JKIICE Journal of the Korea Institute of Information and Communication Engineering

한국정보통신학회논문지(J. Korea Inst. Inf. Commun. Eng.) Vol. 21, No. 4 : 708~717 Apr. 2017

스마트 강의 시스템을 위한 시간차 검출 방식의 실시간 음원 추적 기법 구현

강민수 · 오우진*

Implementation of Real-time Sound-location Tracking Method using TDoA for Smart Lecture System

Minsoo Kang · Woojin Oh*

School of Electronics, Kumoh National Institute of Technology, Kumi, 39177, Korea

요 약

음원 추적은 지능형 CCTV, 화상회의시스템 , 음성 명령 인식 등에서 널리 활용되고 있다. 본 논문에서는 스마트 강의시스템에 적합한 천정 부착형 직교배열 마이크로 음성 신호의 도착 시간차인 TDoA(Time Difference of Arrival)를 이용하는 실시간 음원추적 기법을 제안한다. TDoA를 위한 점 음원과 평행음원 모델을 분석하고, 3개의 선형배열 마이크를 이용하여 상호 상관 방안을 제안하였다. 또한 직교축에 십자 배열된 5개의 마이크를 사용하여 전방위(omni-direction)에서 음원 추적이 가능함을 보였다. 무음구간을 제거하기 위하여 수신 에너지를 이용하였으며 상호 상관을 부호로 연산하여 계산량을 줄이고 추정 결과에 미디언 필터(Median Filter)를 적용하여 안정도를 높였다. 제안된 시스템은 고속 MCU인 TMS320F379D와 MEMs마이크 모듈로 구현하여, 방향 검출 성능이 백색잡음이나 음악 등에 대해서는 0.5도, 음성에 대해서는 6.5도의 정밀도를 보였다.

ABSTRACT

Tracking of sound-location is widely used in various area such as intelligent CCTV, video conference and voice commander. In this paper we introduce the real-time sound-location tracking method for smart lecture system using TDoA(Time Difference of Arrival) with orthogonal microphone array on the ceiling. Through discussion on some models of TDoA detection, cross correlation method using linear microphone array is proposed. Orthogonal array with 5 microphone could detect omni direction of sound-location. For real-time detection we adopt the threshold of received energy for eliminating no-voice interval, signed cross correlation for reducing computational complexity. The detected azimuth angles are processed using median filter for lowering the angle deviation. The proposed system is implemented with high performance MCU of TMS320F379D and MEMs microphone module and shows the accuracy of 0.5 and 6.5 in degree for white noise and lectured voice, respectively.

키워드: 음원추적, 직교배열 마이크, 시간차 검출, 스마트 강의 시스템

Key word: Sound location tracking, Array Microphone, TDoA, Smart Lecture System

Received 09 January 2017, **Revised** 10 January 2017, **Accepted** 02 February 2017

* **Corresponding Author** Woojin-Oh (E-mail:wjoh@kumoh.ac.kr, Tel:+82-54-478-7485)
School of Electronics, Kumoh National Institute of Technology, Kumi, 39177, Korea

Open Access https://doi.org/10.6109/jkiice.2017.21.4.708 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/li-censes/ 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

음원 추적은 지능형 CCTV, 화상회의시스템, 로봇분야 등에서 음원의 방향을 인식하여 사건 검출 및 영상감시, 화자 지향성 마이크, 로봇-인간 인터페이스로 다양하게 활용되고 있다. 대부분의 음원 추적 기술은 사람의 인지방식과 유사하게 2개의 마이크를 사용하여 수신 음성의 전력차, 시간차, 위상차를 이용하는 IID (Interaural Intensity Difference), ITD(Interaural Time Difference), IPD(Interaural Phase Difference)를 사용하고 있다. 음성이 아닌 레이더나 RF 신호와 같은 다른 신호원의 추적은 위상배열을 갖는 송수신기로 RoA(RSSI of Arrival), ToA(Time of Arrival), AoA(Angle of Arrival)를 이용하여 다양하게 연구되고 있다[1-4].

본 논문에서는 멀티미디어 강의시스템에 적합한 음원 추적 방법을 제안한다. 강사 추적 강의 시스템은 음원의 위치를 검출하여 천정에 부착된 다수의 마이크 간에 자동 로밍이 이루어지므로 강사는 마이크를 착용하지 않아도 편리하게 강의가 가능한 장점이 있다.

제안된 음원 추적은 강의 환경인 교단의 크기를 고려하여 배열 마이크를 구성하고 신호의 도착 시간차인 TDoA(Time Difference of Arrival)를 실시간으로 측정하게 된다. TDoA는 직선상에 위치한 3개의 마이크로부터 기준신호에 대하여 상호 상관으로 구해지며, 마이크 배열을 직교축으로 배치하여 전방향에서 음원 추적이 가능함을 보였다. 상호 상관은 부호 연산으로 계산량을 줄였으며, 추정 각도에 대하여 미디언 필터(Median Filter)를 적용하여 오류로 인한 급격한 변화를 줄였다. 무음이나 잡음 구간에서는 음원 추적을 하지 않도록 간단히 수신 에너지를 이용하였다[5,6].

