

음성 신호를 이용한 시간지연 추정에 미치는 영향들에 관한 연구

Factors for Speech Signal Time Delay Estimation

권병호* · 박영진* · 박윤식*

Byoungho Kwon, Youngjin Park and Youn-sik Park

(2008년 2월 29일 접수 ; 2008년 7월 25일 심사완료)

Key Words : Time Delay Estimation(시간지연 추정), The Generalized Cross Correlation Method(일반화된 상호상관 법), Reverberation(잔향), R/D Ratio(잔향-직접 음 에너지 비, reverberant to direct sound energy ratio)

ABSTRACT

Since it needs the light computational load and small database, sound source localization method using time delay of arrival(TDOA method) is applied at many research fields such as a robot auditory system, teleconferencing and so on. Researches for time delay estimation, which is the most important thing of TDOA method, had been studied broadly. However studies about factors for time delay estimation are insufficient, especially in case of real environment application. In 1997, Brandstein and Silverman announced that performance of time delay estimation deteriorates as reverberant time of room increases. Even though reverberant time of room is same, performance of estimation is different as the specific part of signals. In order to know that reason, we studied and analyzed the factors for time delay estimation using speech signal and room impulse response. In result, we can know that performance of time delay estimation is changed by different R/D ratio and signal characteristics in spite of same reverberant time. Also, we define the performance index(PI) to show a similar tendency to R/D ratio, and propose the method to improve the performance of time delay estimation with PI.

1. 서 론

로봇의 대한 연구가 활발히 진행되면서, 사람과 로봇의 원활한 상호작용을 위해 화자의 위치를 인식할 수 있는 음원 위치 추정 방법에 대해 많은 연구가 진행되고 있다. 이런 음원 위치 추정 방법에는 로봇의 머리전달함수(head related transfer function) 데이터베이스를 이용한 방법⁽¹⁾, 다수의 마이크로폰을 이용한 빔포밍(beam-forming) 방법⁽²⁾ 등이 제안되고 있다. 머리전달함수를 이용한 방법은 최소한의 마이크로폰만을 이용할 수 있고, 빔포밍을

이용한 방법은 다화자의 위치를 동시에 추정할 수 있다는 장점이 있지만, 저가의 로봇 생산을 위해 개발되고 있는 사·청각 신호처리용 SoC(system on a chip)와 같은 적은 계산량과 데이터베이스만을 요구하는 시스템에 구현하기에는 적절하지 않다. 이와 같이 적은 계산량이 요구되는 경우에는 대부분 도달 시간지연(time delay of arrival, TDOA)을 이용한 음원 위치 추정 방법이 적용된다^(3,4). 이 방법은 소수의 마이크로폰에 음향 신호가 도달되는 시간지연 값으로부터 음원의 위치를 추정하는 방법으로 로봇의 청각 시스템뿐만 아니라, 자동 원격 회의를 위한 기술 등 많은 분야에 적용되고 있다⁽⁵⁾.

도달시간지연을 이용한 음원 위치 추정 방법에서 가장 중요한 핵심은 마이크로폰들에서 측정된 신호를 이용해 시간지연 값을 정확하게 추정하는 것이

† 교신저자 : 정희원, 한국과학기술원 기계공학과

E-mail : yjpark@kaist.ac.kr

Tel : (042)350-3036, Fax : (042)350-8220

* 정희원, 한국과학기술원 기계공학과

다. 시간지연을 추정하는 방법들은 예전부터 많이 연구되어 왔는데, 1976년에 Knapp와 Carter가 제시한 주파수 영역에서의 가중치 함수를 적용해 상호상관값(cross correlation)을 계산하고 이로부터 시간지연을 추정하는 방법이 가장 널리 사용되고 있다⁽⁶⁾. 뿐만 아니라 최소자승법을 이용하는 방법, Hilbert transform을 이용하는 방법, wavelet transform을 이용하는 방법 등도 연구되었다⁽⁷⁻⁹⁾.

앞서 설명한 시간지연 추정방법들은 이상적인 신호의 경우에 대해 수학적으로 유도되었기 때문에 실제 환경에서 측정된 신호를 이용해 시간지연을 추정할 경우에는 추정 성능이 나빠지게 된다. 그러므로 시간지연 추정 시 그 성능에 영향을 미치는 요인들에 대해 알아보고, 그 특성들을 고려하여 시간지연을 추정함으로써 추정 성능의 정확성을 높일 필요가 있다.

따라서, 이 연구에서는 시간지연 추정 성능에 영향을 미치는 요인들에 대해 공간의 잔향 특성과 측정된 신호의 특성 관점에서 살펴보기로 한다. 공간의 잔향 특성에 따른 시간지연 추정 성능은 측정된 신호에 직접 음(direct wave) 에너지와 잔향 음(reverberation wave) 에너지를 나타내는 R/D 비를 이용하고, 측정된 신호의 특성은 그 신호의 주파수 특성과 신호의 크기변화에 따라 나타나는 시간지연 추정 성능에 대해 살펴보고자 한다.

