

향상된 실시간 음원방향 인지 시스템의 하드웨어 설계

Hardware Design of Enhanced Real-Time Sound Direction Estimation System

김 태 완*, 김 동 훈**, 정 연 모*
(Taewan Kim*, Donghoon Kim**, Yunmo Chung*)

*경희대학교 전자·전파공학과, **LIG넥스원(주) 연구개발본부 Maritime 연구소
(접수일자: 2010년 12월 15일; 수정일자: 2011년 3월 15일; 채택일자: 2011년 3월 31일)

본 논문에서는 십자 형태로 구성된 네 개의 마이크로폰을 이용하여 일반화된 상호 상관 기법을 적용한 음성 도달시간 지연을 측정하여 정확한 음원 방향을 실시간으로 계산하는 방식에 대해 제시하였다. 기존 시스템에서는 마이크로폰 어레이 신호처리를 위해 데이터 수집 장치를 필요로 하므로 시스템을 내장하기 힘들고, 또한 DSP 프로세서를 사용한 음원방향 인지는 마이크로폰의 채널의 수가 늘어날수록 실시간 처리가 어려워지는 두 가지 단점이 있다. 본 논문에서는 이러한 한계를 극복하기 위하여 마이크로폰 어레이 신호처리를 이용한 향상된 음원방향 인지 하드웨어의 개발을 제안 하였다. 공간 구분 기법을 이용한 효율적인 설계 및 검증방식을 제안하였고 이를 통하여 보다 정확한 방향 추정과 설계시간 단축이 가능하다. 최종적으로 음성 코덱과 FPGA를 이용하는 임베디드 시스템을 위해서 사용이 가능한 시스템을 개발하였다. 실험 결과에 의하면 PC 기반이나 DSP 프로세서를 사용한 경우에 비해 보다 빠른 처리 시간을 보였다.

핵심용어: 음원방향 인지, 일반화된 상호 상관 기법, 도달 시간 지연, 마이크로폰 어레이 신호처리, FPGA

투고분야: 음향 신호처리 분야 (1,2)

In this paper, we present a method to estimate an accurate real-time sound source direction based on time delay of arrival by using generalized cross correlation with four cross-type microphones. In general, existing systems have two disadvantages such as system embedding limitation due to the necessity of data acquisition for signal processing from microphone input, and real-time processing difficulty because of the increased number of channels for sound direction estimation using DSP processors. To cope with these disadvantages, the system considered in this paper proposes hardware design for enhanced real-time processing using microphone array signal processing. An accurate direction estimation and its design time reduction is achieved by means of an efficient hardware design using spatial segmentation methods and verification techniques. Finally we develop a system which can be used for embedded systems using a sound codec and an FPGA chip. According to experimental results, the system gives much faster real-time processing time compared with either PC-based systems or the case with DSP processors.

Keywords: Sound direction estimation, Generalized Cross Correlation, Time Delay of Arrival, Microphone array signal processing, Field Programmable Gate Array

ASK subject classification: Acoustic Signal Processing (1,2)

I. 서론

소리의 방향 지각에 대한 연구는 전통적인 음향학적 문제 (acoustic problem)들 중에 하나였으며, 19세기 후반 Rayleigh 와 Alfred M. Mayer 등의 음향학자들로부터 실험 및 수학적 이론화를 통해 설득력 있는 결과를 얻게

되었다. 방향 지각에 대한 문제는 현재 음원 국지화 이론의 밑바탕이 되어왔다 [1-2]. 음원 국지화를 통한 음원이나 표적의 위치를 추정하는 방식은 능동형 시스템 (active system)과 수동형 시스템 (passive system)으로 나눌 수 있다. 능동형 시스템의 경우 표적을 향해 펄스를 전송하여 펄스가 돌아온 시간을 측정함으로써 표적의 위치를 추정하는 방식이며 일반적인 예로 레이더 (radar) 시스템의 경우를 들 수 있다. 수동형의 경우 음원 표적의 소리 정보를 이용하여 수신된 신호간의 시간 지연을 측정하여

위치를 추정하는 시스템이다. 일반적인 예로 SONAR 시스템을 들 수 있다 [3].

