

MUSIC 및 반향 성분 제거 기법을 이용한 음성신호의 입사각 추정

장형욱 · 정상배 · 김영일*

Direction-of-Arrival Estimation of Speech Signals Based on MUSIC and Reverberation Component Reduction

Hyungwook Chang · Sangbae Jeong · Youngil Kim*

Department of Electronics Engineering, Gyeongsang National University, Jinju 660-701, Korea

요 약

본 논문에서는 다중 신호 분류 기반의 음성신호의 입사각 추정 향상 방법을 제안한다. 기본적으로 제안한 방식은 복소 대역통과 필터를 이용하여 신호 분석을 위한 협대역 신호를 생성한다. 또한, 공간 스펙트럼에서의 반향 성분 제거 및 2차 함수 기반의 응답 근사화를 사용하여 추정 각도의 정확도를 향상시켰다. 실험결과 제안한 방법은 일반화된 상호상관도 방식의 입사각 추정 알고리즘보다 검출 오차 및 검출 성공률 측면에서 더 좋은 성능을 보였다.

ABSTRACT

In this paper, we propose a method to improve the performance of the direction-of-arrival (DOA) estimation of a speech source using a multiple signal classification (MUSIC)-based algorithm. Basically, the proposed algorithm utilizes a complex coefficient band pass filter to generate the narrow band signals for signal analysis. Also, reverberation component reduction and quadratic function-based response approximation in MUSIC spatial spectrum are utilized to improve the accuracy of DOA estimation. Experimental results show that the proposed method outperforms the well-known generalized cross-correlation (GCC)-based DOA estimation algorithm in the aspect of the estimation error and success rate, respectively. Abstract should be placed here. These instructions give you guidelines for preparing papers for JICCE.

키워드 : 다중 신호 분류, 음원 위치 추정, 입사각, 상호상관도

Key word : multiple signal classification, sound source localization, direction-of-arrival, cross-correlation

접수일자 : 2014. 04. 17 심사완료일자 : 2014. 05. 09 게재확정일자 : 2014. 05. 14

* **Corresponding Author** Youngil Kim(E-mail:yi@gnu.ac.kr, Tel:+82-55-772-1722)

Department of Electronics Engineering/ERI, Gyeongsang National University, Jinju 660-701, Korea

Open Access <http://dx.doi.org/10.6109/jkiice.2014.18.6.1302>

print ISSN: 2234-4772 online ISSN: 2288-4165

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서 론

음원의 입사각 추정 기법은 보안 시스템, 화상회의, 휴머로이드 로봇의 음성기반 위치 추적, 원전 내부에 분포되어 있는 기기 및 구조물에서 발생하는 누설(leakage)을 감시 등 많은 분야에서 사용되어 왔고, 또한 앞으로 음성명령 시스템을 구축한 스마트 차량 등 더 많은 분야에서 사용될 것이다. 음성 신호의 위치를 추정하는 방법으로는 채널 간 시간차 추정법, 공간 스펙트럼 추정 방법인 MUSIC(MULTIPLE SIGNAL CLASSIFICATION) 알고리즘 등이 있다[1-5].

채널 간 시간차를 이용한 입사각 추정은 계산량이 적게 들고 간단한 장점이 있지만 음원의 위치를 계산하기 위하여 비선형 식을 푸는 과정에서의 불안정성, 추정된 채널 간 시간차의 신뢰도가 떨어질 경우에 추정 자체가 불가능해지는 것, 표본 단위의 시간차 추정에 의한 낮은 정밀도, 가장 우세한 단일 음원만의 추정 등의 단점이 있다. MUSIC 알고리즘을 이용하여 DOA를 추정하는 방법은 고유값 분해(EVD: eigen value decomposition)를 필요하기 때문에 계산량을 많이 필요로 하지만 고해상도의 추정이 가능하고 복수의 음원이 존재하더라도 각각의 DOA를 동시에 추정 가능하다는 장점이 있다[1].

