

마이크로폰 배열 구조를 이용한 음원의 위치 측정에 관한 연구

김진성^{*,}, 김도윤^{*}, 유범재^{*}, 김광배^{*}

한국과학기술연구원 지능로봇연구센터^{*}, 고려대학교 전기공학과^{*}

A Study on Sound Source Localization Using a Microphone Array

Jinsung Kim^{*,}, Do-Yoon Kim^{*}, Bum-Jae You^{*}, Kwang-Bae Kim^{*}

Intelligence Robotics Research Center KIST^{*}, Dept. of Electrical Engineering Korea University^{**}

요약 최근 들어 공장에서만 사용되는 로봇을 탈피하여, 사람과 함께 생활하면서 서비스를 제공할 수 있는 인간 친화적인 로봇이 전 세계적으로 개발되고 있다. 인간과 자연스런 상호작용을 위하여, 화자를 바라보고, 깨끗한 음성신호를 얻는 과정에서 음원의 위치 추적은 필연적이다. 본 연구는 마이크로폰에 도달되는 동일 신호의 도착시간지연(Time Delay of Arrival)정보를 이용하여, 음원 위치를 구하는 기존의 수학적 방법론[1]을 소개하고 검토한다. 모의실험을 통하여, 방법론[1]이 허상 음원을 생성하는 음원 위치 측정임을 보인다.

1. 서 론

인간 친화적인 로봇으로서 휴머노이드(Humanoid)에 관심이 고조되고 있다. 휴머노이드와 인간과의 관계에 있어서 가장 중요한 것은 자연스런 상호작용이다. 원활한 상호작용을 위하여 청각시스템의 연구는 필연적이다. 음원에서 발생하는 음파를 마이크로폰으로 측정하고, 측정된 신호를 이용하여 음원의 거리, 고도와 방위각을 알 수 있다. 이것을 음원의 위치 측정(Sound source localization)이라고 한다[1][2][4].

본 논문은 음원 위치측정에 관한 마이크로폰 배열 구조의 기하학적 해석을 통하여, 3차원 공간 음원 위치 측정에 필요한 최소 마이크로폰 개수가 4개임을 정리한다 [2]. 또한 간결한 음원 위치측정 방법으로 알려진 Seenu 방법론[1]을 소개하고, Seenu 방법론이 음원 위치측정 문제를 수학적 문제로 정의하는데 있어서, 필요충분관계가 아님을 반례를 들어 보인다.

2. 본 론

2.1 기하학적 해석

평면상의 임의의 음원 위치측정을 위해서 최소 3개 이상의 마이크로폰이 필요하다. 기존 마이크로폰과 다른 마이크로폰 사이에 도착시간지연을 2개 방정식으로 나타내어 연립방정식을 풀기 위함이다. 이와 마찬가지로 3차원 공간상의 임의의 음원 위치측정을 위해서는 최소 4개 이상의 마이크로폰이 필요하다. 결과적으로 평면 및 공간상의 좌표는 마이크로폰 간 도착시간지연을 이용하여 좌표 축 변수의 개수 만큼의 방정식을 얻음으로써 해결된다[2].

그림 1과 같이 직각 좌표계에 음원의 위치와 마이크

로폰의 위치를 정의한다. 음원 위치는 (x_s, y_s, z_s) 이고, 임의의 i 번째 마이크로폰 위치는 (x_i, y_i, z_i) 이다.

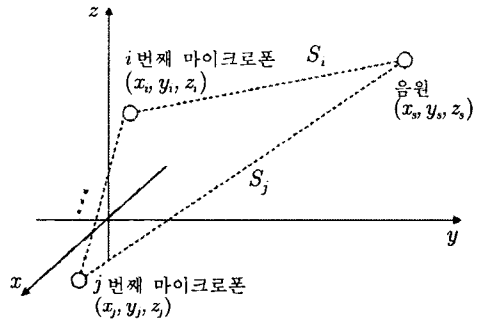


그림 1 마이크로폰 공간좌표 위치

음원과 i 번째 마이크로폰의 거리를 S_i 라고 정의하면,

$$S_i = \sqrt{(x_s - x_i)^2 + (y_s - y_i)^2 + (z_s - z_i)^2} \quad (1)$$

이다. S_i 와 S_j 의 거리차 δd_{ij} 는

$$\begin{aligned} \delta d_{ij} &= S_i - S_j \\ &= \sqrt{(x_s - x_i)^2 + (y_s - y_i)^2 + (z_s - z_i)^2} \\ &\quad - \sqrt{(x_s - x_j)^2 + (y_s - y_j)^2 + (z_s - z_j)^2} \end{aligned} \quad (2)$$

이다. 여기서 δd_{ij} 는 수신된 신호간의 시간지연 δt_{ij} 로부터 구한다:

$$\delta d_{ij} = \delta t_{ij} \cdot v \quad (3)$$

여기서 v 는 음속을 나타낸다. 임의의 i 번째 마이크로폰을 기준으로 나머지 3개의 마이크로폰의 상대적인 시간지연값 δt_{ij} 을 구하고, 식(3)을 통해 δd_{ij} 를 구한다. δd_{ij} 를 식(2)에 대입하여 변수 x_s, y_s, z_s 에 대한 3개의 방정식을 만들 수 있는데, 이로부터 음원의 위치를 계산한다. 3차원 공간 음원 위치 추적은 4개의 마이크로폰을 가지고 해결할 수 있다[1][3].

