with small size microphone array. Experimental results show that the proposed technique provides better performance than the conventional ones, especially in the low frequency band.

Keywords : microphone array, beamforming, post-filter, modulation

I. 서 론

* 학생회원, ** 정회원, 서울대학교 전기공학부
(School of Electrical Engineering, Seoul National University)

※ 이 논문은 산업자원부 지원으로 수행하는 21세기 프론티어 연구개발사업(인간기능 생활지원 지능로봇 기술개발사업)의 일환으로 수행되었습니다.
접수일자: 2005년1월14일, 수정완료일: 2005년4월29일

잡음 제거는 실제 환경에서 음성통신이나 음성인식을 위하여 반드시 필요한 요소 중의 하나이다. 실제 환경에서는 여러 가지 잡음의 영향을 많이 받게 되고, 잡음 환경에 따라서 심각한 성능 저하를 초래하는 경우도 많다. 이러한 잡음 제거의 중요성으로 인해 많은 잡음

제거 알고리즘의 연구가 이루어져 왔다. 잡음 제거 기법으로는 잡음이 섞인 음성신호를 자기회귀모델 (autoregressive model) 등으로 모델링하고, 그 파라미터를 워너 필터나 칼만 필터와 같은 최적필터로 추정하는 모수적 (parametric) 기법^[6] 과 주파수 영역 등으로 변환된 신호의 잡음 성분을 추정하고, 이를 이용하여 깨끗한 음성을 추정하는 비모수적 (non-parametric) 기법^{[7][8]} 등이 있다. 하나의 마이크로폰을 사용하는 잡음 제거 방법은 음성과 잡음의 시간-주파수 영역의 정보만을 이용하기 때문에 잡음신호가 음성신호와 시간적으로 구분되거나 주파수 영역에서 구분되는 경우를 제외하고는 음성의 왜곡 없이 잡음을 제거하는 것이 거의 불가능하다. 그러나 잡음신호와 음성신호가 공간상에서 다른 위치에 존재할 때에는 마이크로폰 어레이를 이용하여 공간 영역에서의 필터링을 적용함으로써 음성신호의 왜곡이 거의 없이 잡음 제거가 가능하다.

마이크로폰 어레이를 이용한 잡음 제거 알고리즘에 대해서는 지금까지 많은 연구가 이루어지고 있다. 마이크로폰 어레이를 이용한 알고리즘은 크게 세 가지로 분류할 수 있는데, 적절한 시간 지연을 주어 원하는 신호의 방향으로 빔을 향하게 하고 각 마이크로폰 신호에 가중치를 곱하여 더하는 delay-and-sum 빔포밍과 주어진 제한 조건을 만족하도록 시간에 따라 적응적으로 가중치를 변화시키는 적응 빔포밍, 그리고 빔포밍과 다채널 적응 워너 포스트 필터링을 결합시키는 방법 등으로 분류 된다^[9]. 이러한 방법들 중에 빔포밍과 포스트 필터링을 결합한 방법은 방향성이 있는 코히런트 (coherent) 잡음과 방향성이 없는 인코히런트 (incoherent) 잡음을 모두 줄일 수 있는 장점이 있어서 많은 연구가 진행되고 있다^{[10][11][12]}. 이 방법은 여러 마이크로폰으로부터 들어온 신호를 빔포밍한 신호와 마이크로폰 신호들을 필터링하여 합한 신호의 평균자승오차를 최소화시키도록 필터의 계수를 바꾸어 나간다. 여기서 빔포밍 알고리즘

은 대개 delay-and-sum 빔포밍을 사용하는데, 이 빔포밍 방법은 비교적 간단하고 성능이 좋은 알고리즘으로 알려져 있지만 음성 신호와 같은 광대역 신호에 적용할 때에는 빔패턴이 주파수에 따라 달라진다는 단점을 갖고 있다. 즉, 저주파 대역에서는 넓은 빔폭을 갖고, 고주파 대역에서는 좁은 빔폭을 갖는 빔패턴을 보인다. 이러한 특성 때문에 저주파 대역에서는 원하는 방향으로의 공간 필터링이 잘 되지 않아서, 잡음 성분이 잘 제거 되지 않는다. 음성신호는 저주파 대역에서 큰 에너지를 갖고, 음성학적 정보도 많이 포함하고 있어서 저주파 대역에서 빔패턴이 제대로 형성되지 않으면 공간적 잡음 제거의 성능이 떨어진다. 이러한 문제를 해결하기 위해서 서브 어레이를 사용하는 방법^{[1][2][3][4]} 이나 다중 빔포밍을 사용하는 방법^[5] 등이 제안되어 왔다. 그러나 이러한 방법들은 주파수에 상관없이 같은 빔폭을 얻기 위해서는 어레이의 크기가 주파수에 따라 조정되어야 한다는 원리를 기반으로 하고 있기 때문에 저주파 대역에서 좁은 빔폭을 얻기 위해서는 어레이의 크기가 커져야 하는 단점이 있어서 실용적이지 못하다는 문제점을 갖고 있다.