기존의 채널 간 시간차 기반 DOA 추정 방법은 시간차 추정 결과의 신뢰성을 높이기 위해서 많은 데이터를 필요로 하는 상호상관도를 이용하는데 반해 잡음과 신호의 수직성을 이용하는 MUSIC 알고리즘은 적은 데이터로도 추정 결과의 신뢰성을 높일 수 있다[1,5]. 제안된 알고리즘은 종래의 MUSIC 스펙트럼 방식에서 2차 함수에 의한 응답 근사화를 이용하여 계산량을 줄였고, 마이크로폰이 존재하는 공간의 특성에 따른 잔향으로 발생하는 반향성분을 제거하여 음원의 DOA 추정 성능을 향상시켰다.

본 논문은 MUSIC 알고리즘을 이용하여 음성신호의 DOA를 추정하는 알고리즘을 제안으로 내용은 다음과 같다. 제 2장에서 이와 관련된 다른 DOA 추정법 및 MUSIC 알고리즘의 이론을 소개한다. 제 3장에서는 제안된 알고리즘의 5단계를 각 단계별로 설명한다. 제 4장에서는 제안된 알고리즘의 최적화 실험 및 실험 결과에 대해서 논하고 마지막으로 제 5장에서 본 논문의 결론을 맺는다.

II. 관련 연구

2.1. 채널 간 시간차를 이용한 DOA 추정법

위치를 측정을 위하여 마이크로폰 간의 시간지연을 이용할 수 있다. 시간지연을 구하는 방법은 여러 가지가 있는데 그 중 상호상관도로 하는 방법이 가장 많이 사용되고 있다. 음원에서 방사된 음파가 두 마이크로폰에서 수신될 때를 식 (1), (2)와 같이 표현될 수 있다.

$$x_1(n) = s(n) + w_1(n) \quad (1)$$

$$x_2(n) = s(n - n_d) + w_2(n) \quad (2)$$

여기서, $x_i(n)$ 는 i 번째 마이크로폰에서 측정된 신호, $s(n)$ 는 음원, $w_i(n)$ 는 i 번째 마이크로폰에서 측정된 잡음, 채널 간 n_d 는 시간지연을 나타낸다. $s(n)$ 는 $w_i(n)$ 와 통계적으로 독립적이라고 가정한다. 측정된 두 신호 $x_1(n)$, $x_2(n)$ 간의 상호상관 값은 식 (3)과 같이 계산할 수 있다[6,7].

$$r_{x_1x_2}(\tau) = \sum_{n=0}^{N-1} x_1(n)x_2(n-\tau) \quad (3)$$

이 때 마이크로폰 사이의 시간지연은 식 (3)에서 $r_{x_1x_2}(\tau)$ 가 최대값을 갖게 하는 표본 단위 τ 가 되고 초 단위의 시간지연은 $\bar{\tau} = \tau/f_s$ 로 표현이 가능하다. 이렇게 구해진 시간지연으로 DOA를 구할 수 있다. 먼저, 원거리 가정(far field assumption), 즉, 음원과 마이크로폰 배열 간의 거리 요소를 배제한 방식을 사용할 때는 음파의 진행 방향의 단위 벡터를 식 (4)와 같이 정의한다[2].

$$\vec{g}_{\theta,\phi} = [-\sin\theta\cos\phi \quad -\sin\theta\sin\phi \quad -\cos\theta]^T \quad (4)$$

여기서, θ 와 ϕ 는 천정각과 방위각을 나타낸다. 그 후, 3차원 좌표 상에서의 각 마이크로폰의 위치를 벡터로 보고 식 (4)와 내적을 취하면 원거리 가정하에서의 채널 간 음원의 지연 거리를 식 (5)와 같이 표현할 수 있다.

$$d_{12} = \vec{g}_{\theta,\phi}^T \vec{p}_1 - \vec{g}_{\theta,\phi}^T \vec{p}_2 = \vec{g}_{\theta,\phi}^T \Delta \vec{p}_{12} \quad (5)$$

마이크로폰의 개수를 M 으로 가정하였을 때 $M(M-1)/2$ 개의 조합에 따른 시간지연을 이용하여 식 (6)과 같은 비용함수를 정의할 수 있다.

$$\xi(\theta, \phi) = \sum_{i < j} (v_s \bar{\tau}_{ij} - d_{ij})^2 \quad (6)$$

여기서, v_s 는 음파의 속도, $\bar{\tau}_{ij}$ 는 i 와 j 번째 마이크로폰에서 측정되는 초 단위의 시간지연을 나타낸다. 식 (6)에 대해서 $\nabla_{\theta,\phi} \xi(\theta, \phi) = 0$ 로 두었을 때 최적의 입사각 \vec{g}_{DOA} 를 식 (7)과 같이 구할 수 있다.

