

마이크로폰 배열로 발생하는 입력 시간차를 이용한 음원의 방향 추정 장치에 관한 연구

윤준호[#], 최기훈*, 유재명**

A Study about Direction Estimate Device of the Sound Source using Input Time Difference by Microphones' Arrangement

Jun Ho Yoon[#], Gi Hun Choi* and Jae Myung Yoo**

ABSTRACT

Human uses level difference and time difference to get space information. Therefore this paper shows that method to presume direction of sound source by time difference and to mark presumed position. The position means direction from geometrical center of sensors to the sound source.

To get the time difference of microphones input level, we will be explained about arrangement of microphones which used for the sensor to take the sound signal. It is included distance among the 3 microphones and distance between microphones and sound source. Secondly, input signals are transmitted to CPU through digital process. CPU is used to DSP(Digital Signal Processor) for manage the signal by real time. Finally, the position of sound source is perceived by an explained algorithm in this paper.

Key Words : Microphone (마이크로폰), Sound Source (음원), DSP (디에스피:Digital Signal Processor), Time Difference (시간차), Open angle(벌림각도)

기호설명

- l = Distance between two microphones [m]
- d = Difference of distances from source to two microphones [m]
- θ = Direction angle of sound source [deg]
- v = Sound speed [m/sec]
- f = Frequency [Hz]

1. 서론

미래의 많은 기계 시스템은 지금보다 우리의 생활과 더욱 가까운 환경에서 활동하게 될 것이다. 그래서 많은 기계 시스템은 인간다움이 요구되어 지고 있다. 따라서 시각, 청각, 촉각 등의 감각기능을 가진 감성 시스템은 매우 매력적이다.

감각 기능 중 청각은 인간과 시스템사이에서 사용되거나 음성이나 단어들을 구별하여 의사 소

[#] 접수일:2003 년 4 월 25; 게재승인일:2004 년 3 월 12 일
[#] 교신저자: 여주대학 자동차과
 E-mail: jhyoon@mail.yeojoo.ac.kr, Tel: (031) 880-5271
 * 중앙대학교 미래신기술 연구소
 ** 중앙대학교 기계공학부

통을 위한 목적으로 사용¹하거나, 음원을 인지하여 음의 위치를 판별하는 음원 정위의 두 가지 기능을 가지고 있다.

이 중 음원 정위에 관한 연구는 1981년 Ruta²가 관측점과 음의 파형의 관계식을 행렬식으로 사용하여 수학적인 모델링을 제시하였으며, 복수 음원 정위 시스템은 음원을 분리하는 방법으로 인간 음성의 스펙트럼 주파수가 시각에 따라 필요없는 부분을 배제 하는 주파수 검출법을 Murai³가 연구하였으며, 1986년 Nagada와 Ohno^{4,5}는 잡음제거 방법을 연구하였다. 그 중 Nagada는 복수음원의 스펙트럼의 위치 결정을 다수의 마이크와 계산 시간이 필요한 최소 직탐방법에 의해 구하는 방법으로 음원을 분리하였다. 그 후 1988년 Sugie는 반사환경에서 단일 음의 파형을 분석하였다.⁶ 1989년 Itow는 다수의 마이크를 이용하여 공간정보를 측정하였다.⁷ 이후 1995년 노이즈가 발생하는 공간에서 휘파람 소리를 인식하여 방향을 검출하는 장치가 Kanagisawa⁸에 의해 개발되어졌으며, 마이크로폰에서 측정된 신호의 스펙트럼에 나타나는 피크의 위상차를 이용하여 음원을 분리하였다.⁹

인간은 두 귀에 입사되는 음의 레벨차와 시간차를 이용하여 공간상에서 음원의 정보를 얻으며, 이렇게 얻어진 음원 정보를 통하여 주변의 환경을 인식할 수 있다.

따라서 본 연구에서는 인간의 귀와 같은 역할을 할 수 있는 마이크로폰을 공간상에 설치하여 마이크로폰에 입력되는 음의 시간차이를 이용하여 음원의 위치를 추정할 수 있는 장치를 제작하고, 추정된 위치를 표시할 수 알고리즘을 설명하고자 한다. 이 장치는 마이크로폰을 사용하여 공간에서 입력되는 음을 입력받을 수 있는 입력부와, DSP(Digital Signal Processor)를 사용하여 입력된 음의 정확한 정보를 실시간(real time)으로 처리하기 위한 중앙처리장치, 그리고 DSP에서 처리된 정보를 표시하기 위한 표시부로 구성되어 있다.