본 논문에서는 서브 어레이나 다중 빔포밍을 사용하지 않고 단순한 등간격 선형 마이크로폰 어레이 구조를 이용하면서 진폭 변조 기법을 이용함으로써 저주파 대역에서 충분한 잡음 제거를 할 수 있는 방법을 제안한다. 즉, 진폭 변조를 통하여 저주파 대역 신호를 고주파 대역으로 옮겨서 빔폭이 넓게 나타나는 문제를 해결할 수 있도록 한다. 이 알고리즘은 저주파 대역의 잡음을 효과적으로 제거하면서도 작은 크기를 가지며 복잡한 계산이 필요 없다는 장점을 갖고 있다. II장에서는 빔포밍에 대한 기본적인 설명과 주파수에 관계없이 일정한 빔폭을 구현하는 기존의 방법들에 대해 알아본다. III장에서 진폭 변조를 이용한 빔포머를 제안하고 기존의 방법과 비교하며, IV장에서 기존의 알고리즘과 비교한 실험 결과를 보인다.

II. 빔포밍

1. Delay-and-sum 빔포머

본 절에서는 우선 그림 1에서 보인 1차원의 등간격 선형 마이크로폰 어레이의 원리에 대해 살펴보기로 한다. 평면파의 신호가 마이크로폰 어레이의 크기에 비해 먼거리 (far field)에서 같은 정도로 감쇄되어 마이크로폰에 도달한다고 가정하면 각 마이크로폰에 도달하는

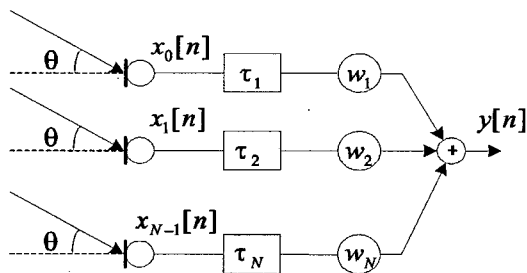


그림 1. delay-and-sum 빔포머의 구조
Fig. 1. The structure of delay-and-sum beamformer.

신호들은 마이크로폰의 위치에 따라 시간 지연만 다른 신호들이 된다. 구체적으로, 마이크로폰 사이의 간격을 d 라고 하고 신호의 속도를 c 라고 하면 1번 마이크로폰에 비해 2번 마이크로폰은 $d\sin\theta/c$ 의 시간만큼 늦게 도달하게 된다. 이러한 도달 시간의 차를 보상해줬을 때의 출력 결과를 다음과 같이 나타낼 수 있다.

$$\begin{aligned} x_{n+1}(t) &= x_n(t - \tau_0(\theta)), \quad \tau_0(\theta) = d\sin\theta/c \\ x_n(t) &= x_0(t - n\tau_0(\theta)) \\ y(t) &= \sum_{n=0}^{N-1} a_n x_n(t + n\tau_0(\theta)) \\ &= \sum_{n=0}^{N-1} a_n x_0(t). \end{aligned} \quad (1)$$

여기서 $a_n = 1/N$ 으로 하면 $y(t) = x_0(t)$ 이 된다. 입사하는 신호의 감쇄가 없다고 가정한다면, 출력은 입사하는 신호 자체가 된다. 즉, θ 방향에 대해서는 감쇄가 없게 된다. 이렇게 θ 방향으로 향하도록 만든 시스템에 ϕ 각도의 입사각을 갖는 다른 신호가 들어온다고 가정하자. 이 경우의 출력을 구해보면 다음과 같다.

$$\begin{aligned} \tau_0(\phi) &= d\sin\phi/c \\ x_n(t) &= x_0(t - n\tau_0(\phi)) \\ y(t, \phi; \theta) &= \sum_{n=0}^{N-1} a_n x_n(t + n\tau_0(\theta)) \\ &= \sum_{n=0}^{N-1} a_n x_0(t + n(\tau_0(\theta) - \tau_0(\phi))) \\ &= \sum_{n=0}^{N-1} a_n x_0(t + nd(\sin\theta - \sin\phi)/c). \end{aligned} \quad (2)$$

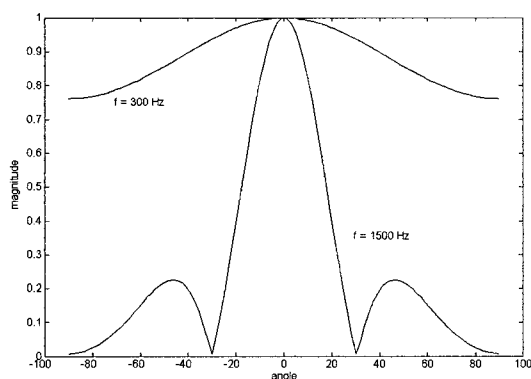


그림 2. 300 Hz 와 1500 Hz 에서의 delay-and-sum 빔포머의 빔패턴
Fig. 2. Beampattern of delay-and-sum beamformer at 300 Hz and 1500 Hz.