제안된 시스템은 고속 MCU인 TMS320F379D에 12비트 ADC가 내장된 MEMs마이크 모듈을 SPI로 연결하여 구현하여 성능을 검증하였다. 백색잡음이나 음악과 같은 연속적인 신호에서는 0.5도의 오차로 추적 성능을 보였으며, 강의 음성에 대해서는 6.5도의 해상도를 보였다.

본 논문의 구성은 II장에서 TDoA기반의 음원 추적 기법을 알아보고, III 장에서 강의시스템에 필요한 오차와 범위에 적합한 TDoA기법을 분석하고 전방위 음원 검출 방안을 제안하였다. 마지막으로 IV장에서 시스템 구현과 성능을 설명하겠다.

Ⅱ. TDoA기반 음원 추적

음원의 방향을 추적하기 위한 수신 모델은 그림 1과 같이 2가지로 나눌 수 있다. 마이크 간의 거리가 d인 경우에 대하여 마이크와 음원간의 각도를 평행 음원과 점 음원의 2가지 모델이 있다.

평행 음원 모델은 그림 1a)와 같으며 음원의 방위각 (azimuth angle) Θ는 직각 삼각형 AMC에서 수선과 이루는 각이며, 이는 ∠C와 같아진다. 평행 음원이므로 B와 C에 위치한 마이크 L, R과 음원의 거리차 r_d는 선분 c와 동일하다. 따라서 수직선에서의 방위각 Θ는 다음과 같은 관계를 갖는다[1].

$$\sin\theta = \frac{c}{d} = \frac{r_d}{d} \tag{1}$$

r_d는 마이크 L과 R의 수신신호에서 첨두치 검출이나 상호상관을 이용하여 시간차인 TDoA로부터 쉽게 구할 수 있다.

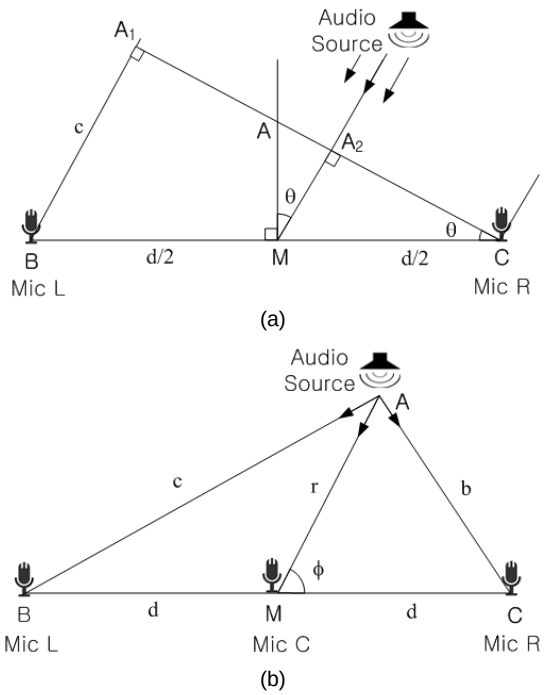


Fig. 1 Sound localization model (a) Parallel source (b) Point source

그림 1b)는 점 음원 모델로 3개의 마이크로 구성된다. 삼각형 ABC에 중선 AM을 연결하면 음원과의 거리인 r이 되며 파푸스 정리(Pappus's theorem)와 코사인 법칙으로부터 각각 다음의 관계식을 얻을 수 있다[2].

$$b^2 + c^2 = 2(r^2 + d^2) \tag{2}$$

$$b^2 = r^2 + d^2 - 2rd\cos\phi \tag{3}$$

여기서 b, c는 그림 2와 같이 각각 마이크 L과 C, C와 R의 TDoA로부터 구할 수 있다.

$$b = r + vt_{TDoA_CR} = r + r_{db} \tag{4a}$$
$$c = r + vt_{TDoA_LC} = r + r_{dc} \tag{4b}$$

v는 음속으로 온도 T℃ 일 때 $(331.5 + 0.6T)$ m/s이다. 이를 식 (2)에 대입하여 음원과의 거리 r은

$$r = \frac{2d^2 - r_{db}^2 - r_{dc}^2}{2(r_{db} + r_{dc})} \tag{5}$$

이 되며 식 (3)에 대입하여 정리하면 다음 식으로 방위각을 얻을 수 있다.

