

4개의 마이크로폰 배열을 이용한 3차원 좌표 기반 실시간 위치 추정 방법

하재승, 김석찬*

부산대학교, *부산대학교

gkwotmd95@pusan.ac.kr, *sckim@pusan.ac.kr

Real-time Location Estimation Based on 3-Dimensional Coordinates Using 4 Microphone Arrays

Jae Seung Ha, Suk Chan Kim*

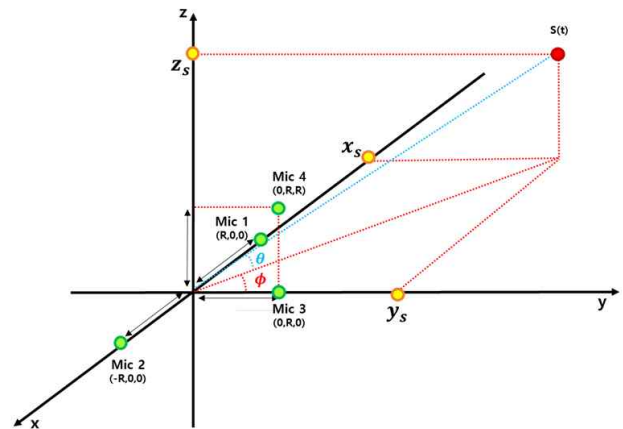
Pusan National Univ., *Pusan National Univ.

요약

본 논문은 실시간으로 4개의 마이크에 도달하는 음원의 시간차를 활용하여 3D 음원 위치 추정 방법을 제안한다. 음원이 있는 교점을 추정하기 위해 모든 경우의 좌표를 구하고 각도를 이용해 최종 좌표를 구분한다. 실내에서 근거리를 측정하고 실외에서 원거리로 측정한 결과를 시뮬레이션 결과와 비교하여 제안한 방법의 성능을 평가한다.

I. 서론

위치 추정 기술이 발전하면서 다양한 환경에서 각기 다른 소스를 이용한 기술이 연구되고 있다. 그 중 음원을 이용한 연구들도 활발히 진행되고 있다. 음원 위치 추정 기술은 2차원, 3차원 공간에서 음원을 찾는 데 사용되는 기술을 말한다. 하지만 음원의 정확한 위치를 파악하기 위해서는 고성능 마이크가 요구되어 소모 마이크가 많을수록 정확한 추정이 가능하지만, 비용이 많이 발생하는 문제점이 생긴다. 따라서 본 연구에서는 3차원 위치에 사용되는 마이크 4개만을 배치하여 각 마이크에 도달하는 시간 지연을 이용한 실시간 음원 위치를 추정을 적용하면서, 평면과 쌍곡선이 겹쳐서 생기는 모든 점을 이용하여 최종적인 좌표 추정 값을 출력하는 방법에 대해서 제안한다.



[그림 1] 마이크 배치도

II. 본론

TDOA (Time Difference Of Arrival)는 마이크 쌍 간의 상대적 시간 차이에 의존된다. 각 마이크 간에 여러 개의 쌍곡선이 생성되는데 이 쌍곡선들의 교점이 되는 부분을 음원의 위치로 추정하는 방법이다. GCC-PHAT (Generalized Cross Correlation Phase Transform)는 주파수 영역에서 두 센서의 CC(Cross Correlation)를 구한 후에 백색화 가중치를 곱하는 과정을 추가하여 신호의 스펙트럼의 영향을 배제하는 효과를 얻는다[2]. 이를 통해 TDOA를 이용한 시간 차이로 위치 추정을 하는데 있어 더 높은 정확도를 이끌어낼 수 있다. 음원의 위치는 음파가 도달하는 방향(방위각, 고도각), 마이크 배열에서 마이크 간 간격에 의해 결정된다. 마이크 별로 도달한 시간차 값(τ)은 식(1)로 표현할 수 있다.

$$\begin{aligned} d_n &= \sqrt{(x_s - x_n)^2 + (y_s - y_n)^2 + (z_s - z_n)^2} \\ d_m &= \sqrt{(x_s - x_m)^2 + (y_s - y_m)^2 + (z_s - z_m)^2} \\ V_{sound} &= 20.05 \sqrt{273.15 + \text{Temperature}(^{\circ}\text{C})} \\ \tau_{nm} &= \frac{(d_n - d_m)}{V_{sound}} \end{aligned} \quad (1)$$

d_n 과 d_m 은 마이크 배열 내에서 각기 다른 위치의 n , m 번 마이크와 음원 간의 거리를 나타내고 V_{sound} 은 음속이다. 식(1)을 이용하여 각도와 거리를 추정하는 계산에 이용한다[1].