이러한 음원 정위 능력을 자율이동로봇에 부여한다면 로봇이 처해있는 환경을 인식할 수 있으며, 어두운 사고현장에서 감지되는 소리만으로도 원하는 소리를 파악하여 위치를 찾아낼 수 있는 인명 구조용으로써 활용할 수 있을 것이다. 게다가 장애물, 복수의 음원, 노이즈가 함께 존재하는 산업현장에서도 방향판단이 가능하며, 원하는 음

파를 인식하며 특정 명령을 받을 수 있을 것이다.

2. 공간음향

2.1 공간인식

3차원 음향 시스템은 주어진 음원과 관련하여, 본래 존재하는 공간속성을 보충하거나 대처하는 과정을 이용하는 것으로 공간지각과 관련된 파라미터의 분류법을 Fig. 1에 도시한다.

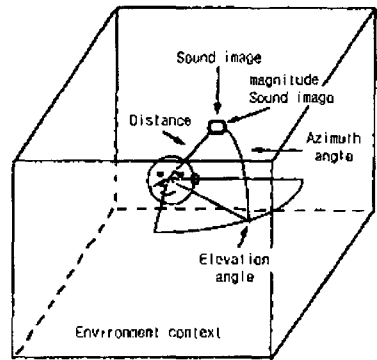


Fig. 1 Terminology of spatial acoustic

Fig. 1에서 알 수 있듯이, 기본적으로 공간지각은 자기 중심 기준계와 관련된다. 즉, 음상의 측정과 방향은 청취자의 위치가 기준이 된다.¹⁰

2.2 음원의 방향 측정

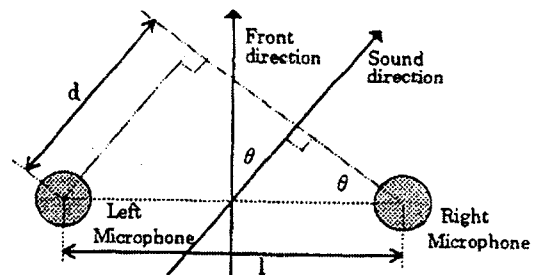


Fig. 2 Relations between d , l , θ

Fig. 2와 같이 정면에 대하여 각도 θ 도 방향으로부터 음이 입사하는 경우

$$\sin \theta = \frac{d}{l} \quad (1)$$

으로 방향각을 알 수 있다. 또한 음원에서 마이크로폰까지의 거리차 d 는 다음과 같은 관계가 있다.

$$d = \frac{v}{2f} \quad (2)$$

만약 두 마이크로폰 사이의 중심을 O 점이라 하고 O 점의 음의 세기(intensity)를 I, 음원으로부터의 O 점까지의 거리를 r 이라 하면, 두 마이크로폰 사이의 세기차 ΔI 는 다음 식으로 구할 수 있다.^{10,11}

$$\Delta I = I \left(\frac{1}{r - \frac{d}{2}} - \frac{1}{r + \frac{d}{2}} \right) \quad (3)$$

3. 실험장치 및 방법

3.1 실험장치 구성

본 연구에서 사용된 실험 장치는 크게 세 부분으로 분류한다. 우선 마이크로폰에서 소리 신호를 입력 받아 신호를 증폭하는 입력부가 있다. 여기서 사용된 마이크로폰은 무지향성 콘덴서 마이크로폰이다. 두 번째는 증폭된 신호를 디지털 값으로 변환하고, 이 변환된 값을 분석하고 처리하여 저장하는 중앙처리부로서 본 연구에서는 실시간 처리를 위해 DSP TMS320C31 를 사용하였다. 마

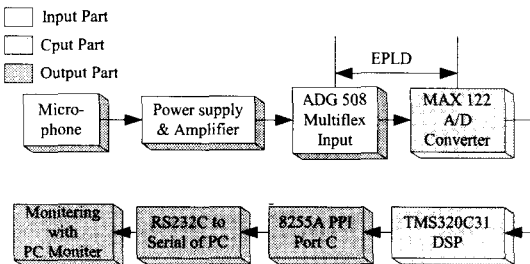


Fig. 3 Block diagram of the system

지막으로, DSP 에서 처리되고, 연산된 데이터 값을 컴퓨터로 모니터링을 할 수 있는 출력부로 구성되어 있다. Fig. 3 에 실험 장치의 블록 다이어그램을 간략히 나타낸다.