이것을 주파수 영역에서 표현하면 다음과 같이 쓸 수 있다.

$$\begin{aligned} Y(\omega, \theta) &= \sum_{n=0}^{N-1} a_n X_0(\omega) e^{-j\omega nd(\sin\phi - \sin\theta)/c} \\ H(\omega, \phi; \theta) &= \frac{Y(\omega, \theta)}{X_0(\omega)} \\ &= \sum_{n=0}^{N-1} a_n e^{-j\omega nd(\sin\phi - \sin\theta)/c}. \end{aligned} \quad (3)$$

$\theta = 0^\circ$ 로 빔포머가 향해 있고, $a_n = 1/N$ 일 때 이는 다음과 같이 표현할 수 있다.

$$\begin{aligned} H(\omega, \phi; 0^\circ) &= \frac{1}{N} \sum_{n=0}^{N-1} e^{-j\omega nd\sin\phi/c} \\ &= \frac{1}{N} \frac{1 - e^{-j\omega Nd\sin\phi/c}}{1 - e^{-j\omega d\sin\phi/c}} \\ &= \frac{1}{N} \frac{e^{-j\omega \frac{Nd}{2c}\sin\phi} (e^{j\omega \frac{Nd}{2c}\sin\phi} - e^{-j\omega \frac{Nd}{2c}\sin\phi})}{e^{-j\omega \frac{d}{2c}\sin\phi} (e^{j\omega \frac{d}{2c}\sin\phi} - e^{-j\omega \frac{d}{2c}\sin\phi})} \end{aligned} \quad (4)$$

$$H(f, \phi; 0^\circ) = \frac{1}{N} e^{-j\pi f \frac{(N-1)d}{c} \sin\phi} \frac{\sin(\pi f N d u / c)}{\sin(\pi f d u / c)} \quad (5)$$

$$H(f, \phi; 0^\circ) = \frac{\text{sinc}(f N d u / c)}{\text{sinc}(f d u / c)}. \quad (6)$$

여기서 $u = \sin\phi$ 이고, 앞의 식에서 구한 빔패턴을 주파수가 300 Hz, 1500 Hz 일 때 각각 그려보면 그림 2와 같이 된다. 지금 빔포머가 0° 에 맞춰져 있으므로 $\phi = 0^\circ$ 에서는 1의 응답을 갖는다. 그리고 ϕ 가 점점 0° 에서 멀어질수록 낮은 응답을 갖게 되어 공간 필터링이 이루어진다. 그러나 그림에서 보듯이 delay-and-sum 빔포머는 주파수에 따라 다른 빔패턴을 가지며, 특히 주파수가 낮아질수록 넓은 빔폭을 갖게 됨을 알 수 있다. 따라서 잡음신호가 다른 방향에서 들어왔을 때, 전체 주파수 영역에 대해 균일하게 감쇄되지 않는다. 즉, 다른 방향에서 오는 신호 중 고주파 대역 성분은 많이 감쇄되고, 저주파 대역 성분은 적게 감쇄되는 것이다.

III. 제안하는 알고리즘

앞에서는 광대역 신호에 대해 주파수에 관계없이 같

은 빔폭을 갖도록 하는 기존의 알고리즘에 대해 알아보았으며 이 알고리즘들은 광대역 신호에 대해 좋은 공간 응답을 나타내고 있음을 알 수 있었다. 그러나 서브 어레이 방법이나 멀티 빔포밍은 저주파 대역에서 충분한 공간 분해능을 얻기 위해서는 수 미터 정도의 긴 마이크로폰 어레이를 필요로 한다. 이렇게 긴 마이크로폰의 경우는 제한적으로밖에 사용할 수 없어서 실용성이 떨어지게 된다. 따라서 본 논문에서는 마이크로폰 어레이의 크기를 키우지 않고서도 저주파 대역에서 좁은 빔폭을 갖도록 하기 위하여 진폭 변조를 통하여 입력 신호를 고주파 대역으로 이동시키는 새로운 방법을 제시한다.

1. 변조 기법을 이용한 빔포밍

진폭 변조는 일반적으로 통신 시스템에서 기저대역 신호를 반송파를 이용해서 높은 주파수 대역으로 옮기는데 사용되는 방법이다. 즉, 높은 주파수의 반송파를 기저대역 신호에 곱하여 기저대역 신호를 반송파의 주파수 대역으로 옮길 수 있다. 진폭 변조의 식은 다음과 같다^[13].

$$c(t) = A_c \cos(2\pi f_c t) \quad (7)$$

$$s(t) = A_c [1 + k_a m(t)] \cos(2\pi f_c t) \\ |k_a m(t)| < 1 \quad \text{for all } t. \quad (8)$$

$m(t)$ 는 기저대역의 신호이고, $c(t)$ 는 반송파이며, $s(t)$ 는 변조된 신호이다. 변조신호에서 위상이 반전되는 것을 막기 위해 $k_a m(t)$ 의 절대값이 1보다 작게 만든다.

