

# 로봇 시스템에 적용될 음원 위치 추정 방법 Sound Source Localization Method Applied to Robot System

권병호† · 박영진\* · 박윤식\*\*

Byoung-ho Kwon, Youngjin Park and Youn-sik Park

**Key Words :** Sound Source Localization(음원 위치 추정), TDOA(Time Difference of Arrival, 도달시간지연), Compensation Method(보상법), SHRTF(Spherical Head Related Transfer Function, 구형머리전달함수)

## ABSTRACT

While various methods for sound source localization have been developed, most of them utilize on the time difference of arrival (TDOA) between microphones or the measured head related transfer functions (HRTF). In case of a real robot implementation, the former has a merit of light computation load to estimate the sound direction but can not consider the effect of platform on TDOAs, while the latter can, because characteristics of robot platform are included in HRTF. However, the latter needs large resources for the HRTF database of a specific robot platform. We propose the compensation method which has the light computation load while the effect of platform on TDOA can be taken into account. The proposed method is used with spherical head related transfer function (SHRTF) on the assumption that robot platform, for example a robot head, installed microphones can be modeled to a sphere. We verify that the proposed method decreases the estimation error caused by the robot platform through the simulation and experiment in real environment.

## 1. 서론

근래에 로봇에 대한 연구가 활발히 진행되고 있는 가운데 사람과 로봇 사이의 원활한 상호작용을 위하여 비전 센서(vision sensor)를 이용한 시각 정보뿐만 아니라 청각 정보를 활용하고자 하는 연구도 많이 진행되고 있다[1]. 그 중에서 화자(speaker)의 음성이 발생한 위치를 알아내서 로봇의 움직임에 활용하고자 하는 음원 위치 추정(sound source localization) 방법에 관한 연구도 진행되고 있다[1~3]. 이런 음원 위치 추정 방법이 실제 로봇 시스템에 신호처리를 위한 SoC(System on a Chip)의 한 모듈로 구현되기 위해서는 계산량이 적어야 되고, 보유하고 있어야 하는 데이터의 양 또한 적어야 한다. 뿐만 아니라 SoC 로 구현된 이후에는 다양한 로봇 시스템에 쉽게 적용될 수 있어야 한다. 따라서 본 연구에서는 이런 목적에 부합될 수 있는 음원 위치 추정 방법을 고안하고자 한다.

기존의 마이크로폰 사이의 도달시간지연(TDOA)이나 빔포밍(Beamforming)을 이용한 음원 위치 추정 방법은 다수의 마이크로폰 어레이를 이용하여

화상회의(videoconferencing)에서 화자의 위치로 카메라가 움직이게 하기 위해 사용되었다[4]. 이를 실제 로봇 시스템에 적용시킬 경우, 마이크로폰이 설치될 로봇 플랫폼의 영향을 고려할 수 없고 너무 많은 마이크로폰이 사용된다는 문제점들이 발생되었다[3]. 또 로봇의 머리전달함수(Head Related Transfer Function, HRTF)를 이용한 음원 위치 추정 방법은 소수의 마이크로폰을 이용하여 3 차원 공간상에 존재하는 음원의 위치를 추정할 수 있다는 장점이 있지만, 특정한 로봇 플랫폼에 대한 머리전달함수 측정의 어려움과 이를 데이터베이스화하고 있어야 하기 때문에 많은 저장 공간이 필요하다[2]. 이런 이유로 SoC 의 모듈로 구현하기에는 어려움이 있다.

따라서 본 연구에서는 소수의 마이크로폰을 이용하여 음원의 위치를 추정하기 위해서 도달시간지연을 이용한 음원 위치 추정 방법을 적용하면서, 로봇 플랫폼의 영향을 고려할 수 있는 방법을 제안하고자 한다. 본 연구에서는 음성신호를 이용해 음원의 위치를 추정하려고 하기 때문에 그에 해당하는 파장(wavelength)을 고려하면 마이크로폰이 설치되는 로봇 플랫폼의 크기 변화가 20cm 내외일 때 이를 구(sphere)로 모델링할 수 있다. 본 연구에서는 구로 모델링 가능한 플랫폼에 대해서 이를 구로 모델링하고 구형머리전달함수(SHRTF)의 해석해를 이용해 플랫폼의 영향을 계산해 마이크로폰 사이의 도달시간지연을 추정할 때 그 영향을 보상해주고자 한다.