$$\vec{g}_{DOA} = \mathbf{A}^{-1} \vec{b} \quad (7)$$

여기서, $\mathbf{A} = \sum_{i < j} \Delta \vec{p}_{ij} \Delta \vec{p}_{ij}^T$, $\vec{b} = v_s \sum_{i < j} \bar{\tau}_{ij} \Delta \vec{p}_{ij}$ 이다. $\vec{g}_{DOA} = [g_x \ g_y \ g_z]^T$ 로 두었을 때 추정 천정각 및 방위각은 식 (8)과 같다.

$$\theta_{DOA} = \cos^{-1}(-g_z), \quad \phi_{DOA} = \tan^{-1} \frac{g_y}{g_x} \quad (8)$$

2.2. 일반화된 상호상관도 기반의 입사각 추정법

음원의 위치추정 과정 중에 도착시간지연의 정확한 검출은 매우 중요한 의미를 가진다. 주변 잡음 및 음향 반향 등의 장애 요소들의 영향을 최소화하여 가급적 음원 위치 추정 오차를 최소화 할 수 있는 다양한 알고리즘이 제안되고 있다. 그 중, 대표적인 알고리즘으로는 주파수 영역에서 두 신호의 상호상관도(cross-correlation)를 이용하여 추정하는 일반화된 상호상관(GCC: generalized cross-correlation)방법이 있다. GCC 방법은 주파수 영역에서 가중치 함수를 활용하며 식 (9)와 같이 표현할 수 있다[4].

$$r_{x_1 x_2}^G(\tau) = \int_{-\pi}^{\pi} \psi(\omega) \hat{R}_{x_1 x_2}(e^{j\omega}) e^{j\omega\tau} d\omega \quad (9)$$

여기서, $\psi(\omega)$ 는 가중치 함수, $\hat{R}_{x_1 x_2}(e^{j\omega})$ 두 마이크로폰에서 수신된 신호의 상호 전력스펙트럼을 나타낸

다. 식 (9)에 의한 채널 간의 시간차를 추정된 후의 DOA 추정은 식 (4)에서 (8)까지의 과정을 동일하게 적용하여 수행한다.

2.3. MUSIC 스펙트럼 기반의 입사각 추정법

협대역 복소신호를 이용하여 DOA를 추정하는 MUSIC 알고리즘은 독립적인 다중소스의 DOA를 찾는 것이 가능하며 다른 알고리즘보다 높은 추정 해상도를 가지고 있다. MUSIC에 의한 DOA추정 방법은 다음과 같다. 먼저, 존재하는 신호원의 개수를 D개, 신호를 수신하는 센서의 개수를 M ($M > D$)개로 가정하여 다채널 신호의 수신 모델을 식 (10)과 같이 표현한다.

$$\vec{x} = \mathbf{A} \vec{s} + \vec{w} \quad (10)$$

여기서, $\vec{x} = [x_1 \ x_2 \ \dots \ x_M]^T$, $\mathbf{A} = [\vec{a}(\theta_1, \phi_1) \ \vec{a}(\theta_2, \phi_2) \ \dots \ \vec{a}(\theta_D, \phi_D)]$, $\vec{s} = [s_1 \ s_2 \ \dots \ s_D]$, $\vec{w} = [w_1 \ w_2 \ \dots \ w_M]^T$ 를 나타내고 $\vec{a}(\theta_i, \phi_i)$ 는 (θ_i, ϕ_i) 에 존재하는 i 번째 신호원에 의해서 정해지는 배열 응답(array response) 벡터, \vec{s} 는 신호원에서 발생하는 신호 벡터, \vec{w} 는 센서에서 발생하는 잡음 벡터, \vec{x} 는 센서에서 수신되는 신호의 벡터이다. MUSIC 기반의 신호 분석을 위해서 먼저, 신호의 공분산 행렬을 식 (11)과 같이 구한다.

$$\mathbf{R} = E[\vec{x} \vec{x}^H] = \mathbf{A} E[\vec{s} \vec{s}^H] \mathbf{A}^H + E[\vec{w} \vec{w}^H] \quad (11)$$