$$\cos\phi = \frac{d^2 + 2r \cdot r_{db} - r_{db}^2}{2rd} \tag{6a}$$

$$\cos(\pi - \phi) = \frac{d^2 + 2r \cdot r_{dc} - r_{dc}^2}{2rd} \tag{6b}$$

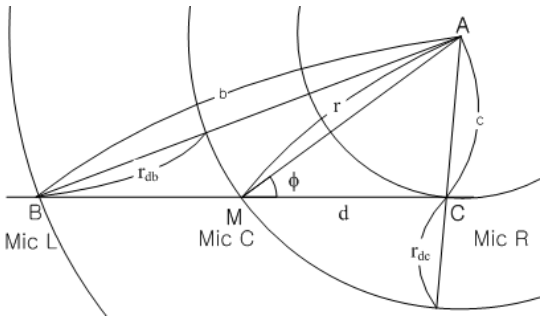
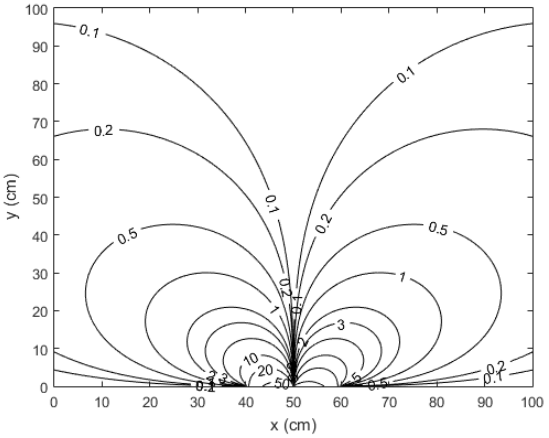
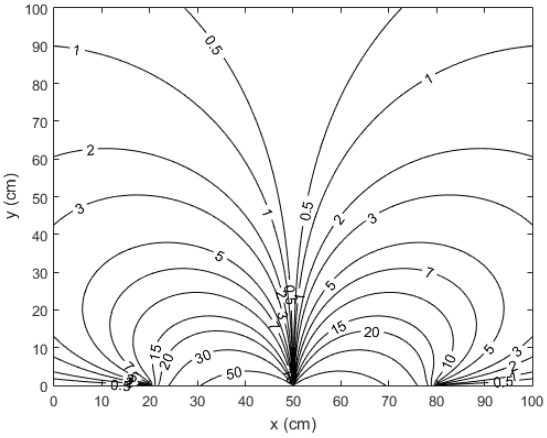


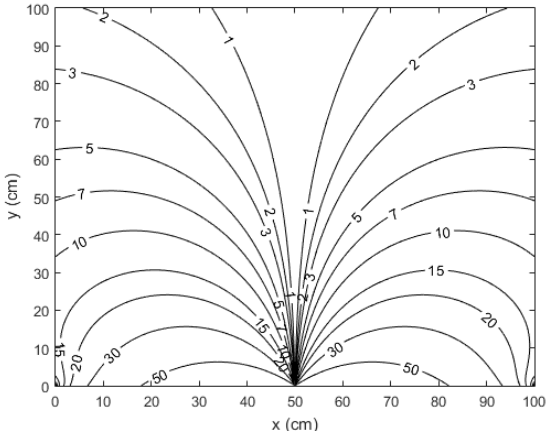
Fig. 2 Example of range distance from sound source



(a)



(b)



(c)

Fig. 3 Azimuth angle error using parallel source model with (a) d=20cm (b) d=60cm (c) d=100cm

TDoA은 방위각이나 거리를 구할 때 정확도에 영향을 미치는 중요한 요소이다. 평행 모델의 경우 각 마이크 L과 R은 중심점 M과의 TDoA가 항상 같은 크기에 부호만 다른 값이 된다. 그러나 점 모델의 경우는 그림 2에서 보인 바와 같이 음원의 위치 A가 M의 수직에 근접하면 같은 부호가 될 수 있다. 평행 모델에 대하여 1m의 범위에 대하여 실제 방위각과 식 (1)으로 계산된 방위각의 오차를 그림 3에 보였다. x축의 중점에서 마이크 L과 R이 양쪽으로 d/2만큼 이격된 위치에 있는 경우를 시뮬레이션 한 것이다. 마이크 간의 거리 d가 증가할수록 오차가 커진다. 참고로 점 모델은 식 (6)으로 계산되며 실제 값과 동일하여 오차가 없다.

다음 장에서는 음원 추적 성능에 영향을 미치는 관련 요소인 TDoA의 측정 방안과 마이크간의 거리 d의 선택 방법 등을 살펴보겠다.

Ⅲ. 제안된 TDoA 기법 및 전방향 음원추적

3.1. 기존의 TDoA검출 기법 분석

음원 추적은 식 (1) 또는 (6)에 의하여 구해지며, 그 성능은 각 마이크와 음원간의 도달 시간차의 정확도에 비례한다. 시간차는 TDoA로 정의하며 음원의 속도를 알고 있으므로 거리차를 구할 수 있다.

TDoA를 측정하기 위한 방법은 위상비교, 첨두치 검출, PN(Pseudo Noise)부호를 이용하는 상관기(correlator) 등의 방법이 있다[7-9].

위상을 검출하는 것은 IQ 채널로 수신하는 경우에 가능하여 음성에 적용하기는 어려우며 RF 신호 검출에 많이 사용된다. 또한 음성의 경우 반송파로 변조하지 않기 때문에 신호의 주파수 대역폭이 20Hz~20KHz로 1000배나 되어 위상차를 검출하기 매우 어렵다.

PN 부호는 상호 상관(Cross Correlation)이 0에 가까운 이진 수열로 CDMA, GPS 등에서 시간 측정, 동기화 등을 위해 사용되고 있다[7]. 음성 신호는 사람이 발성하기 때문에 정해진 수열의 부호화를 적용할 수 없다.

첨두치 검출은 비교기나 신호 검출기를 사용하여 일정 크기 이상의 수신 파형이 처음 도착하는 시간을 관측하는 방식이다[9]. 초음파, 레이저, 광선등과 같은 반사특성이 우수하여 SNR이 높아 신호가 명확히 구분되는 경우에 적합하다.