† 권병호; 한국과학기술원  
E-mail : bhkown@kaist.ac.kr  
Tel. : (042) 869-3056, Fax : (042) 869-8220

\* 한국과학기술원

\*\* 한국과학기술원

## 2. TDOA 를 이용한 음원위치추정 방법

본 장에서는 음원의 위치를 추정하기 위해 기본적으로 적용되는 도달시간지연을 이용한 음원 위치 추정 방법에 대해서 알아본다. 이 방법은 마이크로폰 사이의 시간지연과 마이크로폰이 놓여 있는 기하학적 형상만을 이용하여 간단하게 음원의 위치를 추정할 수 있는 방법이다[5].

Fig. 1 과 같이 두 개의 마이크로폰이 자유음장 (free-field) 공간에 존재할 경우, 임의의 방향,  $\phi$  에서 음원이 발생하게 되면 마이크로폰 사이에는  $d$  거리만큼의 시간 지연이 발생하게 된다. 여기에서  $d$  거리만큼의 시간 지연,  $\tau$  을 알 수 있으면 식 (1), (2)을 이용하여 음원의 위치를 알 수 있게 된다.

$$\tau = \frac{d}{c} = \frac{2r \sin \phi}{c}, \quad \tau : \text{TDOA} \quad (1)$$

$$\phi = \sin^{-1} \frac{\tau c}{2r} \quad (\tau c \leq 2r) \quad c : \text{wave speed} \quad (2)$$

이 때 두 마이크로폰 사이의 시간 지연은 마이크로폰에서 측정된 신호를 이용하여 추정하게 되는데, 본 연구에서는 두 신호의 시간 지연 추정에 있어서 가장 널리 사용되고 있는 Generalized Cross Correlation 방법 중에 PHAT(phase transform) 가중치 함수를 적용한 방법을 이용한다[6].

## 3. 구형머리전달함수를 이용한 보상법

인간과 상호작용을 위해 제작된 로봇들은 인간과 비슷한 형상을 하고 있는데 그 중에서 청각 기능을 담당하는 센서는 로봇의 머리에 해당하는 부분에 설치되는 것이 일반적이다. 이런 로봇 플랫폼을 음성 신호의 파장(wavelength)을 고려할 때, 구(sphere)로 모델링 할 수 있고, 구로 모델링된 플

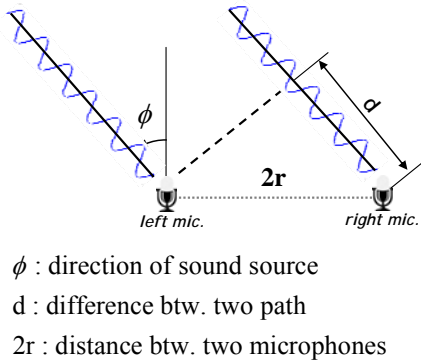


Fig.1 Sound Source Localization method using TDOA in free-field condition

랫폼의 영향을 고려하기 위해서 본 연구에서는 구형머리전달함수(SHRTF)를 이용하고자 한다.

### 3.1 구형머리전달함수(SHRTF)

사람의 머리를 딱딱하고 일정 반경을 가지는 완전 구체로 가정하고 임의의 점음원을 고려하였을 때, 주파수 영역의 이론적인 머리전달함수를 구형머리전달함수(SHRTF)한다. 이에 대한 해석하는 [7]에 자세하게 소개되어 있다. Fig. 2 에서와 같이 주기적인 소리를 발생시키는 공간상의 이상적인 점음원(ideal point source)을 정의하고 그로부터 발생하는 소리의 방사량을  $S_{\omega} e^{-i\omega t}$  라 할 때, 음원의 중심에서 미소거리  $r_{\epsilon}$  만큼 떨어진 지점의 음압은 다음과 같다.