여기서, $E[\cdot]$ 는 기대치를 λ 는 고유값을 나타낸다. 식 (11)을 사용하여 구한 공분산에 대하여 식 (12)를 만족하는 고유벡터와 고유값을 계산하여 EVD를 해준다.

$$\mathbf{R} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^{-1} \quad (12)$$

여기서, $\mathbf{V} = [\vec{v}_1, \vec{v}_2, \dots, \vec{v}_M]$, $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_M\}$ 이고 \vec{v}_i 와 λ_i 는 고유벡터와 고유값을 나타낸다. 센서의 개수가 신호원보다 많기 때문에 고유값이 상대적으로 작은 잡음의 고유벡터가 존재하여 식 (11)은 식 (13)과 같이 나타낼 수 있다[1].

$$R = AE[\vec{s}s^H]A^H + \lambda_{\min}R_0, \quad \lambda_{\min} \geq 0 \quad (13)$$

여기서, λ_{\min} 는 잡음의 고유값을 나타낸다. 따라서, EVD를 수행한 후 고유값이 가장 작은 것에 해당하는 고유벡터 \vec{v}_{\min} 을 센서 잡음을 나타내는 것으로 간주할 수 있다. 잡음의 고유벡터와 $\vec{a}(\theta_i)$ 가 수직임을 이용하여 MUSIC 스펙트럼을 식 (14)과 같이 정의할 수 있다[1].

$$P_{MUSIC}(\theta, \phi) = \frac{1}{\vec{a}^H(\theta, \phi) \vec{v}_{\min} \vec{v}_{\min}^H \vec{a}(\theta, \phi)} \quad (14)$$

따라서, 식 (14)가 최대를 하는 (θ, ϕ) 를 검출하여 DOA로 추정할 수 있다.

III. 제안된 알고리즘

제안된 알고리즘에서는 MUSIC 알고리즘을 사용하여 신호의 DOA를 검출하기 위해 마이크로폰으로 수집되는 음성신호를 복소신호로 변환하였다. MUSIC 기반의 공간 스펙트럼을 계산 후에 반향 성분에 의하여 발생할 수 있는 피크를 거절해 주었고, 계산속도 및 해상도를 고려하여 2차원-2차 보간 다항식을 세워 해상도를 소수점 단위로 향상시켰다. 그림 1에서 상기의 과정을 나타내었다. 그림 1에서 $\vec{x}(n)$ 는 마이크로폰에서 측정되는 신호, $\vec{x}_a(n)$ 는 복소신호, \vec{v}_{\min} 는 최소 고유치를 갖는 고유벡터를 나타낸다. 각 블록에 대한 자세한 설명은 다음과 같다.

3.1. 복소신호의 생성

실수 값을 갖는 신호로부터 그것의 복소 신호를 생성하는 방법으로는 일반적으로 이산 푸리에 변환이 사용될 수 있으나 본 연구에서는 계산량을 줄이기 위해서 복소 계수를 갖는 필터를 활용하였다. 사용된 복소 대역 통과 필터의 생성개념을 (15)에 나타내었다[6].

$$H_a(z) = H_{LPF}(e^{j\omega_0} z) \quad (15)$$

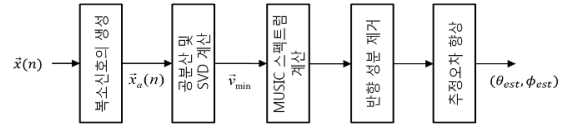


그림 1. 제안된 알고리즘의 구조
Fig. 1 Structure of the proposed algorithm

여기서, ω_0 는 이동하고자 하는 대역의 중심주파수, $H_{LPF}(z)$ 는 실계수를 갖는 저대역 통과 필터이며 그것의 차단 주파수는 생성하고자 하는 복소 대역 통과 필터 대역의 절반이다.

3.2. 공분산 및 EVD 계산

2.3절에 언급한 바와 같이 MUSIC 알고리즘을 사용하기 위해 복소신호 $\vec{x}_a(n)$ 의 공분산을 구한다. 그 공분산의 EVD를 통해 최소의 고유치를 갖는 고유벡터를 음원 혹은 그것의 반사 성분이 아닌 잡음성분에 의한 것으로 간주하여 \vec{v}_{\min} 을 검출한다.