수동형 시스템의 음원 국지화 형태를 기반으로 연구되어 온 음원 지각 방식은 현재 인간과 로봇간의 상호작용 (Human robot interaction: HRI)을 위한 청각정보의 활용에 이용된다. 음원의 위치를 추정하여 카메라의 도움을 받아 자연스러운 로봇의 움직임에 이용된다. 뿐만 아니라 보안이나 감시를 위한 스마트 감시 카메라 및 군사용 비행기나 탱크가 유발한 소음의 위치를 파악하는데 적용에 되고 있다. 연구되고 있는 다수개의 마이크로폰을 이용한 빔포밍 (beamforming) 방법 등이 제안되고 있다. 빔포밍 방법의 경우 너무 많은 마이크로폰이 요구되고 이에 따른 연산의 복잡성을 가지고 있다. 본 논문에서는 십자형태의 4개의 마이크로폰을 이용하여 GCC기법을 적용한 음성 도달시간지연 (TDOA)을 측정하고 음원 위치 추정 방식에 대해 연구하였으며, 마이크로폰 어레이 신호처리를 이용하여 음원 정보의 위치를 획득하는 하드웨어의 개발을 위해 보다 효율적인 설계 방식을 제안하였다. 이러한 설계 방식을 통해 기존 연구들의 마이크로폰어레이를 사용하기 위해 DAQ를 필요로 하였고 그로 인해 시스템을 내장하기 힘든 단점을 극복하기 위해 최종적으로 음성 코덱과 FPGA를 이용한 내장 가능한 음원방향 인지 시스템을 제안 하였다.

II. 음원 방향 추정 모델

기존에 마이크로폰을 이용하여 음원의 방향을 추정하는 방법에 대해 전반과 한계를 언급한다.

2.1. 마이크로폰을 이용한 음원 방향 추정 방식

음원 방향 추정을 위한 가장 기본적인 형태는 그림 1과 같이 두 개의 마이크로폰을 이용한 원거리 모델을 적용하여 각도를 추정할 수 있다.

그림 1에서 두 개의 마이크로폰 M_1 , M_2 간의 거리 차이는 $l \cos \theta$ 이고 거리 차이에 의해 발생하는 음원 신호 입력의 시간 지연은 식 (1)을 이용해 구할 수 있고 해당식으로 각도를 추정할 수 있다.

$$v = l \cos \theta / \Delta t \tag{1}$$

$$\theta = \cos^{-1}(\Delta t \cdot v / l) \tag{2}$$

마이크로폰으로부터의 음원간의 거리차이를 알면 마

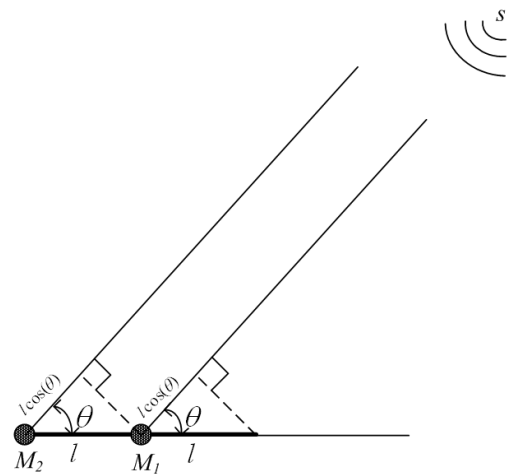


그림 1. 두 개의 마이크로폰 어레이를 이용한 기하학적 분석
Fig. 1. Geometrical analysis using two microphone array.

이크로폰간의 거리와 음원의 속도 (v) 만을 이용하여 음원이 위치한 각도를 인지 할 수 있게 된다. 각도의 정확도는 식 (2) 의 Δt 값을 얼마나 정확하게 측정해 낼 수 있는가에 의해 결정되게 된다. Δt 값을 측정하는 방식들은 ZCPA (zero-crossing and their respective peak amplitudes) [4], CC (Cross-Correlation) 기법, GCC (Generalized Cross-correlation) 기법 등이 대표적인 방식들이다 [5-6]. 일반적으로 음원방향 인지 시스템에서 가장 효과적으로 활용될 수 있는 시간지연 측정 기법은 GCC 기법중 PHAT (Phase Transform) 가중치 함수를 이용한 방식으로 노이즈나 잔향에 강인함을 갖고 있으며 시간 영역상의 CC 기법에 비해 하드웨어 복잡성을 감소시킬 수 있다. 각각의 가중치 함수에 대한 TDOA의 퍼포먼스에 대한 결과는 [7] 에서 확인 할 수 있다. GCC 기법을 이해하기 위해 마이크로폰에 입력되는 신호의 시간 축에서의 수학적 모델은 다음 식 (3), 식 (4)과 같다.

$$x_1(t) = s(t) + n_1(t) \tag{3}$$

$$x_2(t) = \alpha s(t+D) + n_2(t) \tag{4}$$

여기서 x_1 은 마이크로폰1에 입력된 음원의 신호 s 와 유입된 노이즈 n_1 를 더한 값이다. x_2 는 마이크로폰2에 입력된 음원 신호 s 의 감쇄 (α)된 지연된 신호 (D)과 유입된 노이즈 (n_2)를 더한 값이다. 이 두 신호간의 CC는 다음 식 (5)과 같이 나타낼 수 있다.

$$R_{x_1x_2}(\tau) = \frac{1}{T-\tau} \int_{\tau}^T x_1(t)x_2(t-\tau)dt \tag{5}$$

식 (5)에 대한 디지털 형태의 표현은 식 (6)와 같다.

$$R_{ij}(\tau) = \sum_{n=0}^{N-1} x_i[n]x_j[n-\tau] \quad (6)$$

여기서, 식 (5)의 T는 연속시간 환경에서의 시간을 의미하고, 식 (6)에서의 N은 정수로서, 이산시간 신호임을 의미한다.