$$p_{ff}(r, \omega, t) = -i\omega \frac{\rho_0 S_{\omega}}{4\pi r} e^{i(kr - \omega t)}, \quad (3)$$

구 표면에서의 압력,  $p_s$  은 식 (4)와 같이 표현된다.

$$p_s(r, a, \omega, t) = -\frac{i\rho_0 c S_{\omega}}{4\pi a^2} \Psi e^{-i\omega t}, \quad (4)$$

그리고 구형머리전달함수,  $H$  는 다음과 같이 정의된다.

$$H(r, a, \omega, \phi) = \frac{H_2}{H_1} = \frac{p_{\epsilon} p_s}{p_{ff} p_{\epsilon}} = \frac{p_s}{p_{ff}} = -\frac{cr}{\omega a^2} \Psi e^{-i\omega \left(\frac{r}{c}\right)}, \quad (5)$$

$$\text{where, } \Psi(\rho, a, \theta) = \sum_{m=0}^{\infty} (2m+1) P_m(\cos \theta) \frac{h'_m(ka\rho)}{h'_m(ka)}$$

식 (5)로 표현되는 구형머리전달함수의 해석해를 이용하여 시간 지연을 추정함에 있어 구로 모델링된 로봇 플랫폼의 영향을 고려해주고자 한다.

### 3.2 보상법 (Compensation method)

Fig. 2 에서와 같이 구형 플랫폼에 센서가 설치되어 있는 경우, 각 센서들 사이의 시간지연으로부터 음원의 위치를 추정할 때 자유음장에서의 관계식인 식 (2)를 이용하면 위치 추정 오차가 발생하게 된다. 그러므로 플랫폼에 의해 추가적으로 발생하는 시간지연을 추정하고 이를 보상해서 음원의 위치를 추정해야 한다. 본 연구에서는 구형머리전달함수를 이용한 보상법을 제안한다.

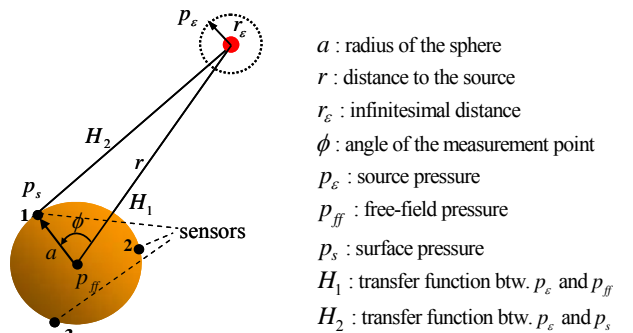


Fig. 2 Spherical Head Related Transfer Function

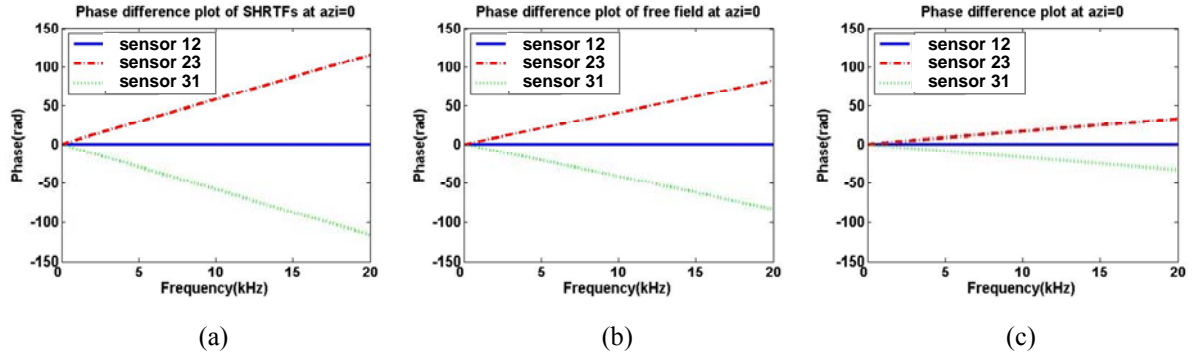


Fig. 3 Phase difference between sensors when source is located at  $0^\circ$ , (a) case of the presence of sphere, (b) case of the free-field, (c) difference between (a) and (b)

