

# 음원 위치 추정용 로봇 인공귀를 위한 마이크로폰 배치법

## Microphone Arrangement for Robot Artificial Ear

이상문† · 박영진\* · 박윤식\*\*

Sangmoon Lee, Youngjin Park, and Youn-sik Park

### 1. 서론

음원 위치 추정이란 마이크로폰 어레이를 통해 측정된 음향 신호를 이용하여 음원의 위치를 추정하는 기술을 일컫는다. 이러한 위치 추정 기술은 군사 및 감시 시스템, 기능형 로봇과 화상회의 등 다양한 분야에 적용된다. 특히 로봇에 적용되는 음원 위치 추정 기술은 로봇의 사용자와 로봇 간의 상호작용을 위한 기반 기술로 로봇의 형태에 따라 다양한 마이크로폰의 배치 및 이에 따른 위치 추정 알고리즘이 적용된다. 소음이 큰 환경에서는 시각 정보에 기반한 사용자의 입술 움직임(lip-reading) 혹은 얼굴 인식(face-detection)을 통하여 발화하는 화자의 위치를 추정할 수 있으나, 제한된 FOV(field of view)를 벗어나 화자가 존재하는 경우, 로봇이 위치한 공간의 조도가 현저히 떨어지는 경우에는 인식이 불가능하다. 따라서 시각 정보와 음성 신호를 같이 이용한 음원 위치 추정 방식이 사용된다. 본 연구에서는 음향 신호만을 이용한 위치 추정 기술에 초점을 둔다.

일반적으로 로봇에 적용되는 음원 위치 추정 방식은 화자의 음성이 마이크로폰에 도달하는 시간 지연차이(Time Delay of Arrivals, TDOA)를 이용한 방식 또는 다수의 마이크로폰을 이용한 빔포밍(beam-forming)방식이 주로 사용되고 있다. 그 외에도 로봇 플랫폼에 의한 음장의 변화를 나타내는 머리전달 함수(Head-related Transfer Function, HRTF)를 이용한 방법이 있다.

본 연구는 미래형 인간형 로봇에 적용 가능한 음원 위치 추정 센서를 개발하는 것으로 보다 자연스러운 화자와 로봇의 상호작용을 위해서 인간형 로봇 산업은 매우 다양한 하드웨어 제약을 요구한다. 사람과 닮은 센서 개발을 위하여 사람의 외이를 모방한 센서를 요구하며 부착 가능한 마이크로폰의 위치를 사람의 귀의 위치로 국한한다. 또한 사람의 음성 신호 추정을 위하여 최대 샘플링주파수는 16kHz로 제한하며 실시간 위치 추정을 요구한다.

이러한 로봇 산업의 요구에 부합하여 로봇 인공귀 기술이 제안되었다. 본 논문에서는 기존의 로봇 인공귀에 사용된 알고리즘 및 이에 사용된 센서의 위치와 달리 Gwon 이 제안한 공간상에 사상된 GCC (Generalized Cross Correlation) 함수를 로봇 인공귀에 적용하였을 때의 결과를 실제 환경에서의 실험을 통하여 살펴보았다.

### 2. 본론

#### 2.1 공간좌표로 사상된 GCC 함수를 이용한 음원 위치 추정 방법

일반적으로 TDOA 를 이용한 음원 위치 추정방법은 각 마이크로폰 쌍에 의해 측정된 GCC 함수를 이용하여 GCC 값이 최대가 되는 시간지연차이를 추정한다. 이는 다음 식 (1)에 의해 추정이 가능하다.

$$\tau_i = \max_{\tau} R_i(\tau) \tag{1}$$

$R_i(\tau)$  는  $i$  번째 마이크로폰 쌍의 신호를 이용한 cross correlation 함수를 나타낸다. 이 때 측정된  $\tau_i$  는 마이크로폰의 기형학적 배치에 따른 사상함수  $\Phi_i(\tau)$  를 이용하여 음원의 위치를 추정할 수 있다. 이는 식 (2)에 의해 표현이 가능하다.

$$\theta_i = \Phi_i(\tau_i) \tag{2}$$

이렇게 각 마이크로폰 쌍에 의해 추정된 음원의 위치( $\theta_i$ )는 다음 식 (3) 에 의해 최종적으로 음원의 위치를 추정한다.

$$\theta_{est} = \frac{1}{N} \sum_{i=1}^N |\theta_i - \Phi_i(\theta)| \tag{3}$$

$N$  은 사용한 마이크로폰 쌍의 수를 나타낸다. 이에 Kwon 은 측정된  $R_i(\tau)$  를 space domain 으로 변환하고 변환된 값들을 합하여 최종적인 음원의 위치를 추정하는 알고리즘을 사용하였다. 이는 다음 식 (4-6)에 의해 표현된다.

$$mGCC_i(\theta) = \Phi(R_i(\tau)) \tag{4}$$

† 교신저자 : 이상문, KAIST 기계공학과

E-mail : smansl@kaist.ac.kr

Tel : (042) 350-3076, Fax : (042) 350-8220

정희원, KAIST 기계공학과

$$sGCC_i(\theta) = \max_{\theta} \sum_{i=1}^N mGCC_i(\theta) \quad (5)$$

$$\theta_{est} = \max_{\theta} sGCC_i(\theta) \quad (6)$$

Kwon 이 제안한 방법을 사용하기 위해서는 기본 로봇 인공귀에 배치된 마이크론의 배치 구조를 사면체 형태로 수정할 필요성이 있다.