3.1.1 마이크로폰 입력부 및 증폭부

인간의 두 귀에 해당되는 것으로 Fig. 4 에서 보듯이 신호 입력단(1)과 전원 공급단(2), 그리고 신호를 증폭(3)하는 장치를 내장하고, 또한 마이크로폰의 입력 신호의 크기를 조정할 수 있는 볼륨(4)을 부착하였다. Output(5)에서는 증폭된 신호를 중앙처리부로 전송한다.

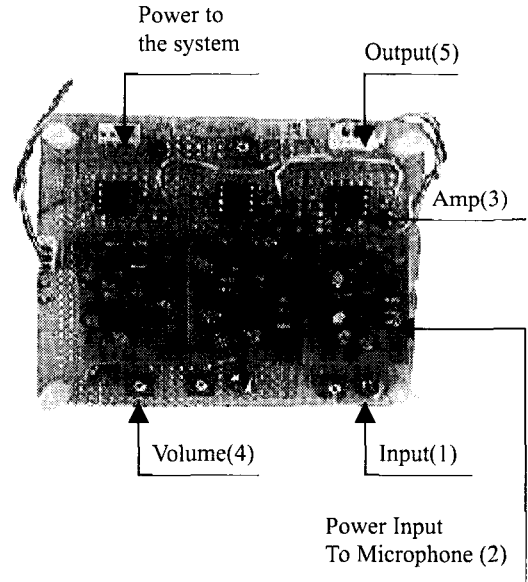


Fig. 4 Input and amplifier of microphones

3.1.2 A/D 컨버터

본 연구에서는 분해능이 비교적 높고 변환 속도가 매우 빠른 12 비트 병렬 접속형 A/D 컨버터를 선택하였다. 이 A/D 컨버터는 축차비교형으로 최고 5MHz 의 클럭을 사용하여 13 클럭주기(2.6 μ s)에 변환을 완료하며, 샘플링 속도는 333[kHz]이다. 그리고 샘플/홀드 회로를 내장하고 있고, 아날로그 입력전압의 범위는 -5[V] ~ +5[V]이다.¹² 또한 변환된 데이터는 12 비트 2 의 보수로 표현되

어 +5[V]가 7FFH, 0[V]가 000H, -5[V]가 800H 로 입력된다.

3.1.3 중앙처리부

DSP(Digital Signal Processor)는 디지털 신호처리 분야에서 실시간으로 매우 빠른 연산을 수행할 수 있도록 개발된 마이크로프로세서이다.

본 연구에서 사용된 DSP 는 TMS320C31 로서, A/D, D/A, 멀티플렉서 그리고 PC 와의 직렬통신 포트를 제공한다. 또한 사용자 프로그램의 입력 및 실행을 위하여 IBM PC 의 병렬 포트용 콘넥터를 내장하고 있으므로, 별도의 라이터기 없이 사용할 수 있도록 설계 되어져있다.¹²

3.1.4 출력부

컴퓨터의 모니터로 방향을 표시하기 위해 출력된 신호는 컴퓨터의 COM 포트와 RS-232C 통신을 한다. 사용된 인터페이스는 8 비트 PIC 로서 본 연구에서는 PIC16F84 와 MAX232C 를 사용한다.

PIC 의 Port B 로 입력된 PPI8255A 의 3 비트 데이터의 상태를 Port A 로 출력하고, 출력된 신호는 MAX232C 를 거쳐 전압 레벨을 변환한 후 컴퓨터의 Com 1 포트로 전송된다. Fig. 5 는 출력부에 관한 세부적인 블록 다이어그램이다.