Fig. 3 (a)는 구형머리전달함수를 이용하여 플랫폼에 위치하는 센서 사이의 위상 차(phase difference)를 주파수 별로 계산한 값이다. Fig. 3 (b)는 자유음장 공간에서 같은 위치에 있는 센서들 사이의 위상 차를 나타낸다. 이 둘의 차가 Fig. 3 (c)에 나타나 있는데, 이는 로봇 플랫폼에 의해 추가적으로 발생하는 위상 차가 된다. 이로부터 다음의 식 (6)을 이용하여 센서들 사이에 보상해 줘야 하는 시간지연 값을 추정할 수 있다.

$$\text{Time Delay} = \frac{1}{2\pi} \left| \frac{d}{df} (\text{Phase of SHRTF}_{\text{sensor1}} - \text{Phase of SHRTF}_{\text{sensor2}}) \right| \quad (6)$$

위에서 구해진 추가적인 시간지연을 고려한 수정된 시간지연과 음원의 위치 사이의 사상함수(mapping function)를 이용하여 오차 없이 음원의 위치를 추정할 수 있다.

### 3.3 보상 값 선정 방법

플랫폼의 영향을 고려하기 위해서는 음원의 위치를 정확하게 알고 있어야 한다. 하지만 실제로는 음원의 위치를 알 수 없기 때문에 각 센서 사이에 보상해줘야 하는 시간지연 값을 결정해야 한다. 이를 위해 본 연구에서는 모든 방향에 대한 각 센서들 사이의 추가적인 시간지연 값을 데이터베이스화하고, 측정된 신호로부터 계산된 상호상관계수 값에 이를 보상했을 때, 보상된 각도에서 실제 음원의 위치를 나타내는 상호상관계수 값이 가장 큰 조합을 찾아내 그 값들로부터 실제 음원의 위치를 추정한다. 이는 플랫폼의 영향이 제대로 반영되었을 경우 실제 음원의 위치를 나타내는 상호상관계수 값이 가장 큰 것이라는 가정하에 이루어지는데 모의실험과 실제 환경에서의 실험에서 검증되었다.

## 4. 모의 실험

모의 실험은 반지름 0.15m의 구형 플랫폼에 Fig. 2와 같이 세 개의 센서가  $120^\circ$  간격으로 설치되어 있을 때, 이들을 이용하여 음원의 위치를 추정함에 있어 보상법을 적용 유무에 따른 성능 비교를 위해 수행되었다. 이를 위해 구형머리전달함수의 시간영역인 SHRIR (Spherical Head Related Impulse Response)를 이용하여 각 위치에서 측정될 신호들을 생성하고 이를 이용해  $0^\circ \sim 180^\circ$  까지  $10^\circ$ 간격으로 음원의 위치를 추정하게 하였다. 이때 사용된 음성 신호는 “안녕하세요”, “권병호” 두 단어가 사용되었으며 그에 대한 결과는 Fig. 4에 나타내었다. 보상법을 적용한 경우에는 모든 경우에서 정확하게 음원의 위치를 추정하고 있지만 그렇지 않은 경우에는 앞 뒤 혼동 (front-back confusion) 현상과 또 다른 추정 오차가 발생함을 알 수 있다.