진폭 변조를 빔포밍에 이용하면 저주파 대역의 신호에 대해 고주파 대역에서의 공간 응답 특성을 얻을 수 있다. 저주파 대역의 신호를 높은 주파수 f_c 를 가진 반송파를 곱하여 신호를 높은 주파수 대역으로 옮긴 후 빔포밍하고, 빔포밍한 결과에 다시 반송파를 곱하고 저주파대역 필터를 통과시켜 복조하면 된다. 이 때, 반송파의 주파수는 기저대역 신호의 가장 높은 주파수보다 높은 값이 되어야 하며, 그 주파수에서의 빔폭이 충분히 좁아야 한다. 반송파의 위상은 각 마이크로폰마다 잡음신호의 위상에 맞춰 정해야 한다. 잡음 신호는 마이크로폰마다 도달 시간에 따른 위상차가 발생하며, 이 위상차를 각 마이크로폰에 곱해지는 반송파에도 반영시켜 주어야 하는 것이다. 잘 알려진 바와 같이 진폭 변복

조의 특성은 반송파 위상의 정확성에 크게 영향을 받으므로 정확히 도달 지연 시간을 추정하여야 제안하는 알고리즘이 효율적으로 동작한다.

도달 지연 시간은 정규 상호 상관도 (Normalized Cross Correlation), LMS 적응 필터, Crosspower Spectrum Phase 등의 방법들로 추정할 수 있다^[14]. 본 논문에서는 정규 상호 상관도를 이용하여 도달 지연 시간을 추정하였다. 정규 상호 상관도를 이용한 도달 지연 시간 추정 방법은 다음과 같은 과정을 통해 구할 수 있다. 먼저, i 번째 마이크로폰을 통해 들어오는 신호는 원신호에 시간 지연과 감쇄, 잡음이 섞여 들어오는 신호로 모델링할 수 있으며, 다음과 같이 나타낼 수 있다.

$$s_i(t) = \alpha_i r(t - \tau_i) + n_i(t) \quad (9)$$

이 때, i 번째 마이크로폰에 들어오는 신호 $s_i(t)$ 와 j 번째 마이크로폰의 신호 $s_j(t)$ 간의 시간 지연 차이는 다음과 같다.

$$\delta_{ij} = \tau_j - \tau_i \quad (10)$$

도달 시간의 차이는 $s_i(t)$ 와 $s_j(t)$ 의 상호 상관도를 이용하여 구할 수 있는데, 이 때의 상호 상관도는 다음과 같이 표현할 수 있다.

$$R_{ij}(\tau) = E[s_i(t)s_j(t+\tau)] \\ = \alpha_i \alpha_j R_{rr}(\tau - \delta_{ij}) + R_{n_i n_j}(\tau). \quad (11)$$

공간적으로 상관성이 없는 (spatially uncorrelated) 잡음이라 가정하면, 서로 다른 마이크로폰간의 잡음 신호 상관도는 0이 되므로, $s_i(t)$ 와 $s_j(t)$ 의 상호 상관도는 $\tau = \delta_{ij}$ 일 때 최대가 된다. 상호 상관도는 아래와 같이 시간 T 동안의 평균값으로 추정할 수 있다.

$$\hat{R}_{ij}(t, \tau) = \frac{1}{T} \int_{t-T/2}^{t+T/2} s_i(u)s_j(u+\tau)du. \quad (12)$$

이것을 정규화하면 다음 식을 얻을 수 있다.

$$\hat{R}_{ij}^{(M)}(t, \tau) \\ = \frac{\int_{t-T/2}^{t+T/2} s_i(u)s_j(u+\tau)du}{\sqrt{\int_{t-T/2}^{t+T/2} s_i^2(u)du} \sqrt{\int_{t-T/2}^{t+T/2} s_j^2(u+\tau)du}}. \quad (13)$$

따라서 시간 t 에서의 두 마이크로폰 간 도달 시간 지연 차이는 다음 식으로부터 구할 수 있다.

$$\hat{\delta}_{ij}(t) = \underset{\tau}{\operatorname{argmax}} \hat{R}_{ij}^{(N)}(t, \tau). \quad (14)$$

다음은, 변조를 이용하여 주파수를 높여서 빔포밍했을 때의 공간 응답을 분석해 본다. 등간격 선형 마이크로폰 어레이에 대해 빔포머가 정면을 향하고 있고, 각 마이크로폰에 들어오는 신호는 시간 지연만 다른 신호이고, 각 마이크로폰의 가중치는 $1/N$ 이라고 가정하면 앞에서 본 바와 같이 delay-and-sum 빔포머의 시간 영역과 주파수 영역에서의 결과는 다음과 같이 나타난다.

$$y(t, \theta) = \frac{1}{N} \sum_{n=0}^{N-1} x_0(t - n\tau_0 \sin \theta) \quad (15)$$

$$\leftrightarrow Y(\omega, \theta) = \frac{1}{N} \sum_{n=0}^{N-1} X_0(\omega) e^{-j\omega n\tau_0 \sin \theta} \quad (16)$$

$$H(\omega, \theta) = \frac{1}{N} \sum_{n=0}^{N-1} e^{-j\omega n\tau_0 \sin \theta} \quad (17)$$

$$\tau_0 = d/c. \quad (18)$$

여기서 입력에 반송파를 곱한 신호는 다음과 같이 나타낼 수 있고

$$s_n(t) = x_0(t - n\tau_0 \sin \theta) \cos(\omega_c(t - n\tau_0 \sin \theta)) \quad (19)$$