마지막으로, 공간의 잔향 특성을 나타내는 R/D 비는 실제 측정된 신호들로부터 추정될 수 없으므로, 이와 유사한 경향을 나타내는 성능지표(PI, performance index)를 정의하고 이를 이용하여 시간지연 추정 성능의 신뢰성을 높이는 방법을 제안하도록 한다.

2. 시간지연 추정 방법

이 논문에서 적용하고 있는 시간지연 추정 방법은 두 신호의 상호스펙트럼밀도함수(cross spectral density function)에 가중치 함수를 곱해 좀 더 정확한 상호상관값으로부터 시간지연을 구하는 일반화된 상호상관값(generalized cross correlation, GCC) 방법이다⁽⁶⁾. 이 방법에 적용될 수 있는 가중치 함수 중에서 잔향의 영향에 가장 강인하다고 알려진 PHAT(phase transform) 가중치 함수를 적용한

GCC 방법을 이용하여 측정된 신호의 시간지연을 추정한다.

이때, 음원 신호는 주로 사람의 음성(speech)이기 때문에 주파수 특성이 시간에 따라 달라지며, 측정 공간이 자유음장공간(free field)이 아니기 때문에 잔향의 영향도 존재하게 된다. 이와 같은 다양한 요인들에 의해 이 논문에서 적용하고자 하는 시간지연 추정방법은 추정 성능에 영향을 받게 된다. 따라서 PHAT 가중치 함수를 이용한 GCC 방법을 이용하여 시간지연을 추정할 때 그 성능에 영향을 미치는 요인들에 알아보도록 한다.

3. 공간의 잔향 특성에 따른 영향

시간지연 추정 성능에 가장 큰 영향을 미치는 요인은 공간의 잔향 특성이다. 마이크로폰이 존재하는 공간의 특성에 따른 시간지연 추정 성능에 대한 연구는 1997년에 Brandstein과 Silverman이 공간의 잔향 시간이 길어질수록 그 공간에서의 시간지연 추정 성능이 나빠짐을 모의실험을 통해 밝혀 놓은바 있다⁽¹⁰⁾. 하지만 동일한 잔향 시간을 갖는 공간에서 측정된 신호를 이용해 시간지연 값을 추정할 때, 신호의 특정 부분에 따라 그 추정 성능에 차이가 발생함을 확인하였다. 이는 시간지연 추정 성능에 미치는 영향이 공간의 잔향 시간 외에도 다른 요인이 존재함을 의미한다. 따라서 이 연구에서는 특정 잔향 시간을 갖는 공간에서 측정된 신호를 이용하여 시간지연을 추정할 때 추정 성능과 잔향 특성 사이의 관계를 R/D 비를 이용하여 살펴보고자 한다.

3.1 R/D 비

R/D 비(reverberant-to-direct sound energy ratio, 잔향-직접 음 에너지 비)는 직접 음(direct sound)

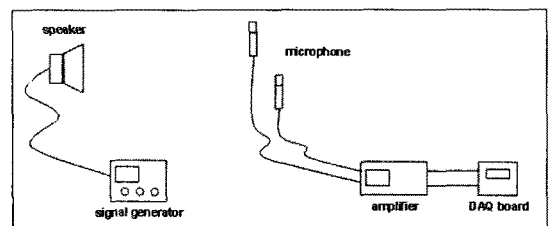


Fig. 1 Equipments to measure room impulse response

에너지에 대한 잔향 음(reverberant sound) 에너지의 비로 정의되며 잔향 강도를 나타낸다. R/D 비는 정규화된 실내 충격 응답 함수, $h(t)$ 을 이용하여 다음과 같이 나타낼 수 있다⁽¹¹⁾.

$$R/D \text{ ratio} = 10 \log \frac{\int_0^{\infty} h^2(t) dt}{\int_0^{\infty} h^2(t) dt} \text{ dB.} \quad (1)$$

여기서 분모 항은 직접 음 에너지를 의미하고, 분자 항은 잔향 음 에너지를 의미한다. 이 연구에서는 직접 측정된 신호를 직접 음과 잔향 음 부분으로 분리하여 이 둘의 에너지 비로부터 R/D 비를 구하고자 한다. 마이크로폰으로부터 측정된 신호에는 직접 음과 잔향 음이 모두 포함되어 있기 때문에, 이 둘을 직접적으로 분리할 수가 없다. 따라서 실제 공간에서 측정된 실내 충격 응답 함수를 직접 음 부분과 잔향 음 부분으로 분리하여 임의의 음성신호를 convolution함으로써 측정될 위치에서의 직접 음과 잔향 음 신호를 생성해 정의에 따라 R/D 비를 구한다.