## 5. 실제 환경에서의 실험

모의 실험을 통해 플랫폼에 설치된 센서들을 이용하여 음원의 위치를 추정할 때 플랫폼의 영향을 고려해야 함을 알았다. 이를 실제 로봇 플랫폼을 이용하여 실제 사용환경에서의 실험을 통해 검증하였다. 사용된 플랫폼은 Fig. 5 (a)와 같고, 이를 구로 모델링한 것이 (b)와 같다. R은 구의 반지름, d는 구 중심에서 센서가 존재하는 면까지의 거리, r은 센서가 존재하는 원의 반지름,  $\theta$ 는 음원의 수평각이고,  $\Psi$ 는 음원과 센서의 위치벡터 사이의 사잇각을 나타낸다. 이들 값들을 이용하여 구형머리전달함수의 해석해를 이용해 각 센서들 사이에 보상해 줘야 하는 시간지연 값을 계산하였다. 실험환경과 실험에 적용된 조건들은 각각 Fig. 6

과 table 1 에 나타내었다.

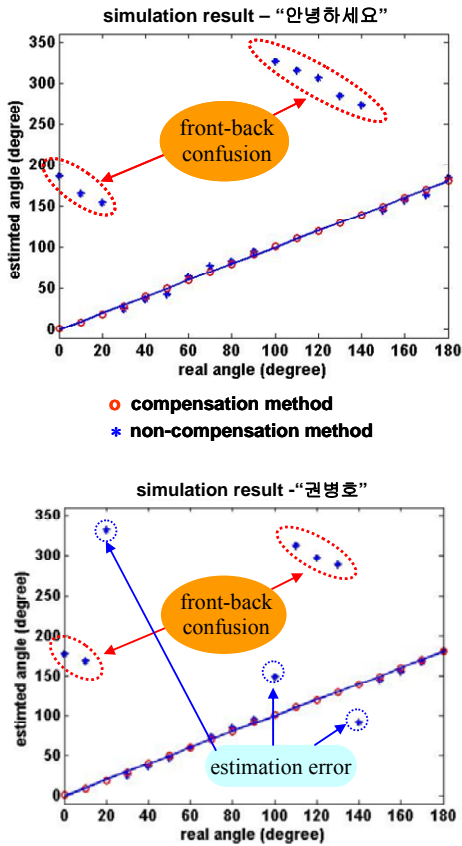


Fig. 4 simulation result

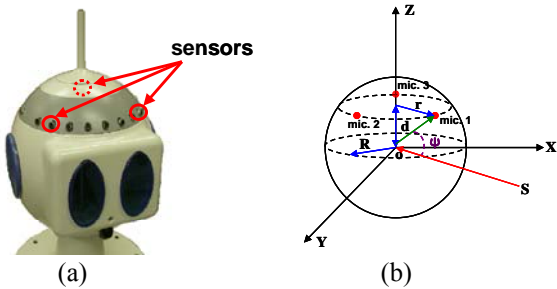


Fig. 5 (a) Robot platform; Infotainment Robot Platform Version 1, (b) modeling to sphere



Fig. 6 Experiment in real environment

Table. 1 Experiment condition

Radius of sphere model, R	0.146m
Offset distance, d	0.09m
Distance btw. sensors	0.20m
Distance btw. Speaker and robot	3m
Speech signal	“안녕하세요”
Background noise power	37dB
SNR	32dB, 26dB, 20dB
Sampling frequency	16000Hz
Measurement point and times	0° ~ 330° as each 30° and 5 times

위와 같은 조건하에서 SNR 별 실험결과는 Fig. 7 과 같다. 추정된 음원의 위치를  $\pm 5^\circ$  범위 내의 대표 값으로 표현하였으면 마커(marker)의 크기는 추정한 횟수와 비례하도록 하였다. 위 결과에서 보면, 보상법을 적용한 경우에는 모