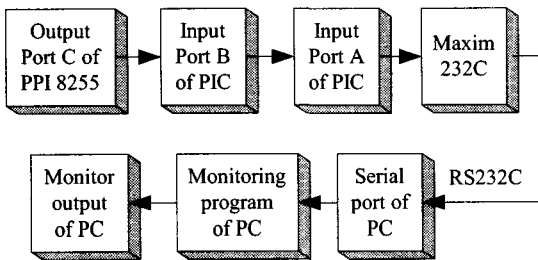


Fig. 5 Block diagram of Output part

3.2 실험방법

3.2.1 마이크로폰 배열

마이크로폰의 배치를 결정하기 위해 첫번째로 음원과 마이크로폰의 거리를 측정하였다. 본 실험에서는 공간에서 발생하는 음원의 크기를 60dB, 70dB, 그리고 80dB 로 변화시켜가며 음원과 마이크로폰의 거리관계를 측정하였다.

두 번째 실험에서는 마이크로폰끼리 간섭받지

않는 최소 거리를 측정하는 실험으로 2 개의 마이크로폰의 거리를 변화시켜 측정하였다. 이것은 마이크로폰에 입력되는 파형의 시간차와 레벨차를 측정하기 위한 것으로서, 공간지각에 매우 중요한 단서로 사용할 수 있다. 이 때 거리는 1, 5, 10, 15, 20, 25, 30, 40 cm 로 변화된다.

그리고, 마지막으로 두 번째 실험에서 측정된 거리를 유지하면서 마이크로폰을 회전시켰을 경우, 즉 벌림 각도(open angle)가 존재하는 경우에 측정되는 시간차를 두번째 실험과 비교하였다. 여기서 벌림 각도란 마이크로폰의 축이 중심축으로부터 시계 혹은 반시계 방향으로 회전되어 있는 경우 각도가 존재하며, 두 마이크로폰의 축 각도이다.

3.2.2 디지털 처리

3 개의 마이크에는 각각에 볼륨을 달아 입력신호의 크기를 조정할 수 있다. 볼륨은 100[kΩ]의 가변저항을 사용하였다. 입력 신호는 앞의 입력부에서 보여진 OP-Amp 로 증폭을 거쳐 -5[V] ~ +5[V]까지의 레벨을 가진다.

마이크로폰에 입력된 3 개의 신호는 아날로그 멀티플렉서를 통해 A/D 컨버터로 보내어져 디지털 값으로 변환된다. 본 연구에서는 ADG508 아날로그 멀티플렉서에서 0.05 초마다 3 개의 마이크로폰을 순서대로 선택하여 A/D 컨버터로 입력이 되고, 입력된 신호를 디지털 값으로 변환한다. 이때의 변환 시간은 대략 2.6 μs이다. 그러나 본 연구에서 사용된 변환 시간은 인터럽트 발생 루틴을 이용해서 변환이 완료될 때까지 대기한다. 이것은 변환되는 시간에 관계없이 입력 신호에 따라 빠르게 디지털 값으로 변환을 하기 위한 것으로서, 만약 입력된 신호가 2.6 μs보다 빠르게 변환이 완료되면, 나머지 시간은 대기하도록 하여, 최적의 변환을 하도록 하였다.

디지털 값으로 변환된 데이터들은 3 개의 레지스터_0 에 저장되며 레지스터_0 에 저장된 값들을 비교하여, 가장 큰 값을 새로운 레지스터_1 에 저장한다. 여기서 두 번째 인터럽트를 사용하여 레지스터_1 에 디지털 값으로 저장된 값들과 새로 들어오는 디지털 값을 다시 비교한다. 즉, 출력되고 있는 값과 0.05 초마다 샘플링 되어 새로 입력되는 값을 다시 비교하여 최종적으로 가장 큰 값을 출력함으로써 소리가 나는 방향, 음원의 방향을 추정할 수 있다.