식 (6)와 같이 표현된 CC 방식은 n^2 크기의 하드웨어 복잡성 (complexity)을 갖는다. 이러한 복잡성을 감소시킬 수 있는 방식으로는 GCC기법을 이용할 수 있다. GCC 기법의 경우 FFT를 이용한 변환과 역변환을 고려했을 경우 하드웨어 복잡성은 약 $3n \log_2 n$ 크기로 감소된다.

$$R_{x_1x_2}(\tau) = \int_{-\infty}^{\infty} X_1(f)X_2(f)^* e^{j2\pi f\tau} df \quad (7)$$

$X_1(f)$ 와 $X_2(f)$ 는 x_1 과 x_2 의 신호가 주파수 영역으로 변환된 식이다. 주파수 영역으로 변환된 두 식의 곱을 시간 영역으로 역변환된 결과가 CC기법과 동일한 값을 갖는다. 이에 대한 디지털 형태의 표현은 식 (8)과 같다.

$$R_{x_1x_2}(\tau) = \sum_{k=0}^{N-1} X_1(k)X_2(k)^* e^{j2\pi k\tau/N} \quad (8)$$

GCC기법의 경우에 식 (8)에 가중치 함수가 포함된 형태로 변형하면 식 (9)와 같다.

$$R_{x_1x_2}(\tau) = \sum_{k=0}^{N-1} \Psi_g(k) \hat{G}_{x_1x_2}(k) e^{j2\pi k\tau/N}, \quad (9)$$

$$(\hat{G}_{x_1x_2}(k) = X_1(k)X_2(k)^*)$$

몇 가지 GCC기법중에서 음원 국지화 시스템에서 높은 신뢰성을 보여주는 가중치 함수는 PHAT 방식이 주로 이용된다. PHAT에 해당하는 가중치 함수는 식 (10)과 같다.

$$\Psi_g(k) = 1/|G_{x_1x_2}(k)| \quad (10)$$

GCC기법의 가중치 함수는 CC, Roth impulse Responce, SCOT, PHAT, Eckart, ML or HT 가 있으며 각 가중치 함수의 특징은 [6]를 참고 할 수 있다. GCC연산을 통해 계산된 값 중 가장 큰 값에 해당하는 샘플이나 시간이 두 신호간의 지연시간을 의미한다. 수식으로 나타내면 식 (11)과 같다.

$$\Delta T_{12} = \operatorname{argmax}_{\tau} R_{x_1x_2}(\tau) \quad (11)$$

CC 기법의 주요 특성은 주파수가 높아질수록 식 (11)에 적용되어야할 최대값을 파악하기 어렵다. 이런 특성을 피하기 위해 LPF (Low Pass Filter)를 적용하여 예방 할 수 있다.

2.2. 기존 음원방향 인지 시스템과 한계

기존 음원방향 인지 시스템의 대부분의 형태는 DAQ를 이용하여 마이크로폰에 입력된 신호를 PC에 전달하여 신호 분석을 통해 음원의 방향 정보를 획득하는 방식들이 대부분 이용되고 있다 [8-10]. 이러한 시스템은 입력 신호를 수신하기 위한 마이크로폰이 많아질 경우에 DAQ가 요구하는 채널의 수가 증가하며, 음향 분석을 위한 충분한 샘플링 속도와 비트수를 만족 시키는 DAQ는 상당히 고가의 장비이기 때문에 로봇이나 감시용으로 사용하기에 적합하지 않다. DAQ 자체의 부피도 상당하기 때문에 원하는 시스템에 내장하기 어렵고 DAQ의 사용은 높은 전력을 소비하게 되어 충분한 전력 공급이 어려운 모바일 로봇과 같은 시스템에는 적용이 불가능하다.