또한 이 방식은 사전에 정해진 기준 신호를 송신하고, 지연되어 도착하는 수신 신호를 검출하는 방식으로 본 논문에 필요한 강사의 일반 음성을 추적하는데 적용하기 어렵다.

멀티미디어 강의 시스템에서의 음원 추정은 단일 화자에 대한 수신 시간차를 찾는 것이므로 자기상관(auto correlation)으로 지연값을 예측하는 문제로 접근할 수 있다. 먼저 자기상관은 지연 τ가 0이 아닌 경우에 상관값 $R_x(\tau)$ 는 다음의 성질을 갖는다.

$$R_x(0) \geq R_x(\tau) \tag{7}$$

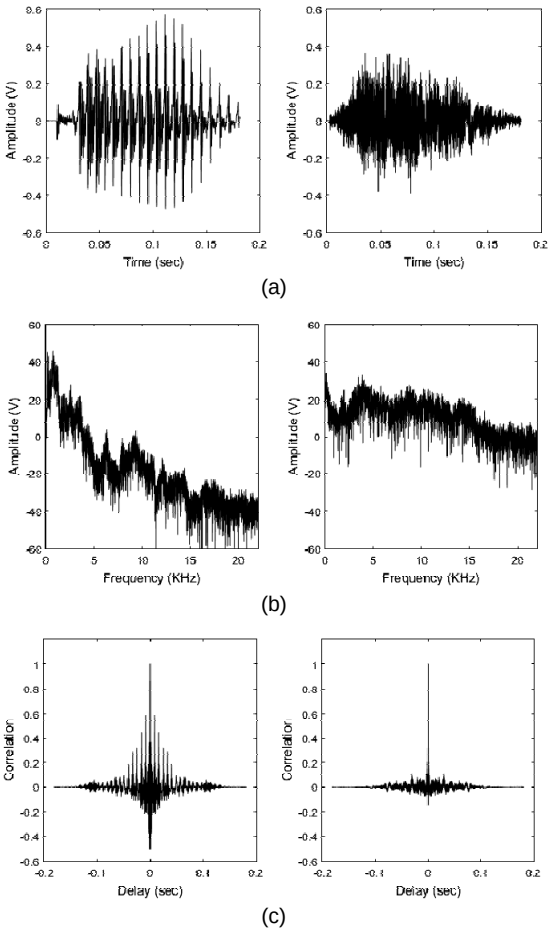


Fig. 4 Comparison of voiced (left) and unvoiced (right) signals: (a) Waveform (b) Spectrum (c) Normalized Correlation

의사 잡음인 경우에는 $R_x(\tau)$ 는 0에 가까운 값을 갖고 있어 수신 신호 간에 상관을 구하면 정확한 시간지연을 추정할 수 있다.

음성신호의 상관특성을 살펴보자. 그림 4에 유성음(Voiced)과 무성음(Unvoiced)에 대하여 신호파형, 스펙트럼, 자기상관을 보였다. 유성음은 '아', 무성음은 '츠'를 사용하여, 44.1KHz 표본화, 8000개의 데이터를 이용하였다. 유성음과 무성음 모두 정규 상관(normalized correlation)값은 지연이 0일 때 최대치를 보이고 있어 상관으로 지연 시간을 측정하는 것이 가능할 것이다. 다만 무성음의 경우 백색잡음과 유사하게 일치하는 경우에만 큰 값을, 나머지의 경우는 거의 0에 가까운 값을 보이고 있다. 그림 4(c)에서 유성음의 경우 피치(pitch)에 의해 주기적으로 반복되는 상관특성이 나타나므로 상관에 사용하는 기준신호의 길이를 적절히 선택해야 한다.

3.2. 제안된 TDoA 검출 상관기

본 논문에서는 멀티미디어 강의 시스템에 적합한 강사 위치 추적을 10도 이내의 정확도로 구현하는 것을 목표로 하고 있다. 검출 범위는 강단의 크기를 고려하여 6x3m로 설정하였다. 평행 음원 모델은 계산이 간단하지만 그림 3에 보인 바와 같이 0.5m 이내의 근거리 영역에서는 큰 오차를 보이므로 3개의 마이크를 사용하는 점 음원 모델을 선택하였다.

먼저 마이크간의 거리 d를 결정하자. 반적인 대화는 대역폭이 300~3400Hz이지만 가장 민감도가 높은 대역인 1KHz를 기준으로 선택하였다. 실내온도인 20도에서 음속은 343.5m/s이 되고, 음성 주파수가 1KHz인 경우 파장은 34.35cm가 된다. 위상 배열 안테나의 경우에는 상호 보완된 수신을 위한 다이버시티(diversity) 효과나 빔포밍(Beamforming)을 위해 반파장의 1/n배를 사용하지만 시간차 검출 환경에서는 동일한 세기로 검출되도록 파장의 정수배 간격으로 설정해야 한다. 본 논문에서는 마이크 간격을 파장에 근사한 34.5cm로 설정하였다.