3.3. MUSIC 스펙트럼 계산

벡터의 내적 개념을 이용한 식 (14)을 사용하여 천정 각 θ 와 방위각 ϕ 에 따른 $P_{MUSIC}(\theta, \phi)$ 값을 분석하여 음원 혹은 그것의 반사 성분에 의해 발생하는 피크를 찾는다.

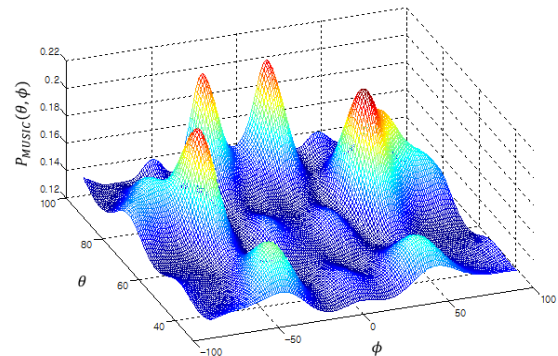


그림 2. MUSIC 기반 공간 스펙트럼에서 반향 성분에 의해 발생한 여러 개의 피크 예시
Fig. 2 Example of multiple peaks by echo components in a MUSIC-based spatial spectrum

3.4. 반향 성분 제거

음성 신호는 벽면이나 여러 물체에 반사되는 반향 성분에 의하여 그림 2에 나타난 바와 같이 MUSIC 기반의 공간 스펙트럼에 여러 개의 피크가 발생된다. 피크 중 가장 큰 점이 찾고자하는 DOA이면 문제가 없지만 반향 성분에 의한 것이 가장 더 클 수 있다. 다만, 음성 신호의 직접 경로에 의한 성분보다 반향에 의한 성분이 그 크기가 상당히 줄어들 것임을 예측할 수 있다. 이러한 개념을 바탕으로 본 연구에서는 검출된 각 피크의 방향으로 MUSIC 알고리즘 계산에 사용된 다채널 입력 신호를 조정(steering)한 후 합한 다음에 에너지가 최대인 것을 방향을 최종 음원의 위치로 선정하였다.

3.5. 추정오차 향상

음원의 위치 추정을 위해서 식 (14)의 공간 스펙트럼은 지정한 정밀도에 의해서 계산될 수 있다. 일반적으로는 정수 단위의 각도 정밀도를 생각할 수 있는데, 정밀도를 2배 높이면 계산량은 4배가 증가하게 되는 단점이 있다. 따라서, 본 연구에서는 식 (16)과 같이 2차원 2차 함수를 활용한 함수의 근사화 기법을 통해서 주어진 추정 정밀도를 높일 수 있는 간단한 방식을 활용하였다.

$$f(\theta, \phi) = \alpha_1\theta^2 + \alpha_2\theta + \alpha_3\phi^2 + \alpha_4\phi + \alpha_5\theta\phi + \alpha_6 \quad (16)$$

식 (16)은 검출된 피크 주변의 6개의 측정된 응답을 사용하여 식 (14)의 MUSIC 스펙트럼을 $f(\theta, \phi)$ 로 근사화 시키는 개념이다. 식 (16)에서 $f(\theta, \phi)$ 가 최대가 될 때 즉, 정밀도가 향상된 천정각 θ 와 방위각 ϕ 는 식 (17)과 같이 구할 수 있다.

$$\begin{bmatrix} \theta_{est} \\ \phi_{est} \end{bmatrix} = \begin{bmatrix} 2\alpha_1 & \alpha_5 \\ \alpha_5 & 2\alpha_3 \end{bmatrix}^{-1} \begin{bmatrix} -\alpha_2 \\ -\alpha_4 \end{bmatrix} \quad (17)$$

IV. 실험 및 결과

4.1. 다채널 DB 수집 및 실험 조건

마이크로폰 배열 중심으로부터 지면상의 거리를 2 m로 두고 천정각 60, 90도 방위각 0, 45, -45도의 위치의 음원을 수집하였다. 마이크로폰의 수는 8 개이고 다채널 마이크로폰은 그림 3과 같이 $p_1 \sim p_8$ 의 위치에 배치