3.2 R/D 비 계산을 위한 실내 충격 응답 함수 분리

실내 충격 응답 함수는 4m×5m×3m 크기의 측정 공간에서 Fig. 1과 같은 장비들을 이용하여 측정하였다. 신호 발생기에서 20 Hz ~ 20 kHz random white noise를 스피커를 통해 내보내 주고, 이를 두 개의 마이크로폰을 이용하여 그 신호를 측정 한 후, 입력신호와 측정된 신호 사이의 전달함수를 구하고, 측정된 전달함수의 역 푸리에 변환을 통해서 실내 충격 응답 함수를 계산하였다. 시간지연 추정을 위해 두 개의 마이크로폰을 이용하였으며 두 마이크로폰 사이의 거리는 0.3m이고, 시간지연이 발생되도록 스피커의 위치는 한 쪽 마이크로폰으로 5° 치우치게 위치시켰다. 위와 같은 실험으로 측정된 실내 충격 응답 함수는 Fig. 2와 같다. 실제 측정된 실내 충격 응답 함수는 1.5초이나, ISO 규정에 따라 계산된 잔향 시간이 약 0.28초이므로, Fig. 2에서는 0.3초까지만 나타내었다. 측정된 실내 충격 응답 함수를 직접 음 부분과 잔향 음 부분으로 나누기 위해서는 식(1)의 정의와 같이 실내 충격 응답 함수를 0⁻~0⁺와 0⁺~∞ 시간 부분으로 나눠야 한다. 하지만 이는 물리적으로 분리가 불가능 하기 때문에, Fig. 3

과 같이 직접 음 부분은 첫 번째 peak와 notch 이후에 실내 충격 응답 함수 값이 영(zero)이 될 때까지로 정하였고, 잔향 음 부분은 직접 음 부분 이후부터 나머지 실내 충격 응답 함수 부분으로 하였다.

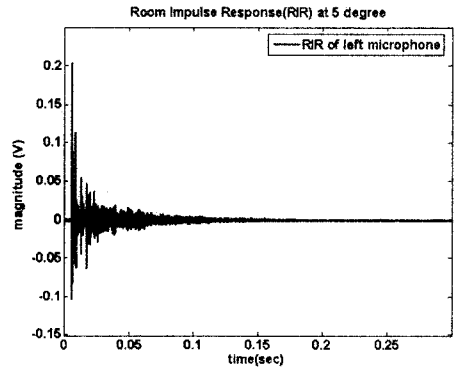
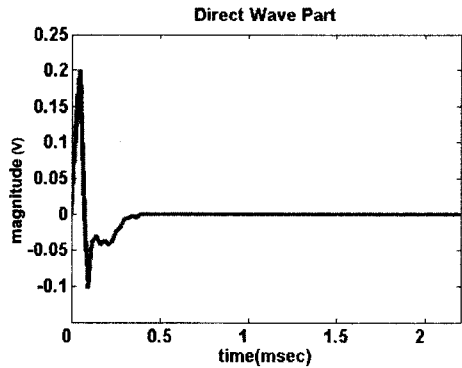
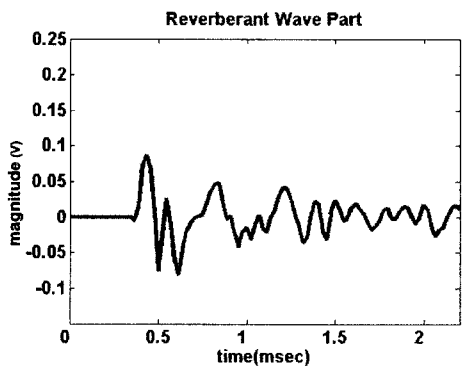


Fig. 2 Measured room impulse response(left microphone)



(a)



(b)

Fig. 3 Separation of room impulse response; (a) direct wave (b) reverberant wave

이는 음파의 첫 번째 밀함(compression)과 소함(rarefaction)부분만을 직접 음으로 고려하기 때문에 이와 같이 분리하였다. 직접 음 부분과 잔향 음 부

분으로 분리된 실내 충격 응답 함수에 임의의 음성 신호를 convolution하여 마이크로폰 위치에서 측정된 직접 음과 잔향 음에 해당하는 신호들을 만들어 내면 Fig. 4와 같다.

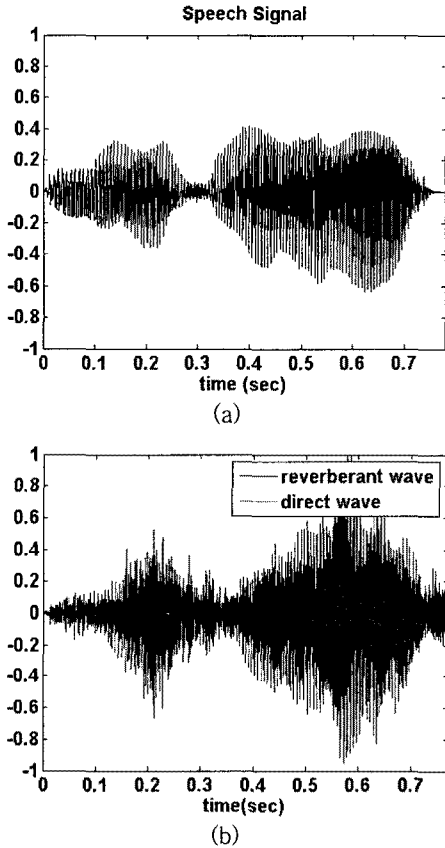


Fig. 4 Generation of direct and reverberant signal using room impulse response; (a) original speech signal, (b) generated signals