그림 1. (좌) 기존의 로봇 인공귀에 부착된 마이크론의 배치 (우) 변경된 마이크론의 배치

## 2.2 구형머리전달함수와 실험 데이터를 이용한 사상함수

Gwon 이 제안한 음원 위치 추정 알고리즘에는 추정된  $R_i(\tau)$  를 공간상으로 사상하기 위한  $\Phi_i(\tau)$  함수가 필요하다. 이를 결정하는 방법에는 이론적인 구형머리전달함수를 이용한 방법과 실험 데이터를 통해 추정하는 방법을 생각 할 수 있다.

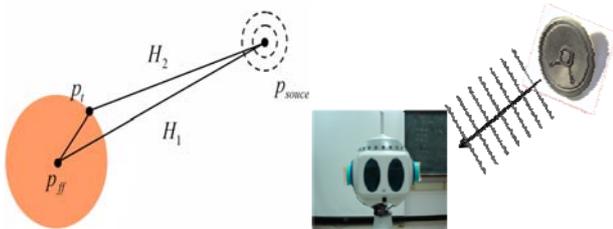


그림 2. (좌) 구형머리전달함수를 이용한 사상함수 (우) 실험 데이터를 이용한 사상함수

## 2.3 실험에서의 음원 위치 추정 결과

제안된 알고리즘을 로봇 인공귀에 적용하였을 때의 음원 위치 추정 성능을 살펴보기 위하여 본 연구에서는 vertical-polar 좌표를 이용한다.

음원이 (수평각, 고도각)=(350°, 20°), (100°, 10°)에 위치하였을 때의 각  $sGCC_i(\theta_{ele}, \theta_{azi})$  를 살펴보고 하자. 그림 3은 구형머리전달함수에 기반한 사상함수를 사용하였을 때의 음원 위치 추정 결과를 나타낸다.

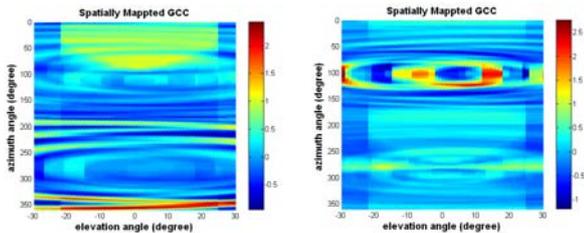


그림 3. 음원이 (수평각, 고도각)=(350°, 20°)일 때의 sGCC(좌), (100°, 10°) 일 때의 sGCC(우)

그림 4는 실험 데이터에 기반한 사상함수를 사용하였을 때의 음원 위치 추정 결과를 나타낸다.

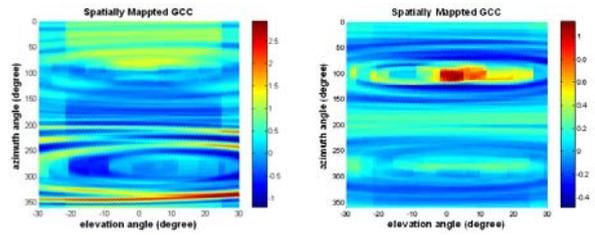


그림 4. 음원이 (수평각, 고도각)=(350°, 20°)일 때의 sGCC(좌), (100°, 10°) 일 때의 sGCC(우)

## 3. 결론

본 연구에서는 기존의 로봇 인공귀에 Gwon 이 제안한 공간상에 사상된 GCC 함수를 이용한 위치 추정 방법을 적용하였고 이를 위하여 마이크론 어레이를 사면체구조로 배치하였다. 구형머리전달함수와 실험 데이터에 기반한 사상함수를 이용하였을 때의 위치 추정 성능은 다음과 같다.

표 1. 음원 위치 추정 성능

mean(std)	수평각 추정오차	고도각 추정오차
구형머리전달함수를 이용하였을 때	8.3°(6.2°)	19.3°(12.5°)
실험 데이터를 이용하였을 때	7.3°(5.2°)	14.3°(8.5°)

실험데이터에 기반한 사상함수가 로봇 인공귀의 외형에 의한 음장을 구형머리전달함수 기반의 사상함수보다 정확히 표현하기 때문에 상대적으로 높은 성능을 보여주는 것을 알 수 있다. 기존의 로봇 인공귀가 가진 위치 추정 방식은 Kwon 이 제안한 방법을 이용하였을 때 보다 높은 성능을 가지지만 단일 음원이 존재할 때에만 적용이 가능하다는 한계점이 있다. 환경의 변화에 따라 서로 다른 위치 추정 방식을 선택할 수 있을 것으로 예상된다.

## 후 기

본 연구는 두뇌 한국 21 프로젝트와 인간기능 생활지원 지능로봇 기술개발 사업단의 지원으로 수행되었음.