하였다. 각 위치는 $p_1 = [0.27, 0, 0]$, $p_2 = [-0.135, -0.234, 0]$, $p_3 = [-0.135, 0.234, 0]$, $p_4 = [0, 0, 0.27]$, $p_5 = [0.09, 0, 0]$, $p_6 = [-0.045, -0.078, 0]$, $p_7 = [-0.045, 0.078, 0]$, $p_8 = [0, 0, 0.09]$ 이다. 성능 측정을 위해서 음성 신호로 남성과 여성의 발성이 섞인 100개의 한국어 고립어가 사용되었으며 고품질 음향 스피커를 통해서 재생되었다. 모든 데이터는 표본화율 16 kHz, 해상도 16 bit로 저장되었다. 음원 및 장비의 전체적인 배치는 그림 4와 같고 지면에서 녹음 장비와의 수직 높이는 약 1 m로 하였다.

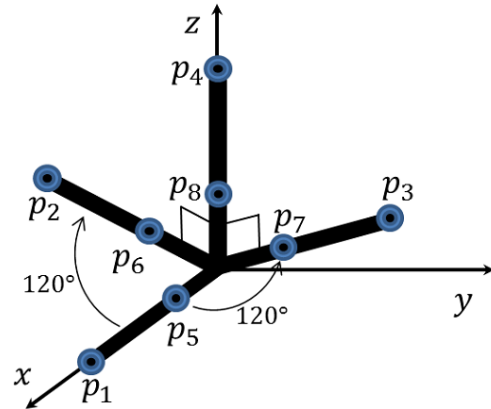


그림 3. 다채널 마이크로폰의 배치
Fig. 3 Placement of multi-channel microphones

4.2. 제안된 방식에서의 복소 필터 최적화

제안된 방식의 그림 1에서 복소 신호 생성을 위한 필터의 중심주파수 w_0 및 대역폭은 실험적으로 최적 조건을 검출하였다. 이를 위해서 먼저, 천정각 θ , 방위각 ϕ 의 오차가 20° 미만일 때를 추정 성공으로 간주하였으며 중심 주파수 및 대역폭을 가변하면서 식 (18)을 적용하여 평균 추정 오차를 계산하였다.

$$e_{avr} = \sqrt{\frac{\sum_{i=1}^N \{(\theta_{est}^{(i)} - \theta_{ans}^{(i)})^2 + (\phi_{est}^{(i)} - \phi_{ans}^{(i)})^2\}}{N}} \quad (18)$$

여기서, N 은 실험 데이터 중 허용오차 범위 내의 개수, $\theta_{est}^{(i)}$, $\phi_{est}^{(i)}$ 는 i 번째 데이터의 추정 각도, $\theta_{ans}^{(i)}$, $\phi_{ans}^{(i)}$ 는 i 번째 데이터의 실제 각도를 나타낸다.



그림 4. 다채널 DB 수집 환경 (방위각 0도, 천정각 90도 예시)
 Fig. 4 Environment of multi-channel DB collection (Examples of azimuth 0°, elevation 90°)

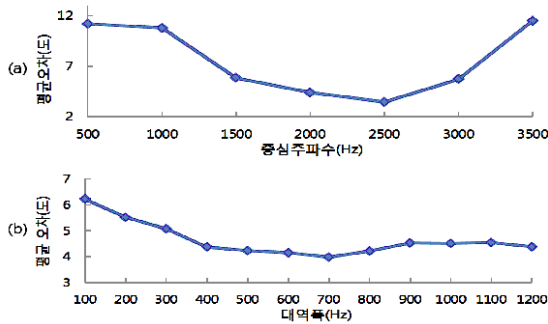


그림 5. 중심주파수 및 대역폭에 따른 평균 추정오차
 Fig. 5 Average estimation error according to center frequency and bandwidth

실험을 위해서 사용된 식 (15)의 $H_{LPF}(z)$ 는 통과 대역의 리플이 0.1 dB, 차단 대역의 감쇄가 40 dB인 6차의 엘립틱(elliptic) 필터였다[6]. 먼저, 그림 5(a)는 대역폭이 700 Hz일 때 500 Hz간격으로 중심주파수 변화에 따른 추정에 성공한 추정 각도의 평균오차를 나타내고 최적의 중심주파수는 2.5 kHz인 것을 알 수 있다. 그림 5(b)는 중심주파수가 2.5 kHz일 때 200 Hz간격으로 대역폭 변화에 따른 추정에 성공한 추정 각도의 평균오차를 나타내고 최적의 대역폭은 700 Hz인 것을 알 수 있다.