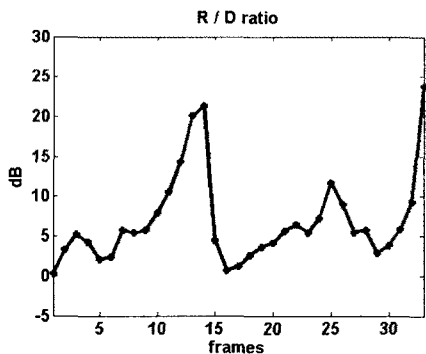


Fig. 5 R/D ratio at each frame of speech signal; 1 frame is about 23 msec(512 samples when sampling frequency is 22050 Hz)

3.3 R/D 비와 시간지연 추정 성능

앞서 설명한 방법으로 만들어진 신호들의 직접 음과 잔향 음을 이용하여 일정한 시간간격(약 23 msec, 512 samples)마다 R/D 비를 구해보면 Fig. 5와 같다. 또한 동일한 시간간격(frame)의 두 신호들을 이용하여 상호상관값을 계산한다. 이때 직접 음만을 이용할 경우와 잔향 음이 포함되어 있는 전체 신호를 이용한 경우에 대해 각각 나타내면 Fig. 6과 같다. Fig. 6에서 x축은 각 시간 별로 상호상관값을 구하기 위한 시간간격, y축은 상호상관값의 시간지연을 나타내며, 상호상관값의 크기는 밝기의 차이로 표현되어 있다. 즉, 상호상관값 peak이 주위보다 밝은 것이 시간지연을 명확하게 나타낸다고 할 수 있

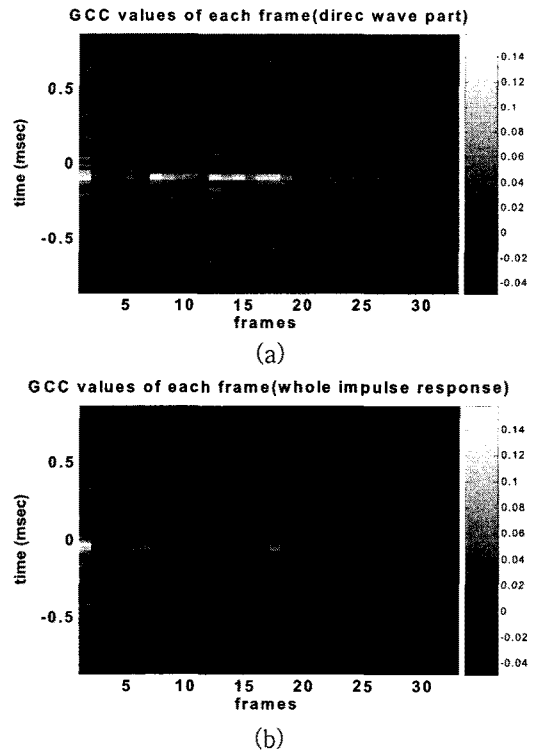


Fig. 6 Cross correlations at each frame of speech signal; (a) case of direct wave only, (b) case of whole wave(included reverberant wave)

다. Fig. 6의 (a)에서는 신호의 직접 음만을 사용했기 때문에 실제 두 신호의 시간지연을 나타내는 0.045 msec 시간에서 상호상관값 peak이 명확하게 나타나지만, (b)에서는 잔향의 영향으로 인해 전체적으로 그 명확성이 떨어지며, 10~15 frame 사이에서는 직접 음만을 이용했을 때와 그 차이가 두드러진다. 이를 Fig. 5의 R/D 비와 비교해보면, R/D 비가 큰 부분에서는 시간지연 추정을 위한 상호상관값 peak의 명확성이 떨어지며 그로 인해 추정오차가 발생하게 됨을 알 수 있다. 뿐만 아니라 잔향의 영향이 포함되어있지 않는 직접 음만을 사용한 경우에도, 2~5, 28~33 frame 사이에서는 상호상관값 peak의 명확성이 떨어지는데, 이는 다음 장에서 신호의 특성과 관련하여 살펴보도록 한다.

4. 신호의 특성에 따른 영향

앞장에서 잔향의 영향이 없는 직접 음만을 사용할 경우에도 상호상관값 peak이 명확하게 나타나지 않는 부분이 있음을 알 수 있었다. 이 장에서는 이를 음성 신호의 주파수 특성과 신호의 크기 변화의 관점에서 살펴보고자 한다. 시간지연 추정을 위해 사용된 음성 신호의 스펙트로그램 (spectrogram)을 구해보면 Fig. 7과 같다. Fig. 7에서와 같이 음성 신호의 주파수 성분이 넓게 분포하고 있는 6~20 frame 사이에서는 Fig. 6의 (a)에서 보는 것과 같이 상호상관값 peak이 명확하게 나타나지만, 2~5 frame 사이에서와 같이 그렇지 않은 부분에서는 직접 음만을 이용한 경우에도 상호상관값 peak의 명확성이 떨어짐을 알 수 있다. 이로부터 PHAT 가중