2.2 수치적 해석

본 절은 Seenu 방법론[1]을 소개하고 설명한다. 그림 2와 같이 직각 좌표계의 원점에 기준 마이크론을 고정한다[1].

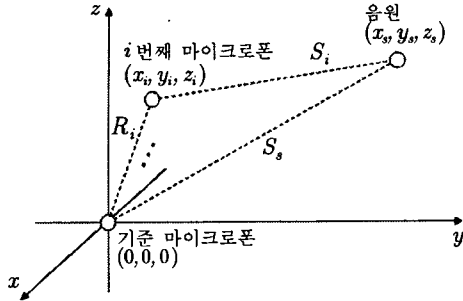


그림 2 기준 마이크론을 원점으로 한 공간좌표계

기준 마이크론과 i 번째 마이크론의 거리를 R_i , 기준 마이크론과 음원의 거리를 S_s , 음원과 i 번째 마이크론의 거리를 S_i 라고 정의하면 다음과 같이 나타낼 수 있다:

$$\begin{aligned} R_i^2 &= x_i^2 + y_i^2 + z_i^2, \quad i = 1, \dots, n \\ S_s^2 &= x_s^2 + y_s^2 + z_s^2 \\ S_i^2 &= (x_s - x_i)^2 + (y_s - y_i)^2 + (z_s - z_i)^2, \quad i = 1, \dots, n \end{aligned} \quad (4)$$

음원 위치에서 발생된 신호가 기준 마이크론과 i 번째 마이크론에 도달되는 도착시간지연차를 δt_i , 거리차를 δd_i 로 정의하면 다음과 같이 나타낼 수 있다:

$$\begin{aligned} \delta t_i &= (S_s - S_i)/v, \quad i = 1, \dots, n, \\ \delta d_i &= (S_s - S_i), \quad i = 1, \dots, n. \end{aligned}$$

문제를 간단히 정리하면, 음의 위치의 음원에서 신호가 발생하여, R_i 와 $\delta t_i, i = 1, \dots, n$ 가 주어졌을 때, 음원 위치 (x_s, y_s, z_s) 를 구하는 것이다.

식(4)로부터,

$$\begin{aligned} S_s^2 - S_i^2 &= x_s^2 - (x_s - x_i)^2 \\ &\quad + y_s^2 - (y_s - y_i)^2 + z_s^2 - (z_s - z_i)^2. \end{aligned} \quad (5)$$

을 간단히 유도할 수 있다. 좌변을 정리하면

$$(S_s + S_i)(S_s - S_i) = (2S_s - \delta d_i)\delta d_i$$

이고, 우변을 정리하면,

$$2x_s x_i + 2y_s y_i + 2z_s z_i - R_i^2.$$

이다. 요약하면 다음과 같다:

$$\begin{aligned} x_s x_i + y_s y_i + z_s z_i &= S_s \delta d_i + (R_i^2 - \delta d_i^2)/2 \\ &, \quad i = 1, \dots, n \end{aligned} \quad (6)$$

식(6)을 벡터로 확장하여 음원 위치 (x_s, y_s, z_s) 를 구하고자 다음을 정의한다:

$$W_i = (R_i^2 - \delta d_i^2)/2, \quad i = 1, \dots, n$$

$$W^T = [W_1, W_2, \dots, W_n],$$

$$D^T = [\delta d_1, \delta d_2, \dots, \delta d_n],$$

$$V^T = [x_s, y_s, z_s],$$

$$A = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \dots & \dots & \dots \\ x_n & y_n & z_n \end{bmatrix}.$$

위의 정의를 이용하여 식(6)을 다음과 같이 정리한다:

$$AV = S_s D + W. \quad (7)$$

식(7)은 $B = (A^T A)^{-1} A^T$, $C = BD$, $Y = BW$ 정의를 이용하여

$$V = S_s C + Y \quad (8)$$

으로 정리 된다. 식(8)에서 직접 음원 위치 (x_s, y_s, z_s) 를 구할 수 없으므로,

$$V^T V = S_s^2 C^T C + S_s (C^T Y + Y^T C) + Y^T Y \quad (9)$$

으로 전개한다. $V^T V = S_s^2$ 이므로,

$$S_s^2 (C^T C - 1) + S_s (C^T Y + Y^T C) + Y^T Y = 0 \quad (10)$$

에서, 양수 S_s 를 구한 뒤, 식(8)을 이용하여 음원 좌표 (x_s, y_s, z_s) 를 찾아낸다.