또한 실시간으로 위치 정보를 탐지하기 어려운 점도 현재 시스템의 한계 중의 하나이다. PC의 프로세서를 이용한 방향 측정 방식이나 DSP 프로세서를 이용한 방식의 경우 연산량이 많아지게 되면 연산결과를 획득하는데 많은 지연이 발생하게 된다. 이러한 지연이 심해지게 되면 음원이 빠르게 움직이거나 짧은 음원이 발생한 경우 방향 탐지의 정확성이 떨어지거나 음원의 위치 정보를 놓치게 된다. DAQ와 실시간 응답성과 관련된 문제뿐만 아니라 음원방향 인지 및 음향 관련 이론의 복잡성과 아날로그 신호의 특성상 기존 FPGA나 HDL 관련 툴에서는 하드웨어 구현 및 검증이 쉽지 않다. 따라서 하드웨어 설계 이전에 보다 정확한 하드웨어로의 구현을 위하여 음원방향 인지 알고리즘의 동작을 미리 알아볼 수 있는 환경이 요구된다.

III. 음원방향 인지 시스템의 하드웨어 설계

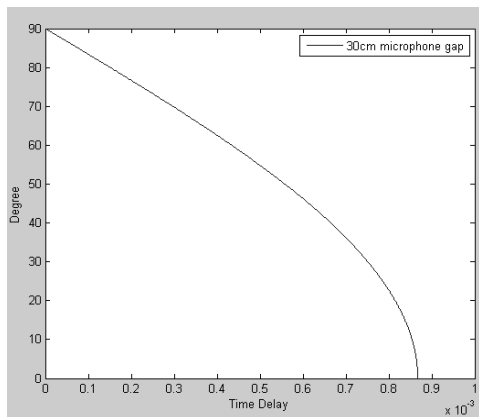
3.1. 음원 방향 추정을 위한 마이크로폰 특징 분석

본 연구를 위해 제작한 마이크로폰 어레이의 구조는 그림 2과 같다.

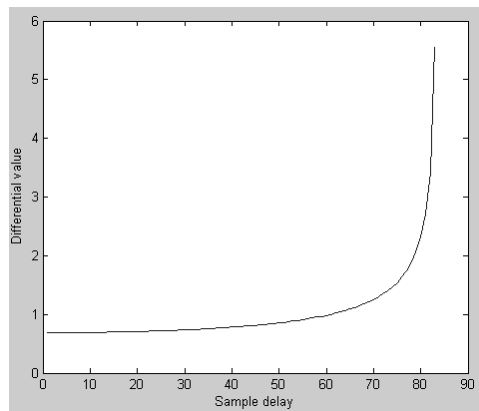
제안한 형태의 마이크로폰 어레이는 전방 (M_F), 후방 (M_B) 한 쌍과 우측 (M_R), 좌측 (M_L) 한 쌍으로 총 네 개의



그림 2. 실험을 위해 제작한 십자 형태의 마이크로폰 어레이 구조
Fig. 2. Cross-type microphone array for experiment.



(a)



(b)

그림 3. 30 cm 간격의 마이크로폰의 시간 지연과 각도의 관계
(a) 시간 지연과 각도의 관계 (b) 샘플들간의 변화율
Fig. 3. Relationship of time delay and degree with 30 cm-microphone gap. (a) Relationship of time delay and degree (b) Differential sample values.

마이크로폰으로 이루어져 있다. 각 마이크로폰 쌍의 간격은 각각 30 cm이며 실내 온도 20도인 경우를 가정하여 하나의 마이크로폰 쌍이 획득하는 음원으로부터 시간 지연 값을 측정하였다.

마이크로폰과 음원이 일직선상에 있는 경우에 이론적인 최대 지연 시간은 0.86458 ms이다. 위와 같은 조건을 디지털 하드웨어로 설계하기 위한 샘플링 비율은 96 kHz를 사용하였으며 최대 지연시간의 경우에 84개의 샘플로

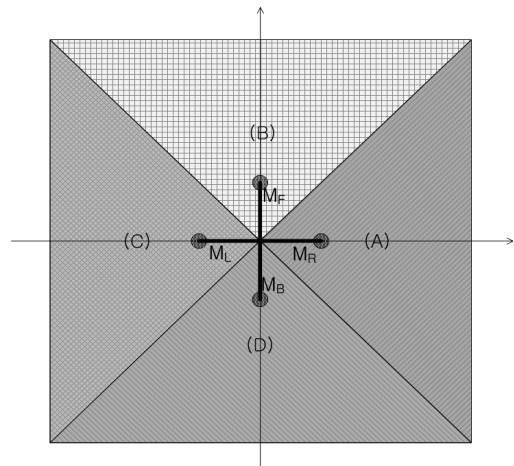


그림 4. 각도 인지를 위한 영역 구분
Fig. 4. Segmentation for degree recognition.

각각의 시간 지연을 나타낼 수 있다. 샘플링 비율이 높아 질수록 더 많은 샘플 개수를 이용하여 각도를 나타낼 수 있으나 96 kHz 이상의 샘플링 가능한 고속 사운드 코덱의 경우 고가이고 전력을 소모가 많은 단점을 가지고 있으며 96 kHz 샘플링 비율의 경우 대부분의 사운드 코덱이 지원하며 가격이 저렴한 장점이 있다. 96 kHz의 샘플링 비율로 각각의 샘플 즉 지연 값에 대한 각도를 알아보면 그림 3과 같다.