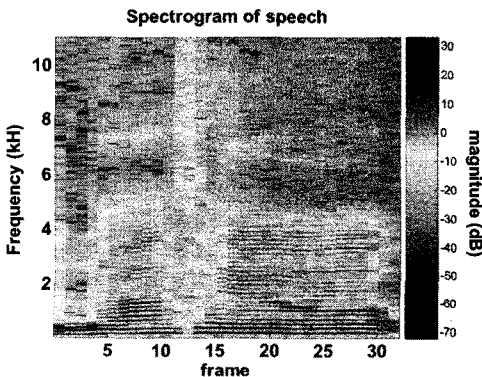


Fig. 7 Spectrogram of speech signal

치 함수를 적용한 GCC 방법을 이용하여 시간지연을 추정할 때, 넓은 범위의 주파수 성분을 갖는 음성 신호를 이용할 경우 그 성능이 좋아짐을 알 수 있다. 이를 확인하기 위해 모든 주파수 대역을 가지는 random white noise 신호를 이용하여 앞의 과정을 반복해 보았다. 이때, 신호의 크기 변화에 따른 시간지연 추정 성능도 알아보기 위해 Fig. 8과 같이 신호의 크기가 변화는 신호를 이용하였다.

앞의 음성신호를 이용한 경우와 마찬가지로 분리된 실내 충격 응답 함수를 이용해 만들어진 신호로부터 각 시간구간별 R/D 비를 구하면 Fig. 9와 같다. 그리고 동일한 간격의 두 신호를 이용하여 상호상관값을 계산하면, 직접 음만을 이용했을 경우에는 Fig. 10의 (a)와 같고, 잔향 음이 포함되어 있는 전체 신호를 이용할 경우에는 Fig. 10의 (b)와 같이 나

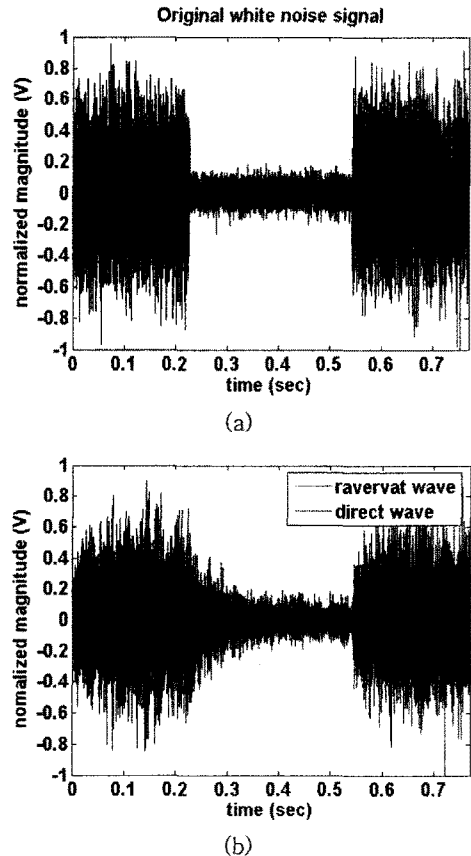


Fig. 8 Generation of direct and reverberant signal using room impulse response; (a) original random white noise signal, (b) generated signals

타난다. 여기에서 알 수 있듯이 전 주파수 대역을 포함하는 random white noise신호의 경우에는 신호

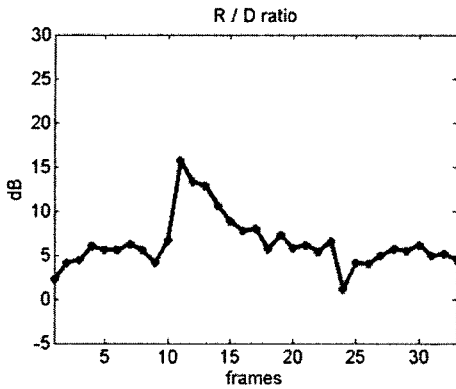


Fig. 9 R/D ratio at each frame of random white noise signal; 1 frame is about 23 msec(512 samples when sampling frequency is 22050 Hz)

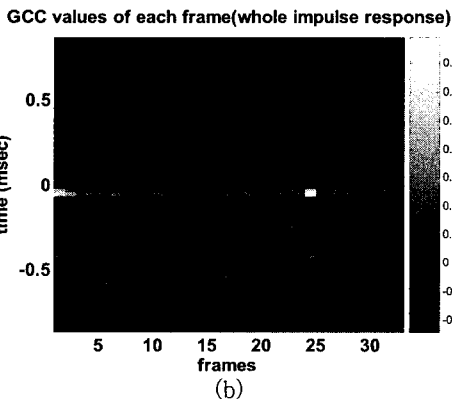
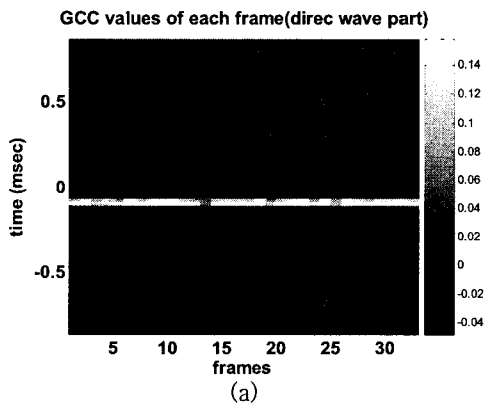


Fig. 10 Cross correlation at each frame of random white noise; (a) case of direct wave only, (b) case of whole wave(included reverberant wave)