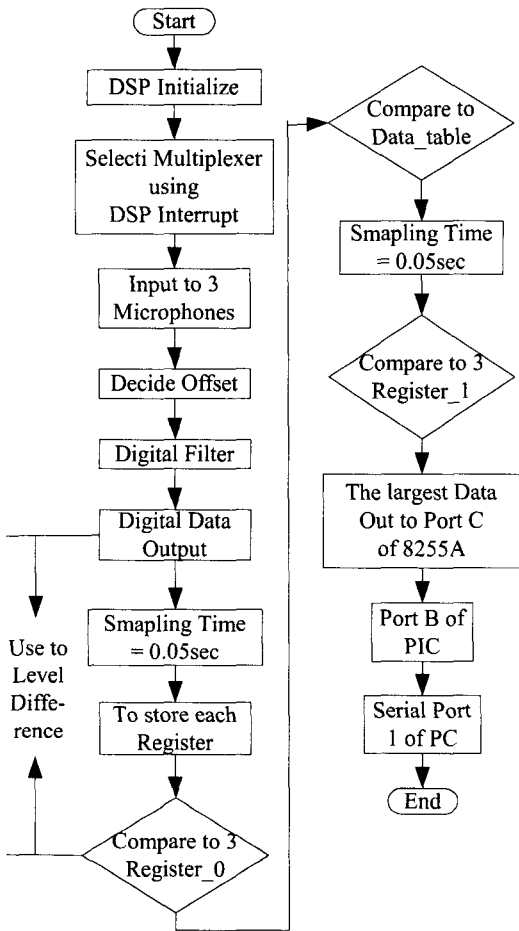


Fig. 6 Flowchart for direction detect

Fig. 6 은 본 논문의 방향감지 알고리즘을 플로우 차트로 나타낸 것이다. 여기서 오프셋(Offset) 조정은 D/A 컨버터를 0V 로 출력하여 A/D 컨버터로 1000 회를 읽어 평균값을 취하는 것으로써 모든 아날로그 값을 읽을 때마다 이를 감해줌으로써 아날로그 입력 값에서 오프셋을 보상하는데 사용된다. 또한 디지털 필터는 전형적인 디지털 저역 통과 필터를 사용하였다. 여기서는 DSP 에 내장된 Timer 0 을 이용하여 입력된 신호를 비교하여 큰 데이터를 레지스터에 저장하고 Timer 1 을 이용하여 저장된 값과 비교하여 가장 큰 데이터 값을 출력으로 하도록 하였다.

3.2.3 모니터 영상 처리

공간상에 배치된 3 개의 마이크로폰에 의해 입력된 신호는 디지털 처리과정을 거쳐 DSP 로 전송된다. 입력된 각 데이터 값들은 레지스터에 저장되고, 저장된 데이터 값들을 비교하여 가장 큰 값을 PPI8255A 의 Port C 로 출력하고 인터페이스 장치를 거쳐 Com 1 포트로 전송된다.

전송된 신호는 컴퓨터내의 프로그램에 의해 입력된 신호의 방향을 모니터에 Center, Left, Right 의 세 방향중 하나로 표시하고, 이 방향 표시로서 조작자는 음원이 발생된 방향을 알 수 있다.

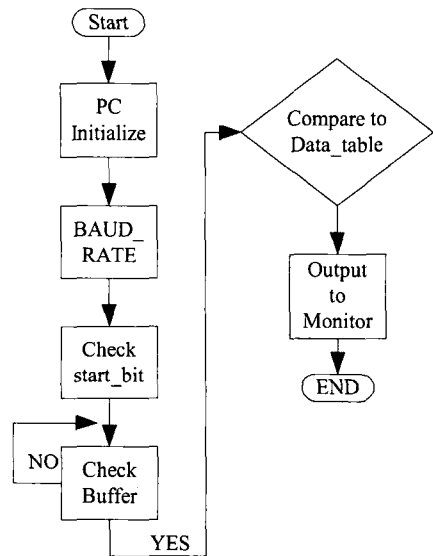


Fig. 7 Flowchart for monitor output

Fig. 7 은 컴퓨터의 Com 1 포트에 입력된 자료를 모니터로 출력하기 위한 프로그램의 간략한 플로우 차트이다

4. 실험결과

먼저 콘텐서 마이크로폰에 입력되는 방향성을 측정 한 결과를 Fig. 8 에 보인다.

콘텐서 마이크로폰에 입력되는 음은 모든 방향에서 거의 유사한 크기의 음이 입력되는 것을 알 수 있다. 또한 전방 0 도와 역방향의 180 도일 때는 Fig. 8 에서 보듯이 큰 차이가 나지 않아 전

방 혹은 후방의 방향을 인식할 수 없음을 알았다.