이것의 가중합은 다음과 같다.

$$\begin{aligned} \tilde{y}(t, \theta) &= \frac{1}{N} \sum_{n=0}^{N-1} x_0(t - n\tau_0 \sin \theta) \cos(\omega_c(t - n\tau_0 \sin \theta)). \end{aligned} \quad (20)$$

이것을 주파수 영역으로 변환하면,

$$\begin{aligned} \tilde{Y}(\omega, \theta) &= \frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{2} [X_0(\omega - \omega_c) + X_0(\omega + \omega_c)] e^{-j\omega n\tau_0 \sin \theta} \\ &= \frac{X_0(\omega - \omega_c) + X_0(\omega + \omega_c)}{2N} \sum_{n=0}^{N-1} e^{-j\omega n\tau_0 \sin \theta} \end{aligned} \quad (21)$$

으로 나타난다.

여기서 ω_b 가 기저대역 신호가 갖는 최대 주파수라면, 반송파의 주파수는 $\omega_c > \omega_b$ 의 조건을 만족해야 한다.

위의 신호를 복조하기 위하여 먼저 반송파를 곱하면 다음과 같이 표현된다.

$$\begin{aligned} \tilde{Y}'(\omega, \theta) &= \frac{X_0(\omega - 2\omega_c) + X_0(\omega)}{4N} \sum_{n=0}^{N-1} e^{-j(\omega - \omega_c)n\tau_0 \sin \theta} \\ &+ \frac{X_0(\omega) + X_0(\omega + 2\omega_c)}{4N} \sum_{n=0}^{N-1} e^{-j(\omega + \omega_c)n\tau_0 \sin \theta}. \end{aligned} \quad (22)$$

여기서 $X_0(\omega - 2\omega_c)$ 성분과 $X_0(\omega + 2\omega_c)$ 성분을 없애기 위해 $2\omega_c - \omega_b$ 의 차단 주파수를 갖는 저역통과 필터로 필터링을 한 결과까지의 과정을 전달함수, 즉 공간 응답으로 나타내면 다음과 같다.

$$\begin{aligned} \tilde{H}(\omega, \theta) &= \tilde{Y}'(\omega, \theta)/X_0(\omega) \\ &= \frac{1}{2N} \sum_{n=0}^{N-1} [e^{-j(\omega - \omega_c)n\tau_0 \sin \theta} + e^{-j(\omega + \omega_c)n\tau_0 \sin \theta}] \\ &= \sum_{n=0}^{N-1} B_n(\omega, \theta) e^{-j\omega n\tau_0 \sin \theta} \end{aligned} \quad (23)$$

$$B_n(\omega, \theta) = \frac{1}{N} \cos(\omega_c n\tau_0 \sin \theta). \quad (24)$$

결국, 식 (23)이 잡음신호 구간에서의 주파수에 따른 공간 응답이 된다. 그림 3은 delay-and-sum 빔포머와 제안하는 빔포머의 공간 응답을 비교한 것이다. 여기서 반송파의 주파수는 3000 Hz 이고, 300 Hz ~ 1000 Hz 구간에 대해 결과를 표시하였다. 그림에서 보듯이 제안하는 빔포머의 빔패턴은 주어진 주파수 대역에서 주파수와 상관없이 동일한 빔폭을 가지며 좁은 빔폭을 가짐을 알 수 있다. 빔폭을 나타내는 방법에는 여러 가지가 있는데, 대표적으로 공간 응답이 처음으로 0이 되는 두 지점 사이를 말하는 null-to-null 빔폭과 파워가 절반으로 줄어드는 지점 사이를 말하는 3-dB 빔폭 (Half Power Beamwidth)이 많이 사용된다. $\theta = 0^\circ$ 로 빔포머가 향해 있을 때, delay-and-sum 빔포머의 경우, 파워가 1/2이 되는 점을 찾기 위해 식 (6)으로부터 테일러 전개를 이용하여 해를 구하면 다음과 같다.

$$|H(f, \phi; 0^\circ)|^2 = \frac{1}{2}$$

$$u \approx \pm \sqrt{\frac{3}{2}} \frac{c}{\pi N d f}$$

$$HPBW_u \approx \sqrt{6} \frac{c}{\pi N d f}$$

$$HPBW_\theta \approx 2 \sin^{-1} \left(\sqrt{\frac{3}{2}} \frac{c}{\pi N d f} \right)$$

제안하는 빔포머의 공간응답의 절대값은 다음과 같이 쓸 수 있다.

$$|\tilde{H}(f, \phi; 0^\circ)| = \frac{1}{2N} \left[\frac{\sin \left(\frac{\pi N d (f + f_c) u}{c} \right)}{\sin \left(\frac{\pi d (f + f_c) u}{c} \right)} + \frac{\sin \left(\frac{\pi N d (f - f_c) u}{c} \right)}{\sin \left(\frac{\pi d (f - f_c) u}{c} \right)} \right]$$