상관기에서 데이터 길이가 길수록 더 큰 첨두치를 얻을 수 있지만 계산량이 늘어나는 문제가 있다. 검출해야할 시간차의 범위가 한정되어 있으므로 적정한 길이를 선택하는 것이 필요하다. 마이크 간의 거리에 의해 최대 수신 시간차는 다음 수식과 같이 정해지며, 34.5cm인 경우 약 1ms가 된다.

$$\tau_{\max} = \frac{d}{v} = \frac{34.5cm}{343.5m/s} \approx 1ms \tag{8}$$

1ms의 시간은 44.1KHz로 표본화 하는 경우 44.1개의 데이터가 된다. 기준신호는 그림 1b)의 중앙 마이크에서 수신된 신호 $x_c(k)$가 되며, L과 R의 마이크 신호 $x_L(k)$, $x_R(k)$와 상호 상관을 각각 구하게 된다. 검출해야 할 시간차는 ±1ms의 범위이므로 최소 2ms의 데이터가 필요하다. 기준 신호의 길이가 N_c일 때 그림 5와 같이 L, R 마이크의 신호는 2ms의 데이터가 추가로 필요하다. 본 논문에서는 기준 신호를 5ms에 해당하는 220개의 데이터, L/R 마이크의 신호는 308개의 데이터를 사용하여 TDoA 검출을 수행하였다.

마이크간의 거리 및 표본화 속도에 따른 각도 해상도를 알아보자. 식 (1)은 평행 음원 모델의 각도이지만, 원거리에서는 오차가 적고 간단하므로 이를 적용하여 해석하겠다. 거리차 해상도 Δr_d는 표본화 간격 T_s와 $\Delta r_d = vT_s$의 관계를 가지므로 각도해상도는 다음과 같이 정의할 수 있다.

$$\theta = \sin^{-1}\left(\frac{n\Delta r_d}{d}\right) = \sin^{-1}\left(n\frac{vT_s}{d}\right) \tag{9}$$

여기서 n은 임의의 정수로 거리차 r_d는 거리차 해상도의 정수배 값이 된다는 것을 의미한다.

마이크간 거리가 34.5cm, 표본화율이 44.1KHz인 경우에 Δr_d는 0.7789cm가 되며 n은 ±44의 범위를 갖는다. n에 따른 추정 각도의 해상도를 그림 6에 보였다. 균등한 거리해상도 값으로부터 각도 추정 오차는 정현파의 특성상 ±90도 근방에서 오차가 커진다. 이를 개선하기 위해 표본화율을 높이는 것이 일반적인 방법이나 한계가 있다.

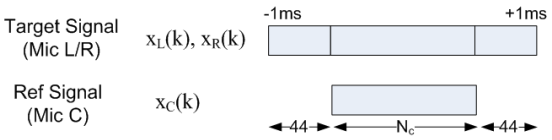


Fig. 5 Correlation data structure for 1ms TDoA

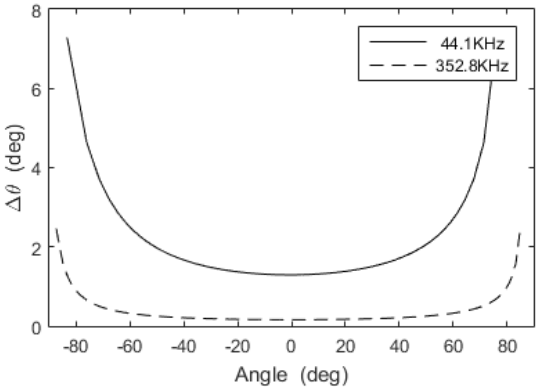


Fig. 6 Resolution of angle estimation

표본화율을 8배인 352.8KHz로 변경해도 그림 6에서처럼 ±90도 근처에서는 여전히 2도 이상의 오차를 보이고 있다. 이 문제를 해결하고 0~360도의 전방위에서 가능한 음원 추적 방안을 살펴보자.

그림 6의 각도 추정 오차는 -45~45도 범위에서 작은 값 보이므로 이 영역을 활용하겠다. 또한 각도 추정에서 sin이나 cos으로 각도를 구하면 0~180도 까지만 판별하게 된다. 이와 같은 특성을 이용하여 두개의 마이크 세트를 그림 7과 같이 직교 배열하면 서로 다른 각도를 검출하게 되며, 이를 조합하여 전방위로 측정이 가능하게 된다. 마이크 세트 1번은 바깥쪽 실선의 방위각을, 마이크 세트 2번은 안쪽 점선을 따라 방위각을 측정하게 된다. 각 마이크에서 측정된 값으로부터 표 1과 같이 영역 A~D를 결정할 수 있다. 오차를 줄이기 위해 최종 음원 추정 방위각은 두 개의 마이크 세트 중에서 45~135도의 범위를 검출한 것을 최종 값으로 사용한다. 이 범위는 그림 7의 가장 안쪽에 있는 원이며 각 마이크 세트의 영역을 각각 구분하여 표시했다. 이 각도 범위에서는 그림 의 결과에 따라 2도 이하의 오차를 보일 것이다. 실제 구현에서 직교배열의 중앙 마이크는 공통이므로 총 5개의 마이크로 성된다.