의 크기에 상관없이 상호상관값 peak의 명확성이 뚜렷하게 나타남을 알 수 있다. 하지만 잔향의 영향이 포함되어 있는 경우에는 신호의 크기가 크다가 작아지는 부분인 10~13 frame에서 명확성이 떨어지며, 신호의 크기가 작다가 커지게 되는 24 frame에서는 명확성이 뚜렷하게 나타나게 된다. 이는 신호의 크기 변화에 따른 잔향의 영향이 달라져 R/D 비가 변화하기 때문이다. 즉, 신호의 절대적인 크기는 시간 지연 추정 성능에 영향을 미치지 않지만, 시간에 따른 크기 변화는 결과적으로 잔향 특성에 영향을 미치므로 시간지연 추정 성능에 영향을 미치고 있음을 알 수 있다.

5. 성능지표를 고려한 시간지연 추정

4장에서는 GCC 방법을 이용하여 시간지연을 추정할 때, 그 추정성능은 시간지연을 나타내는 상호상관값 peak의 명확성으로 나타냈으며, 이에 영향을 미치는 요인으로는 R/D 비가 주요하다는 것을 알아 보았다. 이 장에서는 정성적으로 나타낸 추정 성능을 정량적으로 표현하면서, 동시에 R/D 비와 유사한 경향을 나타내는 성능지표(performance index, PI)를 정의하고, 이를 이용해 시간지연을 추정함으로써 음원 위치 추정의 정확성을 증대시키는 방법에 대해 생각해 본다.

5.1 성능지표

시간지연 추정에서 상호상관값 peak의 명확성을 정량적으로 나타내기 위해 PI를 식 (2)와 같이 정의한다.

$$PI(\%) = \frac{\sum_{\tau}^{\tau_{Max} + \epsilon} |GCC|}{\sum_{\tau}^{\tau_{Max} - \epsilon} |GCC|} \times 100 \quad (2)$$

τ 는 마이크로폰 사이의 거리에 의해 물리적으로 발생할 수 있는 최대 시간지연이고, τ_{Max} 는 상호상관값 peak에서의 시간지연을 의미이다. 또 ϵ 은 시간지연 값의 신뢰구간을 정의하는 변수로 본 논문에서는 τ 의 5%로 정의하였다. 이는 음원의 주파수 대역폭에 따라 달라질 수 있는 값이며⁽¹²⁾, 일반적인 음성신호(4 kHz 미만)를 기준으로 이 논문에서는 정

의하였다. PI는 물리적으로 의미 있는 시간지연 구간의 상호상관값들의 전체 절대값 중에서 신뢰 구간의 절대값들이 차지하는 비율이며, 이는 상호상관값 peak이 존재하는 부분이 다른 시간지연에서의 값들에 비해서 얼마나 주요한지를 나타낸다. 또한 PI는 측정되는 신호의 환경에 따라 그 기준이 변할 수 있는 상대적인 값이다.

일반적으로 R/D 비가 큰 부분의 신호에서 계산된 상호상관값 peak이 실제 두 신호 사이의 시간지연이 아닌 다른 시간에 존재할 경우 즉, 시간지연 추정 오차가 발생한 경우 PI는 상대적으로 작은 값을 가지게 된다. 이와 같은 특성은 Fig. 6의 각 frame별 상호상관값들에 대해 PI를 계산한 Fig. 11로부터 알 수 있다. 직접 음만을 이용한 경우가 잔향 음이 포함되어 있는 경우보다 전반적으로 PI가 높게 나타나며, R/D 비가 상대적으로 큰 10~15 frame사이에서는 PI가 더 낮게 나타남을 알 수 있다. 이로부터 측

정된 신호를 이용해 직접 계산할 수 없는 R/D 비 대신에, 직접 계산이 가능한 PI값을 이용하여 상호상관값 peak의 명확성이 떨어져 시간지연 추정 시 오차를 유발할 수 있는 부분을 선별함으로써 음원 위치 추정의 정확성을 증대시킬 수 있다.

5.2 성능지표를 이용한 음원 위치 추정 실험

앞서 설명한 PI값을 이용하여 음원 위치 추정의 정확성을 높일 수 있음을 검증하기 위해 두 개의 마이크로폰을 이용한 음원 위치 추정 실험을 Table 1과 같은 조건에서 수행하였다. 측정된 음성 신호에

Table 1 Experimental setup

Place	General office room (background noise : 35 dbA, reverberation time : about 1sec)
Sound source	Male voice (5 words at each position)
Microphone	B&K Free-field mic.-Type 4190
Distance between microphones	0.3 m
Distance between mic. & speaker	1 m
Sampling freq.	22050 Hz
Source position	Each 30° from -60° to 60°

Table 2 Experimental results

1. Without PI

Source posi.	Correct estimation	Wrong estimation
-60°	50 %	50 %
-30°	56 %	44 %
0°	61 %	39 %
30°	53 %	47 %
60°	47 %	53 %

2. With PI

Source posi.	HPCE	HPWE	LPCE	LPWE
-60°	31 %	5 %	20 %	44 %
-30°	36 %	6 %	20 %	38 %
0°	42 %	9 %	19 %	30 %
30°	32 %	5 %	21 %	42 %
60°	34 %	7 %	13 %	46 %