이 식으로부터 delay-and-sum 빔포머에서 구한 것과 같은 방법으로 3-dB 빔폭을 구하면 다음과 같다.

$$|H(f, \phi; 0^\circ)|^2 = \frac{1}{2}$$

$$u \approx \pm \sqrt{\frac{3}{2}} \frac{c}{\pi N d \sqrt{f^2 + f_c^2}}$$

$$HPBW_u \approx \sqrt{6} \frac{c}{\pi N d \sqrt{f^2 + f_c^2}}$$

$$HPBW_\theta \approx 2 \sin^{-1} \left(\sqrt{\frac{3}{2}} \frac{c}{\pi N d \sqrt{f^2 + f_c^2}} \right)$$

식 (25)과 식 (27)을 비교해 보면 제안하는 알고리즘에서는 delay-and-sum 빔포머의 수식에 있는 f 대신 $\sqrt{f^2 + f_c^2}$ 이 나타난다. $f_c \gg f$ 이라면 입력신호의 주파수에 상관없이 반송파의 주파수에 따라 빔폭이 결정된다는 것을 알 수 있으며 이 조건이 만족하지 않더라도 $f < \frac{f_c}{2}$ 이므로 빔폭은 입력신호의 주파수의 영향을 적게 받는다. 구체적인 예로 5 cm 간격으로 9개의 마이크로폰으로 어레이를 구성했을 경우, 300 Hz ~ 1000 Hz 구간에 대해 식 (25), (27)로부터의 빔폭 변화를 비교하면, delay-and-sum 빔포머의 경우는 300 Hz 일 때 158° , 1000 Hz 일 때 34.3° 로 크게 변하는 반면, 제안하는 빔포머의 빔폭은 300 Hz 일 때 11.2° , 1000 Hz 일 때 10.7° 로 별로 변하지 않는다. 결과에서 보듯이, 제안한 알고리즘은 어레이의 크기를 증가시키지 않고도 저주파 신호에 대해 고주파 신호에서의 빔폭과 비슷한 결과를 얻을 수 있는 장점이 있다.

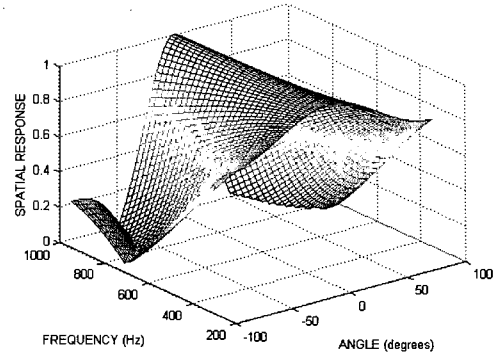


그림 3-(a). delay-and-sum 빔포머의 빔패턴
Fig. 3-(a). Beampattern of delay-and-sum beamformer.

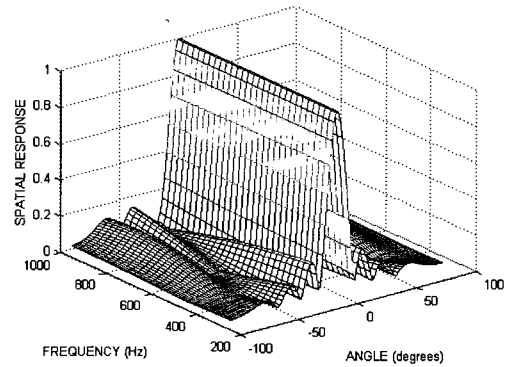


그림 3-(b). 제안하는 알고리즘의 빔패턴
Fig. 3-(b). Beampattern of proposed algorithm.

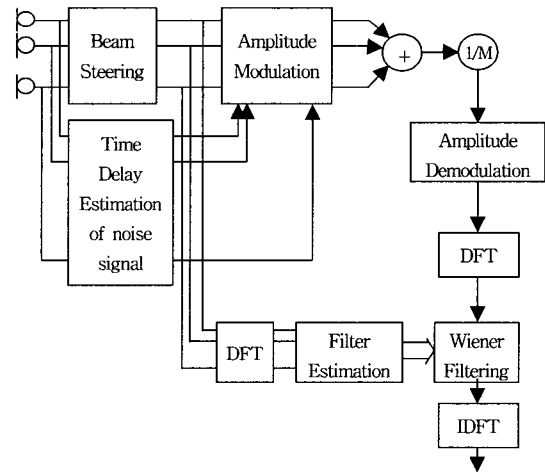


그림 4. 잡음 제거 시스템의 블록 다이어그램
Fig. 4. Block diagram of noise reduction system.