상관의 연산은 기준파형의 길이가 N_c라 할 때 N_c^2의 곱셈이 필요하다. 제안한 상관기는 기준 신호를 부호만 취하여 계산량을 줄이고 있다. 예를 들면 마이크 L에 대한 상관은 다음과 같이 적용하였다.

$$R_L(m) = \sum_{k=0}^{N_c-1} sgn(x_c(k))x_L(k+m) \qquad (10)$$

Table. 1 Omni-directional angle estimation with dual Mic sets shown in Fig. 7

Region	Mic Set 1	Mic Set 2	Decision
A	0~90	0~90	0~90
B	90~180	0~90	90~180
C	90~180	90~180	180~270
D	0~90	90~180	270~360

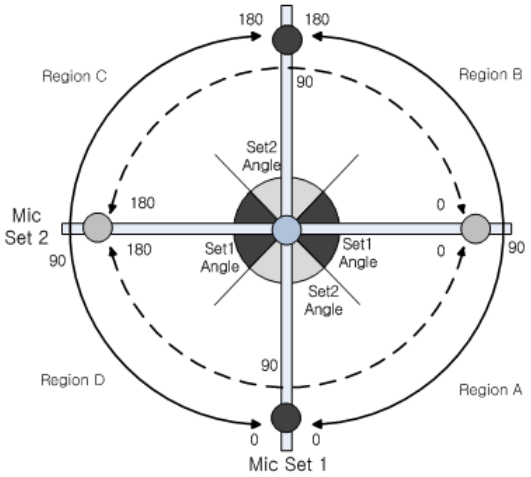


Fig. 7 Omni-directional angle estimation

기준 신호가 충분히 신뢰를 가질 경우 곱하기를 사용하는 것과 유사한 성능을 보인다. m은 정수이며 지연값으로 1ms의 TDoA를 검출하는 경우 앞서 설명한 바와 같이 ±44의 범위를 갖는다.

이를 위하여 일정 크기이상의 음성 구간 데이터가 기준 신호가 되도록 음성 활성도 검출(VAD: Voice Activity Detection) 기법에서 사용하는 짧은 구간의 에너지를 이용하였다[5, 10].

$$E_{VAD} = \sum_{m=0}^{N_E-1} x_c^2(m) \qquad (11)$$

여기서 N_E는 100을 사용했으며, 실제 연산은 이동평균 기법으로 매번의 1번의 곱셈과 덧셈 및 뺄셈만으로 구현이 가능하다. 일반적인 VAD는 음성을 유성음/무성음/무음으로 구분하기 위해 여러 가지 방식이 사용되고 있다. 강의시스템에서는 비교적 노이즈가 적은 환

경에서 강사의 음성을 검출하는 것이 목표이므로 간단히 신호의 에너지를 검출하여 이를 활용한 것이다.

기준신호에 무음구간이 포함될 경우에 TDoA의 검출 오류로 급격한 각도 변화가 발생할 수 있다. 에너지 기준값을 올리면 작은 소리는 검출할 수 없으며, 기준값을 낮추면 많은 오류가 발생 할 수 있다. 본 논문에서는 이러한 오류를 줄이기 위하여 임펄스 잡음에 우수한 성능을 갖고 있는 미디언 필터(Median Filter)를 적용하였다. 길이가 7인 미디언 필터를 사용하여 7개의 관측값들 중에서 1~2개의 오류가 발생해도 일정한 값을 유지함을 확인하였다. 이로서 강의 시스템에서 필요한 안정된 검출 성능을 얻을 수 있었다[11].

Ⅳ. 시스템 구현 및 실험 결과

그림 8은 개발된 시스템의 신호처리부 블록도이다. 멀티미디어 강의 시스템에서의 음원 추적은 강사의 이동 속도가 통상의 걷기 수준인 4Km/h로 한정된다. 따라서 10ms 동안의 이동 거리는 1.11cm이므로 각도 변화는 거의 없기 때문에 두 개의 마이크 세트는 시분할로 데이터를 획득하도록 구성하였다. 마이크 모듈은 MEMs 마이크와 12비트 ADC를 내장한 Digilent사의 Pmod MIC3 모듈을 사용했으며, 직렬통신인 SPI(Serial Peripheral Interface)로 MCU와 연결된다. 두개의 마이크 세트는 GPIO핀으로 CS(Chip Select) 핀으로 제어하여 구분된다. 블록도에서 마이크 모듈을 제외한 나머지 기능은 MCU에서 처리한다.

MCU는 TI사의 TMS320F28379D이며 200MHz의 CPU와 독립적인 연산 모듈인 CLA(Control Law Accelerator)를 각각 2개씩 지원하고, 1024KB 플래시 메모리와 204KB RAM을 내장하여 단일 칩으로 고속 연산을 구현할 수 있는 강점이 있다. 또한 3개의 SPI를 지원하여 본 시스템에 적합한 기능을 대부분 갖추고 있다[12]. 중앙 마이크의 신호에서 VAD_E를 측정하여 기준값 이상인 경우에 전체 마이크의 신호를 저장하도록 동기화한다. 중앙 마이크 신호는 자기상관으로 신호의 신뢰성을 검증하고, L과 R의 마이크 신호와 상호 상관으로 TDoA를 검출한다.