* HPCE : Higher than standard PI and correct estimation
 HPWE : Higher than standard PI and wrong estimation
 LPCE : Lower than standard PI and correct estimation
 LPWE : Lower than standard PI and wrong estimation

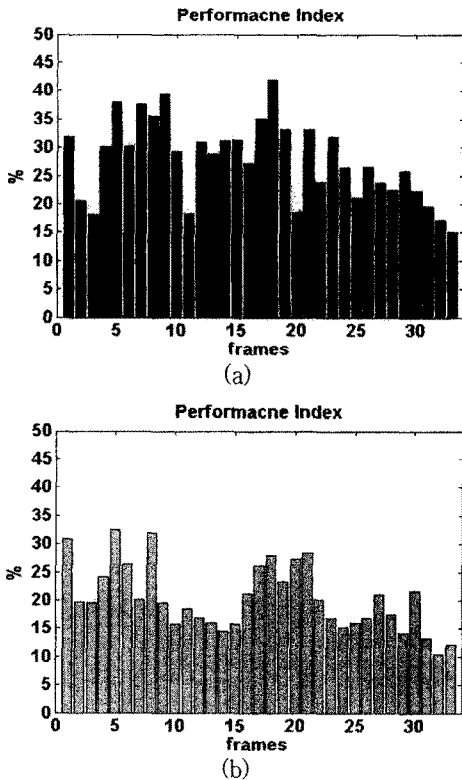


Fig. 11 Performance index at each frame of speech signal; (a) case of direct wave only, (b) case of whole wave(included reverberant wave)

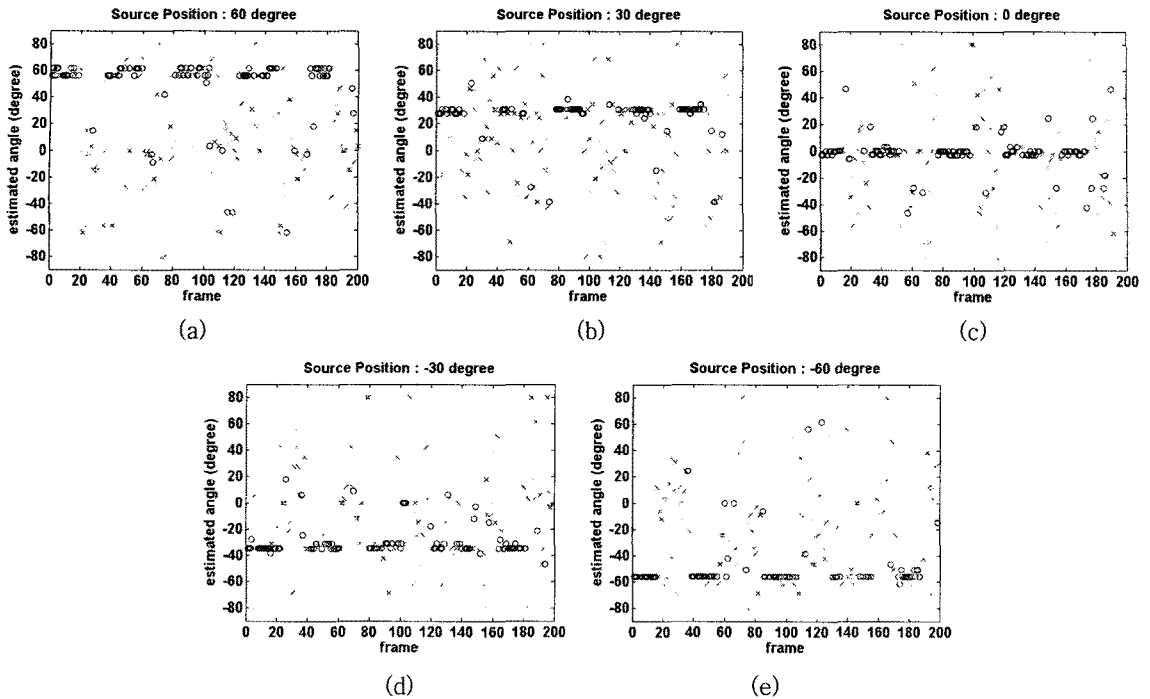


Fig. 12 Experimental results for sound source localization; (a)~(e) estimation results as source position(60°~ -60°)
 o : PI is higher than 20%, x : PI is lower than 20%