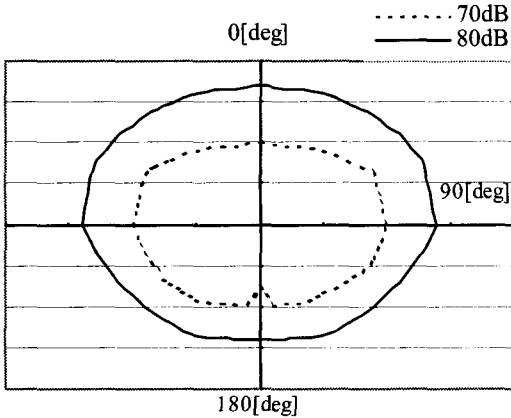


Fig. 8 Microphones' directive on sound magnitude

소음이 존재하는 일상적인 환경에서 제작된 시스템을 이용하여 사람의 말 소리를 인식할 수 있는 거리를 측정하여 Fig. 9 에 나타냈다.

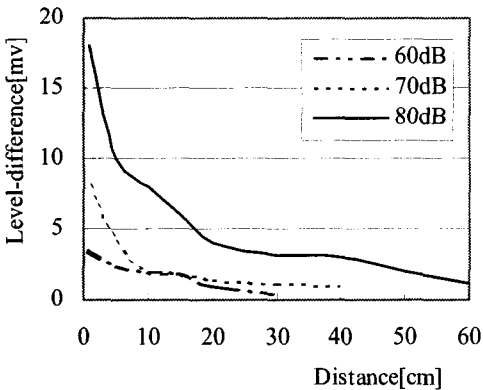


Fig. 9 Distance relationship of sound source and microphone

보통 일상적인 대화를 나누는 사람의 말 소리는 60dB 이다. 그러나 이 크기에서는 음원과 마이크로폰의 거리가 1 cm 이내의 경우에서만 마이크로폰에 측정이 가능한 음이 입력되지만 거리가 너무 짧기 때문에 실제 사용이 불가능한 것으로 판단된다. 음원의 크기가 70dB 인 경우는 15 cm 이내

의 음이 마이크로폰에서 입력된다. 그러나 Fig. 9 에서 보듯이 80dB 의 음은 30~40 cm 떨어져 있는 마이크로폰에서도 측정이 가능한 음을 입력 받을 수 있음을 알았다.

두 마이크로폰의 배열 관계를 알아보기 위하여 70dB 과 80dB 의 음을 발생시켜 두 마이크로폰에 입력되는 시간차이를 측정하여 Fig. 10 에 나타냈다.

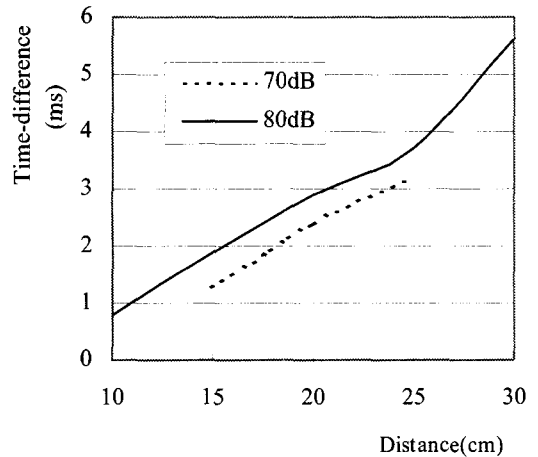


Fig. 10 Inputted time-difference of two microphones

70dB 의 음에서는 15 cm 이내와 25 cm 이상에서는 시간차를 측정하는 것이 불가능했지만 80dB 에서는 10 cm 까지 측정할 수 있었으며 최대 40 cm 까지 마이크로폰에 입력되는 시간차를 측정할 수 있었다.

식(2)에서 입력되는 음의 주파수는 567Hz(80dB)이고, 음파의 속도는 340m/s 일 때 마이크로폰 사이의 거리는 30 cm이다. 이 때의 시간차를 측정된 결과 약 5.6ms 였다.