최종 구현된 시스템은 그림 9와 같다. 그림 9(a)의 DSP보드는 향후의 성능 개선을 위하여 2개의 MCU와 1개의 Altera Cyclone III FPGA, TI사의 개발용 키트를 연결하기 위한 소켓을 탑재하였다. 직교배열 마이크는 그림 9(b)와 같이 평면상에 직교로 배치했으며 음원 추적 결과는 (c)와 같이 외부에 장착된 OLED에 숫자로 표시된다.

표 2는 0~360도에서 1m의 거리, 60dBC의 수신 음향 전력레벨에서 45도 간격으로 음원의 위치를 변화하며 실험한 결과이다. 함수발생기로 백색잡음을 발생시켜 실험한 경우에는 0.5도 이하의 오차로 정확히 검출되고 있다. 음성 신호에 대해서는 강사의 음성과 유사하도록 인터넷 강의와 토론 영상을 사용하였으며, 측정 결과는 6.5도 이내의 오차로 검출되었다. 45도에서 오차가 가장 큰 것은 두 마이크 세트의 유효영역 경계치로 인해

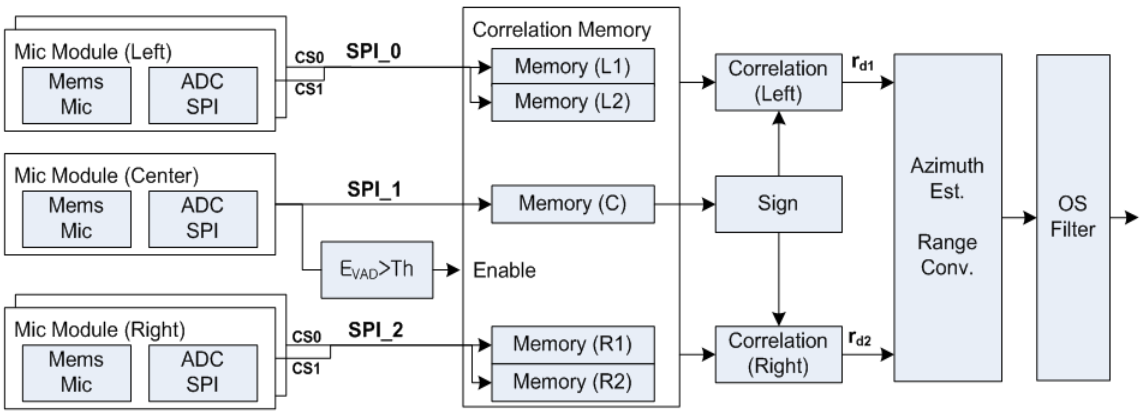
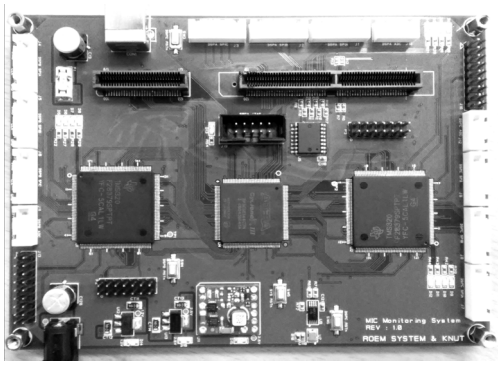


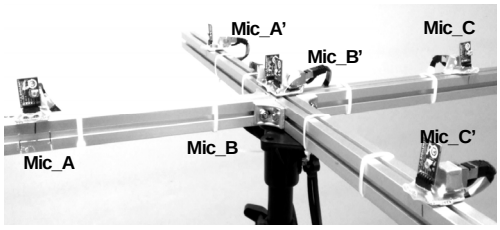
Fig. 8 Block diagram of system

나타난 것이다. 음성 신호는 일정한 피치를 갖고 있기 때문에 시간영역에서 정현파와 비슷하게 관찰된다. 신호의 주기성으로 상관이 다른 위치에서 검출될 수 있으므로 표 2와 같이 변화폭을 보이게 된다.

표 3은 음원을 270도, 1m에 고정하고 수신 전력을 변화하며 측정한 결과이다. 일반적인 강의 환경에서 신호의 크기에 무관하게 동일한 결과를 보여주고 있으나 54dBC이하와 86dBC 이상에서는 검출하지 못하였다. 이것은 마이크의 감도가 낮아서 작은 소리를 검출하지 못한 결과이며, 큰 소리의 경우는 벽면의 반사파가 많아지고, 수신단에서 최대치로 포화되기 때문에 일정한 상관값을 얻기 어려워 나타난 것으로 확인되었다. 정상 동작 영역인 55~85dB는 일반적인 강의 시스템에서 사용하기에는 문제가 없는 것으로 판단된다.



(a)



(b)

(c)

Fig. 9 Pictures of System: (a) DSP board (b) Mic array (c) Result Display

Table. 2 Experiment: varying azimuth with 1m distance, 60dBC sound level

Azimuth (deg)	White Nose	Lecture Voice
0	0.0	-0.8~0.7
45	45.3	44.1~51.5
90	90.5	90.7~93.6
135	135.1	134.5~136.8
180	180.0	180.7
225	225.0	226.0~227.0
270	270.0	270.0~270.7
315	315.1	312.4~315.3

Table. 3 Experiments with varying sound level at 270° azimuth, 1m distance

Level (dBC)	White Nose	Lecture Voice
~54	Undetected	Undetected
55~85	270.0	267.3~273.5
87~	Undetected	Undetected

Ⅴ. 결 론

본 논문에서는 멀티미디어 강의시스템에 적합한 음원추적 방안을 제안하였다. 교단의 천정에 직교배열 마이크를 부착하여 전방위에서 강사의 위치를 추적하여 자동 로밍 마이크를 위한 제어신호를 제공하는 것이다.