2. 포스트 필터링과 변조 기법을 이용한 빔포밍

마이크로폰 어레이로 음성을 취득할 때 발생하는 잡음의 종류를 구분하면, 다른 방향의 음원에서 발생하여 공간 필터링으로 감쇄시킬 수 있는 코히어런트(coherent) 잡음과 마이크로폰 자체의 잡음이나 반향 때

문에 발생하는 인코히어런트 (incoherent) 잡음으로 분류할 수 있다. 지금까지 살펴본 빔포밍은 주로 방향성이 있는 코히어런트 잡음을 제거하는데 효과적이며 인코히어런트 잡음을 보다 효과적으로 제거하기 위해서 빔포밍과 포스트 필터링을 결합한 구조를 갖는 잡음 제거 시스템을 많이 사용한다^{[10][11][12]}. 마이크로폰 어레이에서 포스트 필터링은 최적 잡음 제거 필터에 빔포밍 결과를 통과시키는 것을 의미한다. 여기서는 최소평균자승오차 (minimum mean squared error) 특성을 갖는 위너 필터 (Wiener filter) 를 사용한다.

도달 지연 시간이 보상된 후의 마이크로폰에서의 신호는 다음과 같이 표현할 수 있다.

$$x_i(t) = s(t) + n_i(t), \quad i = 1, \dots, N \quad (28)$$

$s(t)$ 는 목적 신호, $n_i(t)$ 는 잡음신호이다. 이 때, 위너 필터는 다음과 같이 계산된다.

$$\begin{aligned} W(\omega) &= \frac{\Phi_{ys}(\omega)}{\Phi_{yy}(\omega)} \\ &= \frac{\Phi_{ss}(\omega)}{\Phi_{ss}(\omega) + \Phi_{nn}(\omega)} \\ &\approx \frac{2}{N(N-1)} \frac{\sum_{i=0}^{N-2} \sum_{j=i+1}^{N-1} \text{Re}\{X_i(\omega)X_j^*(\omega)\}}{\frac{1}{N} \sum_{i=0}^{N-1} |X_i(\omega)|^2} \end{aligned} \quad (29)$$

여기서, y 는 빔포머의 출력 신호이고, $\Phi_{nn}(\omega)$ 는 빔포머를 통과한 잡음 신호의 전력 스펙트럼 밀도 (power spectral density)이다. 전력 스펙트럼을 추정하는 과정에서 서로 다른 마이크로폰의 잡음 신호는 상관 관계가 없으며, 잡음 신호와 목적 신호도 상관 관계가 없다고 가정

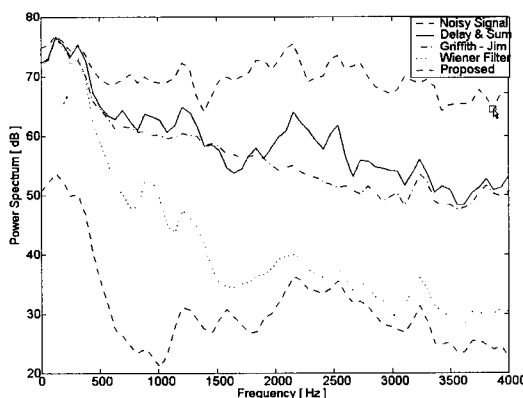


그림 5. 잡음 제거 후 잡음 구간의 전력 스펙트럼 비교

Fig. 5. Noise power spectra at the output of noise reduction systems.

하였다. 그림 4는 빔포머와 포스트 필터를 결합시킨 마이크로폰 어레이의 잡음 제거 시스템을 보여준다.

IV. 실험 결과

제안된 방법의 잡음 제거 성능을 알아보기 위해 우선 등간격 선형 마이크로폰 어레이를 제작하였다. 마이크로폰의 개수는 8개이며 마이크로폰 간의 간격은 5 cm로 정한다. 8 채널을 입력으로 하는 A/D 보드를 사용하고, 마이크로폰은 무지향성의 일렉트렛 마이크로폰을 사용하였으며, 샘플링 주파수는 8000 Hz로 하였다. 발생자는 마이크로폰 어레이의 정면에 위치하고, 그 거리는 1 m 정도로 하며, 컴퓨터 잡음 등이 존재하는 일반 연구실에서 녹음하였다. 방향성 있는 잡음 발생을 위해서 마이크로폰 어레이 정면에서 90° 위치에 스피커를 두고 백색 가우시안 잡음을 재생하였다. 그 거리는 음성 신호와 마찬가지로 1 m 정도로 한다.

신호를 서브밴드로 나누기 위해 코사인 변조 필터뱅크 (cosine-modulated filter bank)를 사용한다^[15]. 이 필터뱅크는 128차 500 Hz의 차단 주파수를 갖는 저역 통과 필터를 코사인 변조하여 4개의 균등한 필터들로 구성한 것이다. 이 때, 각 필터의 대역폭은 1000 Hz가 된다. 4개의 대역 각각에 대해 복소포락 (complex envelope)을 구하여 기저대역으로 옮긴 다음, 같은 주파수 (3000 Hz)의 반송파를 이용하여 진폭 변조를 한 다음, 빔포밍을 한다. 반송파의 주파수는 빔포머의 지향성 (directivity)과 잡음 방향 추정 오차에 대한 강인성을 적절히 고려하여 선택한다. 성능 비교를 위해 delay-and-sum 빔포머, Griffith-Jim 적응 빔포머^[16], 위너 포스트 필터^[10]를 각각 적용하였고, 주파수 영역에서 잡음의 전력 스펙트럼을 그림 5에 나타내었다. 그림에서 알 수 있듯이 제안하는 알고리즘은 저주파 대역(300 Hz ~ 1000 Hz)에서 다른 방법들보다 높은 잡음 제거 성능을 나타냄을 알 수 있다.