4.3. 성능 비교

제안된 알고리즘과 종래의 GCC 기반의 입사각 추정법의 성능이 비교되었다. 종래의 GCC 기반의 입사각 추정법을 위해서 PHAT(Phase Transform) 가중치 함수를 적용하였다.

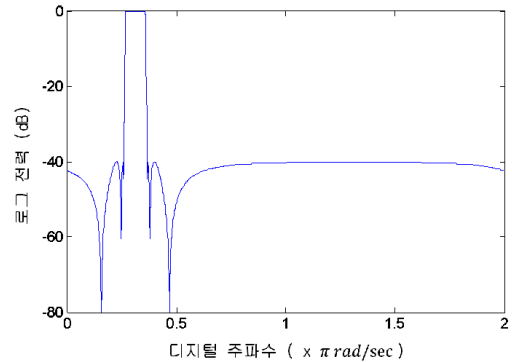


그림 6. 복소신호의 생성에 사용된 복소 계수 대역 통과 필터의 주파수 응답

Fig. 6 Frequency response of the band pass filter with complex coefficients to generate complex signals

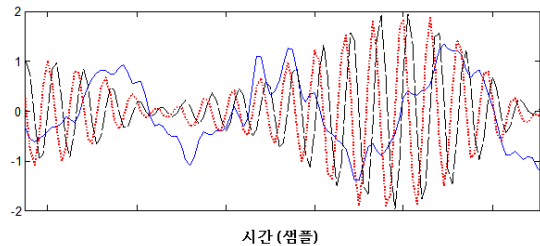


그림 7. 입력된 음성신호(실선) 및 생성된 복소 신호 예시(점선: 실수부, 파선: 허수부, 틱간 간격은 30 샘플)

Fig. 7 Example of Input speech signal(solid) and produced complex signal (dotted: real part, broken: imaginary part, 30 samples for each tick)

또한, 채널 간 시간차 추정이 시행되는 신호의 구간은 에너지가 가장 큰 부분을 검출하여 설정하였고 사용된 이산 푸리에 변환의 크기는 1024였다. 제안된 알고리즘에서는 동등한 조건에서의 성능 비교가 이루어지도록 하기 위해서 GCC에서 사용한 동일한 구간에 대해서 식 (11)의 공분산 행렬을 구하고 EVD 분석을 수행하였다. 복소 신호의 생성은 4.2절에서 얻은 최적 실험결과를 바탕으로 대역폭 700 Hz와 중심주파수 2.5 kHz 인 6차 복소 계수 엘립틱 필터를 사용하였다. 그림 6에 그것의 주파수 응답을 나타내었으며 그림 7에 그것에 의해 생성되는 복소 신호의 예를 나타내었다. 식 (14)의 MUSIC 기반 공간 스펙트럼은 천정각 θ 및 방위각 ϕ 에 대해서 1 도 단위로 계산되었다. 그림 8에서 제안된 알고리즘과 GCC 기반의 입사각 추정법의 성능 측정 결과를 나타내었다.

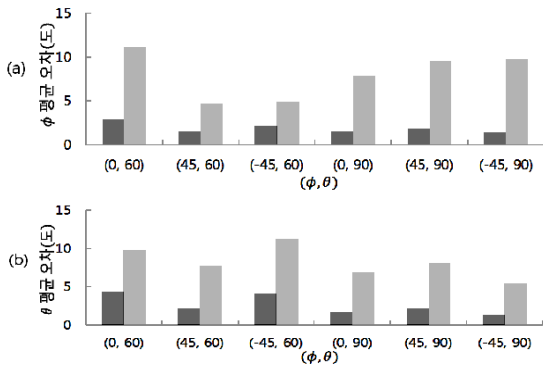


그림 8. 제안된 알고리즘과 GCC 기반의 입사각 추정법과의 비교. 검정색 및 회색은 제안된 알고리즘 및 GCC 기반의 입사각 추정법 (a) 음원의 위치에 따른 방위각 추정 오차의 평균, (b) 음원의 위치에 따른 천정각 오차의 평균