서 각 frame (23 msec)별 시간지연을 계산해 음원의 위치를 추정된 결과는 Fig. 12와 같다. 이는 각 frame별 추정된 음원의 위치를 나타내며, 그 중에서 PI가 20(%)이 넘는 경우에는 'o'로 나타내었고 20보다 낮은 경우에는 'x'로 나타내었다. 여기에서 PI의 신뢰기준은 실험 공간에서 사전 실험을 통해 최적의 성능을 낼 수 있는 값으로 20을 선정하였다. 그림 12의 결과와 같이 미리 정해놓은 PI값을 기준으로 하여 기준 PI보다 낮은 값을 가지는 부분들('x'로 표현된 부분)을 제외시킬 수 있어 음원 위치 추정에서 정확성을 높일 수 있다. 이를 정량적으로 나타내면, Table 2와 같다. 실제 음원의 위치에 대해서 $\pm 10^\circ$ 를 허용 오차범위로 정했을 경우, PI를 이용하지 않은 경우에는 약 50%정도 추정 오차가 발생하게 된다. 하지만 실험공간에서 선정한 기준 PI를 이용하여 음원의 위치를 추정하면, 기준 PI보다 높은 값을 가지면서 정확하게 음원의 위치를 추정한 경우가 약 35%, 잘못된 음원의 위치를 추정한 경우가 약 6%였다. 그리고 기준 PI보다 낮은 값을 가지는 부분(즉, 음원 위치 추정에서 신뢰하지 않는 부분) 중에서 음원의 위치를 정확하게 추정한 경우가 약 20%,

오차가 발생된 경우가 약 40%였다. 20%정도 정확히 추정된 결과를 제외하지만, 약 40%정도의 잘못된 추정된 결과를 제외할 수 있으므로, 이 논문에서 제안한 PI를 이용하면 음원의 위치를 추정 시 그 정확성 및 효율성을 향상시킬 수 있다.

6. 결 론

이 연구에서는 도달시간지연을 이용한 음원 위치 추정 방법에서 핵심이 되는 시간지연 추정 성능에 미치는 영향들에 대해서 알아보았다. PHAT 가중치 함수를 이용한 시간지연 추정에서 그 성능은 주로 공간의 잔향특성에 의해 영향을 받는데, 그 중에서 직접 음 에너지에 대한 잔향 음 에너지의 비로 정의되는 R/D 비를 이용하여 분석해본 결과, R/D 비가 큰 부분에서는 시간지연 추정 성능이 나빠짐을 알 수 있었다. 그리고 신호의 특성측면에서는 측정된 음성 신호가 넓은 주파수 대역을 가질수록 좋은 성능을 보이며, 신호가 작다가 커지는 부분에서 상대적으로 R/D 비가 감소하여 시간지연 추정 성능이 좋아짐을 알 수 있었다. 또한 이 논문에서는 시간지

연 추정성능을 정량적으로 나타내면서, R/D 비와 유사한 경향을 나타내는 PI를 정의하였다. 이를 이용하여 음원의 위치를 추정할 경우 약 40%의 추정 오차 발생부분을 제외시킬 수 있어, 음원 위치 추정에서 정확성과 효율성을 높일 수 있음을 실험을 통해 검증하였다.

후 기

이 연구는 한국과학재단을 통해 교육과학기술부의 국가지정연구실 사업(ROA-2005-000-10112-0), 두뇌 한국 21프로젝트 일환으로 수행하였음.

참 고 문 헌

- (1) Hwang, S. M., Park, Y.-S. and Park, Y. G., 2005, "Detection of Speaker Position for Robot Using HRTF", Proceedings of the KSNVE Annual Autumn Conference, pp. 637~640.
- (2) Sasaki, Y., Kagami, S. and Mizoguchi, H., 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.
- (3) Brandstein, M. S. and Silverman, H. F., 1997, "A Practical Methodology for Speech Source Localization with Microphone Arrays", Computer Speech and Language, Vol. 11, No. 2, pp. 91~126.
- (4) Kwon, B. H., Park, Y. J. and Park, Y.-S., 2007, "Sound Source Localization Method Applied to Robot System", Proceedings of the KSNVE Annual Autumn Conference, KSNVE07A03-02.
- (5) Wang, H. and Chu, P., 1997, "Voice Source Localization For Automatic Camera Pointing System In Videoconferencing", Acoustics, Speech, and Signal Processing, ICASSP-97, Vol. 1, pp. 187~190.
- (6) Knapp, C. H. and Carter, G. C., 1976, "The Generalized Correlation Method for Estimation of Time Delay", IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. AS-24, No. 4, pp. 320~327.
- (7) Vaccaro, R. J., Ramalingam, C. S. and Tufts, D. W., 1992, "Least-squares Time-delay Estimation for Transient Signals in a Multipath Environment", Journal of Acoustic Society of America, Vol. 92, No. 1, pp. 210~218.
- (8) Cabot, R. C., 1981, "A Note on the Application of the Hilbert Transform to Time Delay Estimation", IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-29, No. 3, pp. 607~609.
- (9) Chan, Y. T., So, H. C. and Ching, P. C., 1999, "Approximation Maximum Likelihood Delay Estimation via Orthogonal Wavelet Transform", IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-47, No. 4, pp. 1193~1198.
- (10) Brandstein, M. S. and Silverman, H. F., 1997, "A Robust Method for Speech Signal Time Delay Estimation in Reverberant Rooms", Acoustics, Speech and Signal Processing, Vol. 1, pp. 375~378.
- (11) Ahn, S. T., 1999, "A Study on FIR Filter Design for Natural Sounding Artificial Reverberation", Master's Thesis, KAIST.
- (12) Yoon, J.-R., Ro, Y.-J. and Kim, D.-H., 1998, "Effect of Noise Source Bandwidth on Noise Source Imaging Technique," Proceedings of the KSNVE Annual Autumn Conference, pp. 122~126.