그림 3의 (a)는 시간 지연에 대한 각도 값의 관계를 나타내는 그래프이며 (b)는 샘플간의 차이 즉 (a)의 기울기 값을 그래프로 나타낸 것이다. 그림 3의 그래프를 살펴보면 음원이 두 개의 마이크로폰과 직각인 방향 (90도)에 존재하여 시간 지연이 없는 경우와 두 개 마이크로폰 쌍과 음원이 일직선상 (0도)에 놓이는 경우의 특징들을 확인할 수 있다. 90도에서 약 30도 까지 경우 선형적으로 감소하는 것을 확인할 수 있으나 30도에서 0도 사이에 음원이 위치하는 경우 딜레이 샘플값의 한 두 개의 오차에도 급격한 각도 오차를 갖게 된다. 마이크로폰 쌍과 일직선상에 음원이 위치할 경우 정확한 음원 방향을 획득하는데 한계가 있다. 전방향 음원 추정을 위한보다 정확한 결과를 위해 그림 4와 같이 네 개의 구역 (A, B, C, D)으로 나누어 각도를 인지하는 방식을 이용한다. 영역 구분을 통해 A영역과 C영역은 마이크로폰 M_R와 M_B쌍을 이용하여 각도를 구하고 B영역과 D영역은 M_R과 M_L쌍을 이용하여 각도를 구한다.

하나의 마이크로폰 쌍으로 추정 가능한 각도는 0도에서 180도 사이임으로 전방향 추정이 불가능기 때문에 두 개의 마이크로폰 쌍을 이용하여 전방향 추정이 가능하도록 하였다.

표 1. 음원 존재 가능 구역에 대한 맵핑 테이블
Table 1. Mapping table for sound source region.

	A (0-90)	B (91-180)	C (181-270)	D (271-359)
M_{FB}	0-90	0-90	91-180	91-180
M_{RL}	0-90	91-180	91-180	0-90
Result Output	M_{RL}	$M_{FB}+90$	$M_{FB}+90$	$450-M_{FB}$
	$90-M_{FB}$	M_{RL}	$360-M_{RL}$	$360-M_{RL}$

표 1은 각각의 구역에 해당하는 마이크로폰 쌍이 추정할 수 있는 각도의 범위와 해당 구역에 대한 각도의 결과값에 대해 테이블로 나타낸 것이다.

M_{FB} 는 전방과 후방 마이크로폰 쌍, M_{RL} 은 우측과 좌측의 마이크로폰 쌍을 의미 한다. 그림 4과 같이 구분된 영역에 음원이 존재할 경우 표 1의 조건에 맵핑함으로써 최종 음원 방향을 구해낸다.

3.2. 하드웨어 설계 및 검증

음원방향 인지 시스템의 하드웨어 설계를 위해서는 음원을 입력 받을 수 있는 사운드 코덱이 필요하며 본 연구에서는 최대 샘플링 비율인 96 kHz로 음성 신호를 입력 받는다. 하드웨어의 전반적인 구조는 그림 5과 같다.

본 시스템은 두 개의 WM8731 스테레오 사운드 코덱을 적용하여 음원을 입력 받고, Xilinx Vertex4 FPGA를 이용하여 GCC 연산의 신호처리를 가속화 하였고, ARM s3c2440을 이용하여 FPGA에서 연산된 결과값들을 처리하고 통계적 수치관련 결과값을 획득할 수 있는 시스템을 설계하였다. 시스템 전반의 타이밍 다이어그램은 그림 6, 7과 같다.

FPGA와 사운드 코덱은 I²S 사운드 코덱 인터페이스에 맞추어 통신한다. LRCK 신호에 따라 16-비트의 음성 데이터를 시리얼로 전송한다. start 신호가 high가 되면 Left채널의 데이터와 Right채널의 데이터가 FFT 연산을 시행하게 된다. FFT 연산과정은 데이터 로딩과 프로세싱 그리고 데이터 출력의 과정을 거치며 식 (12)에 해당하는 만큼의 사이클 후에 FFT 연산 결과가 출력된다.

$$T_{cycle} = 3N/2 - 2 + (N/2 + 3) \log_2(N) \quad (12)$$

N은 FFT의 사이즈를 의미하며 본 연구에서는 512 사이즈의 FFT로 설계 하였다. FFT 연산이 수행된 두 신호는 한 쪽의 신호를 쉼터 복소화한 후 곱셈연산을 거쳐 IFFT 연산을 수행하고 해당 연산 값이 CC 결과값에 해당된다.