먼저 식(1)로 나타낸 mic 1, 2, 4간의 시간차 값인 τ_{12} , τ_{14} , τ_{24} 를 두 쌍씩 연결하여 3차원 상의 직선인 $Ax + By + Cz = D$ 형태로 바꾸면 식(2)이 된다.

$$\begin{aligned} 1) & \left(\frac{1}{\tau_{14}} + \frac{1}{\tau_{24}}\right)x - \left(\frac{1}{\tau_{14}} - \frac{1}{\tau_{24}}\right)y - \left(\frac{1}{\tau_{14}} - \frac{1}{\tau_{24}}\right)z = \left(\frac{V_{sound}^2}{2^*R^*(\tau_{24} - \tau_{14})}\right) + \frac{R}{\tau_{24}} * \left(-\frac{1}{\tau_{14}} - \frac{1}{\tau_2}\right) \\ 2) & \left(\frac{1}{\tau_{14}} - \frac{1}{\tau_{12}}\right)x - \frac{1}{\tau_{14}}y - \frac{1}{\tau_{14}}z = \left(\frac{V_{sound}^2}{2^*R^*(\tau_{14} - \tau_{12})}\right) - \frac{R}{2^*\tau_{14}} \\ 3) & \left(\frac{1}{\tau_{24}} - \frac{1}{\tau_{12}}\right)x - \frac{1}{\tau_{24}}y - \frac{1}{\tau_{24}}z = \left(\frac{V_{sound}^2}{2^*R^*(\tau_{24} - \tau_{12})}\right) - \frac{R}{2^*\tau_{24}} \end{aligned} \quad (2)$$

다음으로 y 를 소거하기 위해 mic 1, 2, 3간의 시간차 값인 τ_{12} , τ_{13} , τ_{23} 을 두 쌍씩 연결하여 $y = ax + b$ 형태로 변환하여 식(3)으로 나타낸다. 식(3)과 식(2)를 연결하여 y 를 소거한다[1].

$$\begin{aligned} 1) & y = \frac{\tau_{23} + \tau_{13}}{\tau_{23} - \tau_{13}} * x - \frac{V_{sound}^2}{2^*R^*\tau_{23}*\tau_{13}} \\ 2) & y = \left(1 - \frac{2^*\tau_{13}}{\tau_{12}}\right) * x + \frac{V_{sound}^2}{2^*R^*(\tau_{12} - \tau_{13})*\tau_{13}} \\ 3) & y = -\left(1 + \frac{2^*\tau_{23}}{\tau_{12}}\right) * x - \frac{V_{sound}^2}{2^*R^*(\tau_{12} - \tau_{23})*\tau_{23}} \end{aligned} \quad (3)$$

마지막으로 τ_{12} , τ_{13} , τ_{23} 를 식(1)을 이용하여 $\alpha x^2 + \beta x + \gamma = 0$ 형태인 식(4)로 변경한다[1].

$$\begin{aligned}
 1) \quad & (1+a^2 - (\frac{2^*R}{V_{sound}^* \tau_{12}})^2 + A_{dd})^* x^2 + (2^*a^*b + B_{dd})^* x + R^2 + b^2 - (\frac{V_{sound}^* \tau_{12}}{2})^2 + C_{dd} = 0 \\
 2) \quad & (1+a^2 - (\frac{(a-1)^*R}{V_{sound}^* \tau_{13}})^2 + A_{dd})^* x^2 + (2^*a^*b^* - R^*(a+1) - 2^*b^*(a-1) \\
 & * (\frac{R}{V_{sound}^* \tau_{13}})^2 + B_{dd})^* x + (b-R)^2 - (\frac{b^*R}{V_{sound}^* \tau_{13}} - \frac{V_{sound}^* \tau_{13}}{2})^2 + C_{dd} = 0 \\
 3) \quad & (1+a^2 - (\frac{(a+1)^*R}{V_{sound}^* \tau_{23}})^2 + A_{dd})^* x^2 + (2^*a^*b^* - R^*(a-1) - 2^*b^*(a+1) \\
 & * (\frac{R}{V_{sound}^* \tau_{23}})^2 + B_{dd})^* x + (b-R)^2 - (\frac{b^*R}{V_{sound}^* \tau_{23}} - \frac{V_{sound}^* \tau_{23}}{2})^2 + C_{dd} = 0
 \end{aligned} \quad (4)$$

$$\begin{aligned}
 A_{dd} &= \frac{(A + \alpha^* B)^2}{C^2} \\
 B_{dd} &= \frac{-2^*(A + \alpha^* B)^*(D - B^* \beta)}{C^2} \\
 C_{dd} &= \frac{(D - B^* \beta)^2}{C^2}
 \end{aligned} \quad (5)$$

식(2)와 식(4)를 연립하면 mic 1, 2, 3에 대한 평면과 mic 1, 2, 4에 대한 평면 그리고 3d 쌍곡선들이 겹치는 점이 27개가 도출된다. 식(4)에 근의 공식을 이용하면 54개의 점이 나온다. $-B + \sqrt{B^2 - 4^*A^*C}$ 와 $-B - \sqrt{B^2 - 4^*A^*C}$ 기준으로 x, y, z 를 분류한 후 각각을 평균하면 최종적인 근의 공식 좌표가 2개 구해진다. 구해진 좌표를 이용하여 식(6)에서 방위각 ϕ 을 구하고 고도각 θ 는 mic 3, 4의 시간차 값인 τ_{34} 를 이용하여 구한다.