2.3 적용 및 토의

본 절은 방법론[1]이 음원의 위치 측정 문제를 수학적 문제로 정의하는 과정에 있어서, 필요충분조건이 되지 않을음을 보인다. [1]의 예제로서 (1,0,0), (0,1,0), (0,0,1)에 위치한 마이크론과 원점에 위치한 기준 마이크론을 가지고 3차원 음원 추적의 경우를 근거로 제시하고 있다. 그림 3은 $x > 0, y > 0, z > 0$ 의 영역에서 음원이 발생했을 때 위치를 정확히 찾아냄을 보여준다.

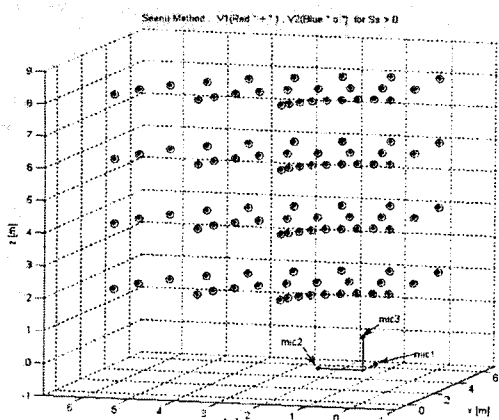


그림 3 정확한 음원 위치 측정의 사례

그러나 그림 4와 같이 전방향의 음원에서 신호가 발생하였을 경우, 양수 S_s 가 2개 존재하는 영역이 생김을 확인할 수 있다. 구체적인 반례 반례로서 음원의 위치가 (-2.1, 5.6, -6) 경우, S_s 는 8.4853과 1.8917, 즉 2개가 존재하기에 음원 추적의 결과로 (-2.1, 5.6, -6)과 (-0.1, 1.49, -1.16)이 나온다. 하나의 허상 음원 위치를 추가하여 생성해 내는 것이다. 전 방향 음원에 대한 위치 측정을 한 결과, 상당한 수의 허상 음원 위치가 생성됨을 확인하였다.

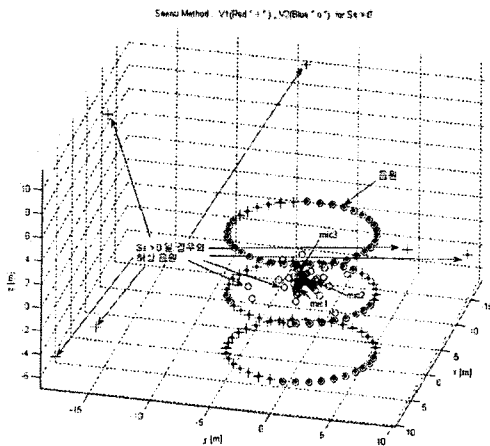


그림 4 허상음원이 생성하는 경우

3. 결 론

3차원 공간의 음원 위치 측정을 위한 마이크로폰 배열 구조의 기하학적 해석 방법과 수치적 해석 방법을 다루었다. 마이크로폰 간 도착시간지연을 이용하고, 3차원 공간좌표 축 변수의 개수인 3개의 방정식을 얻기 위하여 최소 4개의 마이크로폰이 필요함을 정리하였다. 수치적 해석 방법인 Seenu 방법론[1]을 소개하여, 공간상의 음원추적이 간단히 해결됨을 보였다. 그러나 음원 추적 과정에서 허상 음원이 생성됨을 보임으로서 공간상의 음원

추적 문제가 수학적 문제로 정의 될 때, 필요충분조건으로 정의되지 않음을 보였다.

앞으로 연구과제는, 음원 추적이 전 방향에서 가능하도록 마이크로폰 배열 구조를 제안하고, 음원의 위치 측정 문제를 필요충분조건 수학적 문제로 정의하는 것이다.

[참 고 문 헌]

- [1] Seenu S. Reddi, "An Exact Solution to Range Computation with Time Delay Information for Arbitrary Array Geometries", IEEE Transactions on Signal Processing, Vol.41, No.1, pp.485-486, January 1993.
- [2] 배진만, "인간의 청각시스템을 응용한 음원위치 추정에 관한 연구", 경희대학교 전자공학과 석사학위논문, 2004.
- [3] 이호철, "시간지연추정을 이용한 음원위치추적에 관한 연구", 국민대학교 자동차공학전문대학원 차체 및 새시 전공 석사학위 논문, 2000.
- [4] Yiteng Huang et al. "Real-Time Passive Source Localization : A Practical Linear-Correction Least-Squares Approach", IEEE Transactions on Speech and Audio Processing, Vol.9, No.8, pp.943-955, November 2001.