V. 결론

본 논문에서는 등간격 선형 마이크로폰 어레이를 이용한 잡음 제거 시스템을 제안하였다. 일반적인 등간격 마이크로폰 어레이에서 delay-and-sum 빔포밍의 단점인 저주파 대역에서의 잡음 제거 성능 저하를 개선하기 위해 진폭 변조 기법을 이용하였다. 신호를 4 주파수 대역으로 나누고 각 대역 신호를 기저대역 신호로 바꾼

다음 진폭 변조 빔포밍을 함으로써 균일한 빔폭을 갖도록 하였다. 제안하는 알고리즘은 기존의 방법들과는 달리 어레이의 크기를 변화시키지 않으면서도 저주파 대역에서의 좁은 빔폭을 얻을 수 있어서 비교적 작은 크기의 어레이로 좋은 잡음 제거 성능을 얻을 수 있다.

참고 문헌

- [1] D. B. Ward, R. A. Kennedy, and R. C. Williamson, "Theory and design of broadband sensor arrays with frequency invariant far-field beam patterns," J. Acoust. Soc. Amer., vol. 97, no. 2, pp. 1023-1034, Feb. 1995.
- [2] M. Brandstein and D. B. Ward, Microphone Arrays : Signal Processing Techniques and Applications, Springer-Verlag, May 2001.
- [3] T. Chou, "Frequency-Independent Beamformer with Low Response Error," in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 5, pp. 2995-2998, 1995.
- [4] R. J. Webster and T. N. Lang, "Prescribed Sidelobes for the Constant Beamwidth Array," IEEE Trans. Acoustics, Speech and Signal Processing, vol. 38, no. 4, April 1990.
- [5] M. M. Goodwin and G. W. Elko, "Constant Beamwidth Beamforming," in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 1, pp. 169-172, 1993.
- [6] J. S. Lim and A. V. Oppenheim, "All-pole modeling of degraded speech." IEEE Trans. Acoust., Speech, Signal Processing, 26(3):197 - 210, June 1978.
- [7] S. F. Boll, "Suppression of Acoustic Noise in Speech Using Spectral Subtraction," IEEE Trans. Acoust., Speech, Signal Processing, 27(2):113 - 120, April 1979.
- [8] Y. Ephraim and D. Malah, "Speech Enhancement Using a Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator," IEEE Trans. Acoust., Speech, Signal Processing, 32(6):1109 - 1121, December 1984.
- [9] S. Fischer, K. U. Simmer, "Beamforming Microphone Arrays for Speech Acquisition in Noise Environments," Speech Communication, 20, pp. 215-227, 1996.
- [10] R. Zelinski, "A microphone array with adaptive post-filtering for noise reduction in reverberant rooms," in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 5, pp. 2578-2581, 1988.
- [11] C. Marro, Y. Mahieux, and K. U. Simmer, "Analysis of noise reduction and dereverberation techniques based on microphone arrays with postfiltering," IEEE Trans. on Speech and Audio Processing, vol. 6, no. 3, May 1998.
- [12] I. Cohen, "Analysis of Two-Channel Generalized Sidelobe Canceller (GSC) With Post-Filtering," IEEE Trans. Speech and Audio Processing, vol. 11, no. 6, November 2003.
- [13] S. Haykin, Communication Systems 4th ed., John Wiley & Sons, Inc., New York, U.S.A., 2001.
- [14] M. Omologo and P. Svaizer, "Acoustic Event Localization Using a Crosspower-spectrum Phase based Technique," in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 2, pp. 273-276, 1994.
- [15] P. Vaidyanathan, Multirate Systems and Filter Banks, Englewood Cliffs, NJ: Prentice Hall, 1993.
- [16] L. J. Griffiths, C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," IEEE Trans. on Antennas Propagation, vol. 30, no. 1, pp. 27-34, Jan. 1982.

저 자 소 개



김 기 백(학생회원)

1994년 서울대학교 전자공학과
학사 졸업.

1996년 서울대학교 전자공학과
석사 졸업.

1996년~2000년 LG 전자기술원
연구원.

2000년~2003년 (주)보이스웨어 연구원

2003년~현재 서울대학교 전기공학부 박사과정.

<주관심분야 : Array Processing, Speech
Enhancement>



조 남 익(정회원)

1986년 서울대학교 제어계측학과
학사 졸업.

1988년 서울대학교 제어계측학과
석사 졸업.

1992년 서울대학교 제어계측학과
박사 졸업.

1991년~1994년 제어계측신기술연구센터 연구원.

1994년~1998년 서울시립대학교 전기공학부 교수.

1999년~현재 서울대학교 전기공학부 교수

<주관심분야 : 신호 처리, 영상 처리, 적응 필터>