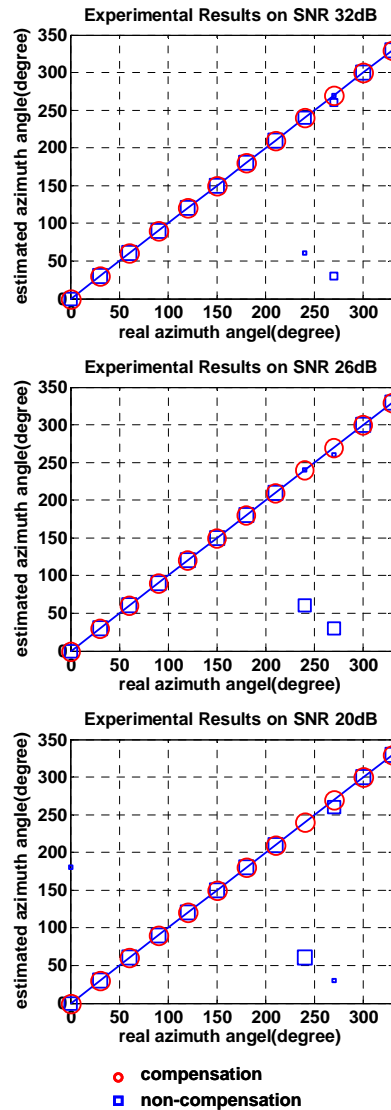


Fig. 7 Experimental results as SNR

든 각도에서 오차  $\pm 5^\circ$  내에서 음원의 위치를 추정하는 반면, 그렇지 않은 경우에는 앞 뒤 혼동 현상이 발생하는 것을 알 수 있다. 또 SNR 이 낮아질수록 그 성능이 나빠지는 것도 알 수 있다. 여기에서 음원의 추정 오차를 정량적으로 나타내지 않고 있는데 이는 보상법 유무에 따른 음원 위치 추정 성능에 차이만을 보여주기 위함이다.

## 6. 결 론

본 연구에서는 로봇 시스템에 적용될 음원 위치 추정 방법 중에서 소수의 마이크로폰을 사용하면서 적은 계산 량과 데이터만으로 플랫폼의 영향을 고려할 수 있는 방법을 제안하였다. 시간지연을 이용하여 음원의 위치를 추정함에 있어서 플랫폼의 영향을 보상해주는 보상법을 제안하였고, 플랫폼이 구(sphere)로 모델링 될 수 있는 경우에는 구형머리전달함수를 이용하여 보상하는 방법에 대해 설명하였다. 이를 모의 실험과 실제 환경에서의 실험으로 검증하였는데, 플랫폼의 영향을 고려한 경우에는 앞 뒤 혼동현상 없이 음원의 위치를 정확하게 추정할 수 있음을 알 수 있었다. 본 연구에서는 음원이 평면상에 존재하는 경우에 대해서만 그 성능을 검증하였는데, 앞으로는 이를 3차원 공간상에 존재하는 음원의 위치를 추정할 수 있도록 해야 할 것이다.

## 후 기

이 논문은 2007 년도 정부(과학기술부)의 재원으로 한국과학재단의 국가지정연구실사업(M1050000112-05J0000-1121), 두뇌 한국 21 프로젝트와 정보통신부 및 정보통신연구진흥원의 IT 신성장동력핵심기술개발사업[2007-S001-01, 지능형 로봇 시·청각 신호처리 SoC] 사업의 일환으로 수행하였음.

## 참고문헌

(1) K. Nakadai, et al, 2000 "Active Audition for Humanoid," Proceedings of the 17<sup>th</sup> National Conference on Artificial Intelligence and 12<sup>th</sup> Conference on Innovative Applications of Artificial Intelligence, pp. 832-839  
 (2) 황성목, 2005 "머리전달함수를 이용한 로봇의 화자 위치 추정," 추계학술대회논문집, 한국소음진동공학회, pp. 637-640  
 (3) Y. Sasaki, S. Kagami and H. Mizoguchi, 2006, "Multiple Sound Source Mapping for a Mobile Robot by Self-motion Triangulation," Proceeding of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China

(4) H. Wang and P. Chu, 1997 "Voice Source Localization for Automatic Camera Pointing System in Videoconferencing," Proceedings of the 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '97) Vol. 1, pp. 187.

(5) M. S. Brandstein and H. Silverman, 1997 "A practical methodology for speech source localization with microphone arrays," Computer Speech and Language, 11(2):91-126.

(6) C. H. Knnapp and G. C. Cater, 1976, "The generalized correlation method for estimation of time delay," in proceeding of IEEE Trnas. on acoustics, speech and signal processing, Vol. Assp-24, No. 4