따라서 본 연구에서는 80dB 음원을 공간상에서 발생시켜 30 cm 떨어진 두 마이크로폰에 입력되는 시간의 차이로 방향을 추정한다.

또한 마이크로폰의 벌림 각도(open angle)가 존재하는 경우와 그렇지 않은 경우의 시간차를 비교하여 Fig. 11 에 나타낸다. 이 때 X 축은 두 마이크로폰의 벌림각도[deg]를, Y 축은 시간차이(ms)를 나타낸다.

벌림각도가 존재하지 않는 경우에 마이크로폰 1 과 마이크로폰 2 에 입력되는 시간의 차이는

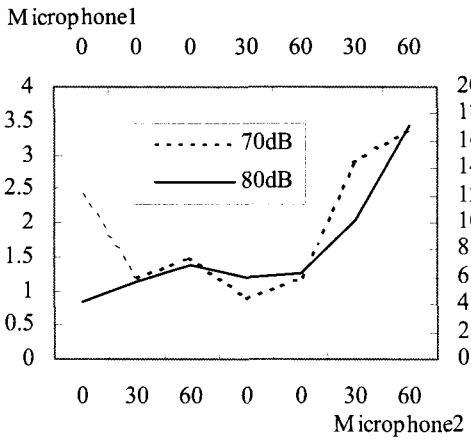


Fig. 11 Direction of each microphone

70dB 에서는 2.4ms, 80dB 에서는 4.2ms 이다. 그러나 벌림 각도가 120 도인 경우는 각각 3.4ms, 17.2ms 로 측정되었다.

즉 두 마이크로폰에 입력되는 시간차는 마이크로폰의 배열과 관련이 있고, 이것은 벌림 각도를 120° 도로 하였을 경우 최대의 시간차를 가진다.

Fig. 12 에 이상의 실험 결과를 종합하여 공간

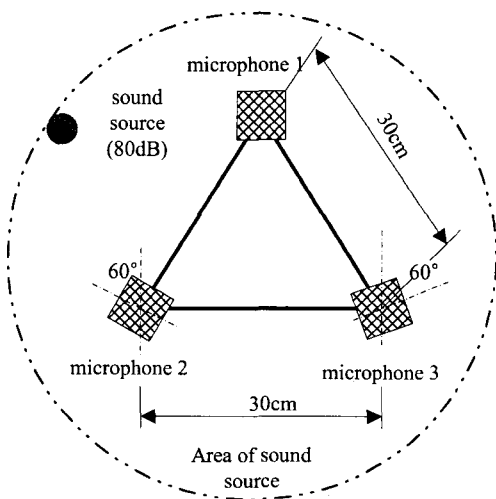


Fig. 12 Range of sound source and distance of each microphones

상에 실제 마이크로폰이 배치된 영역을 간단히 도시한다.

배치된 마이크로폰에 입력되는 음의 세기를 알아보기 위해 80dB 의 음을 여러 각도에서 입력한 후 각 마이크로폰에 입력된 음의 세기를 전압값으로 변환하여 Fig. 13 에 나타낸다. 수행된 데이터들은 100 번 반복되어 입력된 평균값이다.

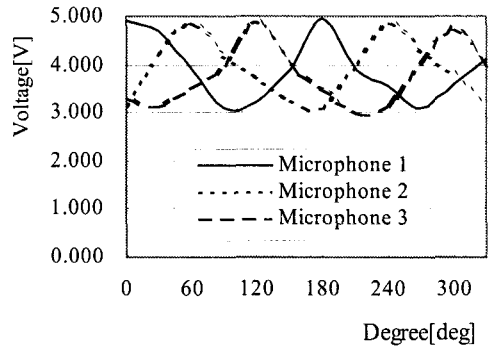


Fig. 13 Inputted digital data to each microphones

Fig. 13 에서 보듯이 여러 각도에서 음이 방사될 때 각 마이크로폰에 입력되는 값들의 차는 1.8V 이내의 매우 적은 시간차를 갖는다.