기존 HDL개발툴에서 음향 신호처리에 대한 시뮬레이션과 검증이 어렵기 때문에 본 시스템의 설계는 Matlab

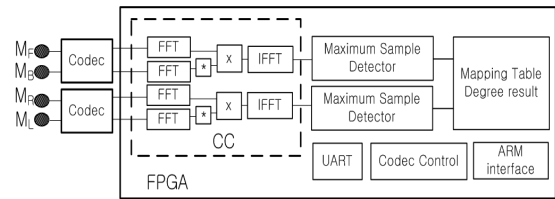


그림 5. 전반적인 하드웨어 구조
Fig. 5. Overall hardware configuration.

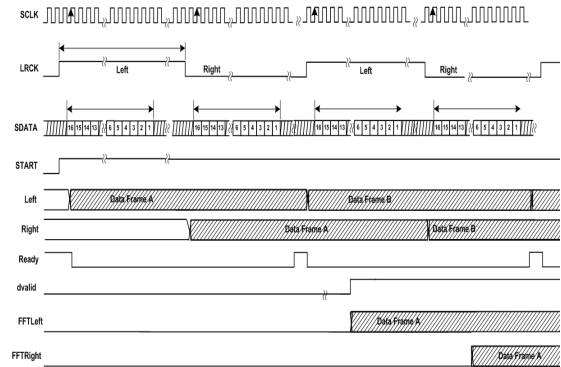


그림 6. 시스템 전반의 타이밍 다이어그램
Fig. 6. Timing diagram of overall system.

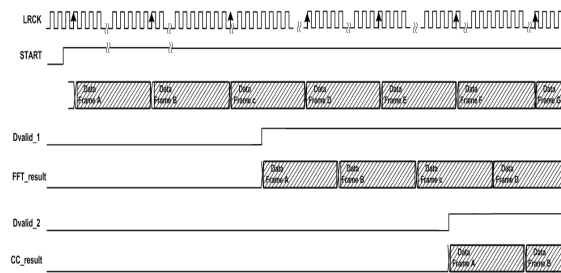
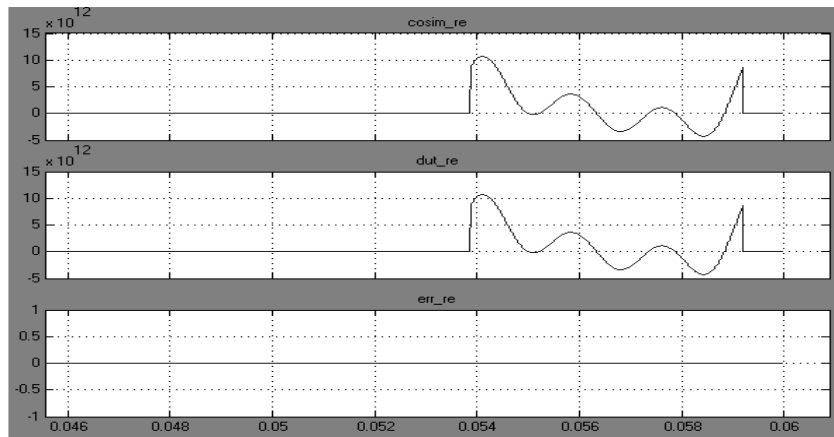
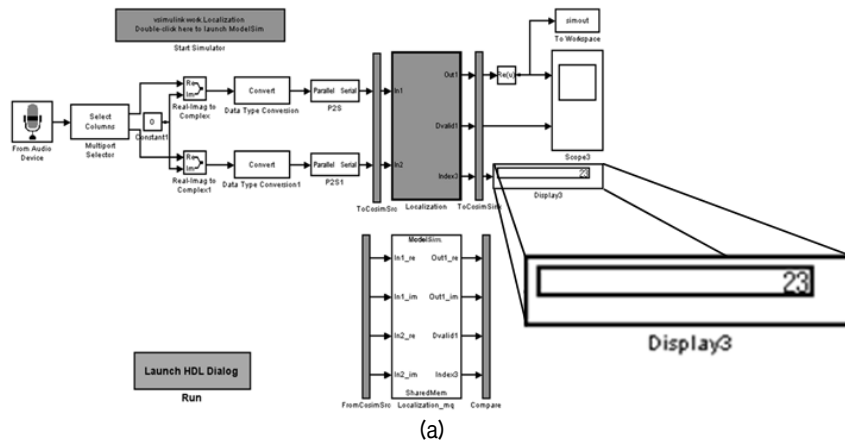


그림 7. CC 연산의 타이밍 다이어그램
Fig. 7. Timing diagram of the CC calculation.