$$\begin{aligned}
 \phi &= \cos^{-1}(\frac{x}{\sqrt{x^2 + y^2 + z^2}}) \\
 \text{if } \phi &> \frac{\pi}{2} : \\
 &\quad \text{if } a > 0 : \\
 &\quad \quad \text{quadrant} = 2 \\
 &\quad \text{else} : \\
 &\quad \quad \text{quadrant} = 3 \\
 \text{else} \\
 &\quad \text{if } a > 0 : \\
 &\quad \quad \text{quadrant} = 1 \\
 &\quad \text{else} : \\
 &\quad \quad \text{quadrant} = 4 \\
 \text{if } \text{quadrant} &= 3 \text{ or } 4 \\
 \phi &= -\phi \\
 \theta &= \cos^{-1}(\sqrt{1 - (\frac{V_{sound}^* \tau_{34}}{R})^2})
 \end{aligned} \quad (6)$$

최종적인 (x, y, z) 좌표는 식(6)의 방위각 ϕ 을 이용하여 정한다.

III. 실험 환경 및 결과

실험은 ADC(Analog-to-digital converter)와 2개의 PC를 이용하여 프로그램 상으로 실시간 음원 데이터를 받아서 알고리즘에 적용했다. 마이크는 지표면으로부터 0.5m 정도 높이의 삼각대를 세우고, 그 위에 고정시켜 그림 1와 같이 배치하였다. 마이크 간 거리 R은 1m이며 실내에서는 스피커로 드론 음원을 틀어 진행하였고, 실외에서는 Mavic 2 - DJI 드론을 대상으로 측정했다. 실험결과들은 다음과 같다. 표 [1]은 실내, 표 [2]는 실외에서 측정한 값이다. 실내는 장애물이 없는 부분에서 측정했다. 거리는 87 %정도의 정확도를 보이고, 각도는 83 %정도의 정확도를 보인다. 실외는 드론을 이용하여 5m 거리에서 높이를 올려가며 실험을 진행하였다. 3m 높이에서는 방위각과 고도각의 범위의 오차가 적지만 높이 올라 갈수록 오차가 늘어나는 것이 확인되었다. 거리 또한 높이 3m 이상으로는 Variance가 큰 결과가 보였다.

표 1. 실험 결과 [실내]

좌표		x	y	z	Φ	Θ
x=2 y=0.5 z=1	실험	1.71	0.58	1.30	39.95	24.14
	이론	1.89	0.50	0.97	30.03	13.30
x=1.5 y=1.2 z=1.5	실험	1.22	1.19	1.33	52.87	33.18
	이론	1.49	1.22	1.49	52.28	32.41
x=2 y=-0.5 z=1.5	실험	2.05	-0.405	1.38	314.32	19.84
	이론	1.961	-0.485	1.481	319.47	21.46
x=2 y=-1.6 z=1.5	실험	1.54	-1.40	1.41	312.27	16.35
	이론	1.49	1.22	1.49	307.72	32.41

표 2. 실험 결과 [실외]

좌표		x	y	z	Φ	Θ
x=4.5 y=2.1 z=3	실험	2.73	1.64	1.64	40.44	27.03
	이론	4.44	2.07	2.95	39.17	28.24
x=4.5 y=2.1 z=5	실험	0.59	0.49	0.62	54.29	50.88
	이론	4.46	2.08	4.96	50.32	44.08
x=4.5 y=2.1 z=7	실험	5.28	4.12	2.3	60.52	66.20
	이론	4.47	2.09	6.9	58.37	54.47

IV. 결론

본 논문은 마이크 배치에 TDOA와 GCC-PHAT를 적용하여 최소한의 마이크 수로 나올 수 있는 모든 좌표 도출하고, 각도를 이용하여 최종적인 음원의 위치를 추정하는 방법을 제안했다. 제안하는 방법은 실내에서는 준수한 정확도를 보였지만, 잡음이 심한 실외에서는 오차가 심한 것이 확인되었다. 추후에는 잡음에 강한 알고리즘으로 이를 보완하여 연구를 진행할 예정이다.

ACKNOWLEDGMENT

“This research was financially supported by the Institute of Ciiil Military Technology Cooperation funded by the Defense Acquisition Program Administration and Ministry of Trade, Industry and Energy of Korean government.”

참 고 문 헌

- [1] Ali Pourmohammad and Seyed Mohammad Ahadi "Real Time High Accuracy 3-D PHAT-Based Sound Source Localization Using a Simple 4-Microphone Arrangement" in IEEE SYSTEMS JOURNAL Sept 2012
- [2] Lim Jun-Seok, Cheong MyoungJun and Kim, Seongil "Improved generalized cross correlation-phase transform based time delay estimation by frequency domain autocorrelation" in the Journal of the Acoustical Society of Korea pp271-275, Sept 2018