음성의 TDoA를 점 음원과 평행 음원 모델로 분석하고, 3개의 마이크 세트로 기준신호에 대하여 상호상관 방안을 제안하였다. 무음구간을 제거하기 위하여 수신 에너지를 이용하였으며, 상호 상관을 부호로 연산하여 계산량을 줄이고, 추정 각도에 미디언 필터를 적용하여 안정된 각도를 얻었다. 전방위 검출을 위하여 평면에 십자 배열된 5개의 마이크로 음원 추적이 가능함을 보였다. 제안된 시스템은 고속 MCU인 TMS320F379D와 12비트 ADC가 내장된 MEMs마이크 모듈로 구현하여 실시간 동작을 확인하였다. 백색잡음과 음성으로 실험하여 각각 0.5도와 6.5도 이하의 오차를 확인하였다.

이를 멀티미디어 강의 시스템의 강사 추적에 활용하여 다수의 마이크로 자동 로밍이 가능해지므로 강사는 마이크 없이 자유롭게 움직이는 것이 가능해진다. 또한 2개의 마이크를 더 추가하여 x-y-z 축으로 확장하면 방

위각과 고도각(Elevation Angle)을 추정이 가능할 것이다. 마이크 로밍과 3D 공간 추적은 향후의 연구과제로 진행 중에 있다.

ACKNOWLEDGMENTS

This paper was supported by Research Fund, Kumoh National Institute of Technology

REFERENCES

[1] J. Murray, S. Wermter and H. Erwin, “Auditory robotic tracking of sound sources using hybrid cross-correlation and recurrent networks,” in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005)*, Edmonton, AB, Canada, pp. 3554-3559, Aug. 2005.

[2] B. Lee and B. Choi, “Spherical localization of sound source using triangular microphone Array,” in *Proceedings of the Korean Automatic Control Conference (KACC 2009)*, Institute of Control, Robotics and Systems, Busan, Korea, pp.360-363, 2009.

[3] J. Han, S. Han, and Lee, “Sound source tracking control of a mobile robot using a microphone airay,” *Journal of Institute of Control, Robotics and Systems (Korea)*, vol. 18, no. 4, pp. 343-352, Apr. 2012.

[4] S. Oh and K. Park, “Optimal acoustic sound localization system based on a tetrahedron-shaped microphone array,” *Journal of Korea Information Scientists and Engineers*, vol. 43, no. 1, pp. 13-26, Jan. 2016.

[5] J. Ramirez, J. M. Górriz, and J. C. Segura, *Voice activity detection. fundamentals and speech recognition system robustness*. I-TECH, Vienna, Austria, 2007.

[6] Y. A. Huang, J. Benesty (Eds.), *Audio signal processing for next-Generation multimedia communication systems*, Kluwer Academic Publishers, Boston, USA, 2004.

[7] J. Kim and D. Eom, “TDoA-based practical localization using precision time-synchronization,” *The Journal of Korean Institute of Communications and Information Sciences*, vol. 38, no. 2, pp. 141-154, Feb. 2013.

[8] S. Lee and H. Choi, “On the speaker"s position estimation using TDOA algorithm in vehicle environments”, *Journal of Digital Contents Society (Korea)*, vol. 7, no. 2, pp. 71-79, Apr. 2016.

[9] R.C. Luo, C. H. Huang, and T. T. Lin, “Human tracking and following using sound source localization for multisensor based mobile assistive companion robot,” *in Proceedings of IECON 2010-36th Annual Conference on IEEE Industrial Electronics Society*. Glendale, AZ, USA, IEEE, pp. 1552-1557, 2010.

[10] M. H. Moattar and M. M. Homayounpour, “A simple but efficient real-time voice activity detection algorithm,” *in Proceedings of 17th European Signal Processing Conference* (EUSIPCO), Glasgow, Scotland, IEEE, pp.2549-2553, Aug. 2009.

[11] L. Rabiner, M. Sambur, and C. Schmidt, “Applications of a nonlinear smoothing algorithm to speech processing,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 23, no. 6, pp. 552-557, Dec. 1975.

[12] Texas Instruments Inc. TMS320F2837xD Dual-core Delfino Microcontrollers: Technical Reference Manual. [Internet]. Available: http://www.ti.com/lit/pdf/spruhm8.

강민수(Minsoo Kang)

2015년: 금오공과대학교 전자공학부 (학사)
2015년~현재: 금오공과대학교 전자공학과 (석사과정)
※관심분야 : 신호처리, DSP 설계

오우진(Woojin Oh)

1996년: 한국과학기술원 전기및전자공학과 (공학박사)
1996년~1998년: SK 텔레콤 중앙연구원 (선임연구원)
1998년~현재: 금오공과대학교 전자공학부 (교수)
※관심분야: 신호처리, DSP 및 FPGA 시스템 설계 및 구현, 차량용 레이더 신호처리