의 Simulink를 이용하였고 HDL work flow advisor 기능을 이용하여 설계하고 시뮬레이션 및 검증하였다. Matlab의 HDL변환 가능한 블록들을 통해 설계하고 실제 이론값과 HDL변환된 블록의 결과를 비교하여 설계상의 오류가 있는지 검증하였다. 그림 8은 GCC 연산 및 최대 시간 지연 샘플의 검출을 위해 300 Hz 정현파 신호가 마이크로폰 어레이의 75도에 위치하였을 경우 최대 지연 샘플값을 측정하는 것이다. 이 경우 두 신호간에 23 샘플만큼 지연되어 입력되며 해당 샘플값을 맵핑테이블과 비교하였을 때도 74도로 거의 일치함을 확인할 수 있었다. 그림 8의 (a)는 최대 지연값에 대한 샘플값을 검증하고 HDL을 이용한 하드웨어를 설계 블록과 시뮬레이션 블록을 비교하기 위한 전체 시스템의 구성이다. 그림 8의 (b)는 (a)에서 연산된 결과의 시뮬레이션값과 하드웨어 구현한 연산 값의 오차를 나타낸 그래프로 하드웨어 설계 시 시뮬레이션과



Messages		
↕ /Localization/dk	S0	
↕ /Localization/reset	S1	
↕ /Localization/dk_en...	S1	
↕ /Localization/In1_re	-197	1.9724 1.9768 1.9808 1.9845 1.9877 1.9906 1.9931 1.9952 1.9970 1.9983 1.9993 1.9999 1.0000 1.9999 1.9993 1.9983 1.9970 1.9952 1.9931 1.9906
↕ /Localization/In1_im	0	0
↕ /Localization/In2_re	6990	-1.8527 1.8423 1.8315 1.8205 1.8091 1.7974 1.7854 1.7731 1.7605 1.7476 1.7344 1.7209 1.7072 1.6931 1.6789 1.6643 1.6495 1.6344 1.6191 1.6036
↕ /Localization/In2_im	0	0
↕ /Localization/ce_out	S1	
↕ /Localization/Out1_re	0	49... 5160... 5413... 5665... 5916... 6166... 6415... 6661... 6906... 7149... 7389... 7626... 7861... 8093... 8321... 8546... 0
↕ /Localization/Out1_im	0	97... 19608... 19393... 19116... 18776... 18375... 17912... 17387... 16802... 16156... 15450... 14685... 13862... 12980... 12042... 11048... 0
↕ /Localization/Dvald1	S0	
↕ /Localization/Index3	23	0

테이블 맵핑용 결과 인덱스 (인덱스 23은 74도(degree)에 해당)

그림 8. Simulink환경에서의 하드웨어 설계 시뮬레이션 및 검증 (a) CC 검증을 위한 하드웨어 구성 및 테스트벤치 (b) 하드웨어를 통한 결과와 이론 결과의 비교 (c) CC 및 최대 샘플값 결과의 타이밍 다이어그램
 Fig. 8. Hardware design simulation and verification with simulink. (a) Hardware configuration and test bench for CC verification (b) Comparison of hardware result and ideal result (c) Timing diagrams of CC and maximum delay sample.

의 오차가 없음을 확인 할 수 있었다. 그림 (c)는 하드웨어로 설계된 GCC 연산을 통해 얻어진 최대 지연 샘플값의 타이밍 다이어그램이다.

IV. 실험 및 성능 평가

본 연구에서 제안된 방법의 음원 방향 추정의 성능 검

증을 위해 네 개의 마이크로폰을 십자 형태의 어레이로 구성하여 마이크로폰과 동일한 높이에 존재하는 음원의 방향 추정 실험을 하였다. 실험 환경은 정확한 검증을 위한 무전파실험실과 일반적인 룸 환경에서 실험하였다. 실험 조건과 실험에 사용된 하드웨어의 구성은 표 2과 같다. 실험에서 사용한 음원은 일반 스피커를 사용하여 반복되는 음악 소리와 협대역 음원을 주파수 발생기를 이용하