본 연구에서 제시된 알고리즘에 의해 방향을

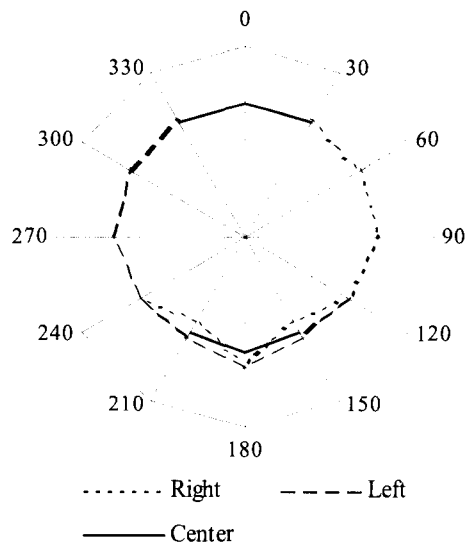


Fig. 14 Direction of sound source

관별한 결과를 측정하여 Fig. 14 나타낸다.

Fig. 14 에서 보듯이 전방부에 음원이 존재 하는 경우에는 상당히 높은 인식률을 보였지만 30 도 근방과 330 도 부근에서 음원이 발생하는 경우에는 전방 혹은 좌우의 방향을 판별하는 것이 불가능하여 실제 모니터로 출력했을 때 ‘Center-Left’ 혹은 ‘Center-Right’가 번갈아 가며 출력이 되었다. 또한 120 도에서부터 240 도 사이에 음원이 입력되는 경우 3 개의 마이크로폰에 입력되는 음의 시간차가 거의 비슷하여 방향 추정이 매우 어려웠다.

5. 결론

본 연구에서는 음원 정위 방법을 이용하여 공간에서 발생하는 음의 방향을 추정하는 방법에 대하여 연구하였다. 음원의 위치와 방향을 판단하는 프로그램을 작성하고 실험 장치를 제작하여 그 실효성을 실험하여 다음과 같은 결론을 얻었다.

1. 본 연구를 통하여, 실시간(Real time)으로 음향신호를 처리할 수 있는 실험장치를 개발하였다.
2. 음원의 대략적인 방향을 판단하는 것은 마이크 간의 신호레벨 비교로 간단히 알 수 있다.
3. 음원의 방향과 거리를 판단하는 인간의 방법은 경험과 학습이라는 차원에서 뉴럴 네트워크를 이용한 학습제어 이론의 도입이 바람직하다고 판단한다.

참고문헌

1. 이종락, "음성인식기술의 현황과 전망," 한국통신 연구 개발단, 1999.
2. 柳田益造, 角所收, "一般的たたみ翔み行列を用いた音原波形分離の定式化," 昭和 56 年 日本音響學會講演論文集 2-7-7, pp. 499 - 500, 1981.
3. 永田仁史, 安部正人, "最小値探索による音源位置とスペクトルの推定," 信學技報, EA86 - 3, 19 86.
4. 村井伸行, 森田龍彌, "複數話者の音聲分離," 第 29 回自動制御連合講演會論文集, pp. 611 - 612, 1986.
5. 小野隆彦, 片山庇之, "クロススペクトル法による固定音源位置の推定における平均回數の最適

化," 信學技報, EA 86 - 25, 1986.

6. 黃 捷, 大西 昇, 杉江 昇, "生體に示唆を得た音源正位システ一反響のある環境での單一音源正位一," 電子情報通信學會論文誌, A J71-A-10, p p.1780-1789, 1988.
7. Itow, T., "Measurement of spatial information in sound fields by closely four point microphone method," J. the Acoust. Soc. Jpn., 10, 2, pp. 101 - 110, 1989.
8. Kanagisawa, Y., Akihisa, O., Vuta, S., "An operator interface for an autonomous mobile robot using whistle sound and a source direction detection system," Proc. of the 1995 IEEE IECON, Vol.2, pp. 1118 - 1123, 1995.
9. 김경환, 김연훈, 곽윤근, "감성로봇을 위한 음원의 위치측정 및 분리," 한국정밀공학회지 제 20 권, 제 5 호, pp. 116 - 123, 2003.
10. 강성훈, "방송음향총론," 기전, pp. 160, 2000.
11. 이순요, 양순모, "가상현실형 감성공학," 청문각, pp. 432 - 433, 1997.
12. 윤덕용, "TMS320C31 마스터," Ohm 사, pp. 213 - 220, pp. 41, pp. 472 - 487, 1998.