Fig. 8 Comparative performance of the proposed and the GCC-based DOA estimation algorithm. (a) Average azimuth estimation error according to sound source locations, (b) Average elevation estimation error according to sound source locations

그림 8에서 알 수 있듯이 제안된 입사각 추정 알고리즘이 모든 조건에서 전반적으로 더 낮은 성능을 나타내었다. 제안된 방식은 θ 및 ϕ 추정에 있어서 종래의 GCC 기법에 비해서 평균적으로 4.79°, 6.92° 만큼의 더 낮은 추정오차를 나타내었다. 상기의 성능치는 4.2절에서 언급한 바와 같이 추정 오차가 20° 미만인 추정 성공의 경우 대해서만 구한 것이다. 추가적으로 추정 성공률에 대한 성능은 제안된 방식은 95.33% 인데 반해 GCC 방식은 9.18%에 불과하였다. 즉, 실환경에서의 성능 측정 결과 종래의 GCC 방식은 추정 오차보다 추정 성공률에서 매우 불리함을 지적할 수 있다. 이러한 결과는 종래의 GCC 방식이 제안된 방식과는 달리 반향 성분 등에서 오는 성능 저하를 대처할 수 없기 때문에 얻어진 것으로 판단된다.

V. 결 론

본 논문에서는 MUSIC 알고리즘을 기반으로 반향성분 거절, 2차 함수 기반 DOA 추정 정밀도 향상에 관한 방법을 제안하였다. 실환경 조건에서의 실험결과 제안한 방법의 평균 추정 오차는 GCC 기반의 입사각 추정

법에 비해 천정각 4.79°, 방위각 6.92° 감소하였고 추정 성공률은 월등히 향상됨을 알 수 있었다. 향후, 제안된 DOA 추정 알고리즘을 기반으로 복수의 음원이 존재할 때 검출된 피크 방향으로 빔포밍을 수행한 후 그 결과에 대한 상호상관도를 측정하여 반향성분의 DOA를 거절할 수 있는 알고리즘을 구현할 예정이다.

감사의 글

이 논문은 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업지원을 받아 수행된 것임(2011-0011800).

REFERENCES

- [1] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antenna and Propagation*, vol. AP-34, no. 3, pp. 276-280, 1986.
- [2] I. Frire and J. Apolinario, "DoA of gunshot signals in a spatial microphone array: Performance of the interpolated Generalized Cross-Correlation method," *Argentine School of Micro-Nanoelectronics Technology and Applications (EAMTA)*, pp. 1-6, 2011.
- [3] S. Zhao, S. Ahmed, Y. Liang, D. Chen, and D. Jones, "A real-time 3D sound localization system with miniature microphone array for virtual reality," *7th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pp. 1853-1857, 2012.
- [4] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-24, no. 4, pp. 320-327, 1976.
- [5] J. Benesity, J. Chen, and Y. Huang, *Microphone Array signal processing*, Springer-Verlag, 2008.
- [6] A. Oppenheim and R. Shafer, *Discrete-time signal processing*, 3rd ed., Pearson, 2009.
- [7] M. Hayes, *Statistical digital signal processing and modeling*, John Wiley and Sons Inc., 1996.



장형욱(Hyungwook Chang)

2013년 경상대학교 전자공학과(학사)
2013년 - 현재 경상대학교 전자공학과(석사과정)
※관심분야 : 음성신호처리, real-time operating system



정상배(Sangbae Jeong)

1997년 부산대학교 전자공학과(학사)
1999년 한국과학기술원 전기 및 전자공학과(석사)
2002년 한국과학기술원 정보통신공학과(박사)
2002년 - 2006년 삼성종합기술원 컴퓨팅랩(책임연구원)
2006년 - 2009년 한국과학기술원 디지털미디어랩(연구조교수)
2009년 - 현재 경상대학교 전자공학과/공학연구원(부교수)
※관심분야 : 음성신호처리, 음성오디오 부호화



김영일(Youngil Kim)

1979년 경북대학교 전자공학과(학사)
1981년 연세대학교 전자공학과(석사)
1985년 연세대학교 전자공학과(박사)
1987년 - 현재 경상대학교 전자공학과/공학연구원(교수)
※관심분야 : 음향공학, 음성신호처리