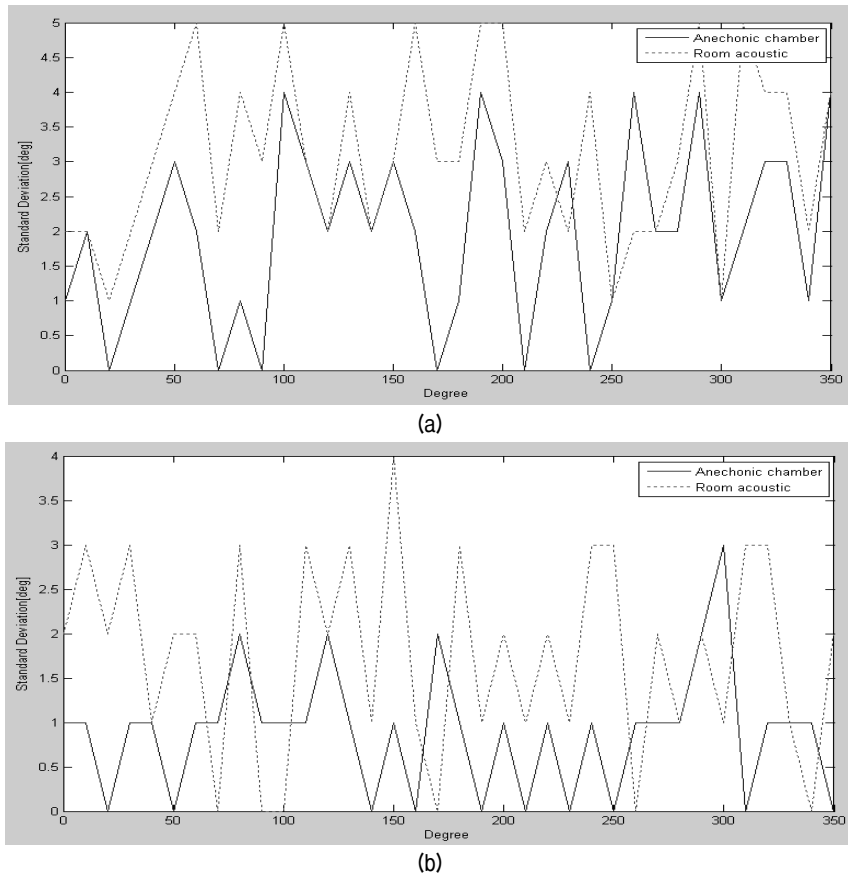


그림 9. 실험 결과 (a) 음악 음원의 표준 편차 (b) 협대역 음원 (500 Hz)의 표준편차
 Fig. 9. Experimental result. (a) Standard deviation of music signal (b) Standard deviation of narrow band signal (500 Hz).

표 2. 실험환경 및 하드웨어 구성
 Table 2. experimental environment and hardware setup.

Place	anechonic chamber, room acoustic
Microphone gap	30 cm (4 condenser type)
Sound source	narrow band (500 Hz), music signal
Distance	2 m
Sampling Frequency	96 kHz
Hardware specification	
FPGA	Altera Cyclone ii (x2)
Sound Codec	WM8731 (x2)
Experimental environment	

였다. 장소의 제약 때문에 음원을 움직이는 대신 고정된 음원에서 마이크로폰 어레이를 회전 하여 음원의 방향 정보를 획득하였다. 시스템에 대한 실험은 80 dBA크기의

협대역 음원 그리고 음악 신호를 이용하였으며 무전파실험실과 일반적인 잔향이 존재한 실험실 내에서 각각 실험하였다. 그림 9는 음악 음원이 존재할 때 전방향에 대한 음원 방향의 각도 표준 편차를 그래프로 나타내었다. 그리고 무전파실험실에서의 결과와 실험실 내부에서의 결과에 대해 비교하였다.

음원 신호는 실험실내에서의 실험결과 최대 각도 오차는 약 5도를 나타내었으며, 무전파실험실에서는 약 4도의 오차를 보였다. 그리고 500 Hz의 협대역 음원은 무전파실험실의 경우에는 3도 그리고 일반적인 인 방의 경우에는 4도를 나타내었다. 비교적 광대역인 음원 신호의 경우보다 협대역 음원의 경우 더 정확한 방향 추정이 가능한 것을 알 수 있다. 음원의 구역을 나누는 기법을 이용하여 기존 시스템들이 마이크로폰 어레이와 일직선에 존재하는 음원의 경우 각도오차가 큰 문제점을 어느 정도 극복하여 전방향에 대해 비교적 안정적인 각도 결과를 획득할 수 있었다.

제안한 시스템의 경우 50 MHz의 클럭을 사용하여 음성이 입력된 후 각도값이 도출 되는 전 과정에 걸리는

시간은 약 69 msec의 시간이 소요되며, 100 MHz의 DSP 프로세서를 사용한 경우 257.3 msec의 처리 시간이 요구된다. DSP 프로세서를 이용한 음원방향 인지 시스템의 경우는 [11]을 참고 하였다. PC 기반으로 음원 국지화 시스템의 실시간 특성을 측정 한 결과는 120 msec의 시간 지연을 보였다. PC기반 설계의 경우 PC 환경은 1.8 GHz 듀얼 CPU 3.25 GB RAM을 이용하였으며, LabVIEW환경에서 리얼타임블럭을 사용하여 전체 시스템의 딜레이를 측정하였다. FPGA를 이용하였을 경우 기본적으로 다른 시스템에 비해 실시간 응답이 좋은 것을 확인 할 수 있으며, 마이크로폰 채널수가 증가할수록 다른 프로세서의 응답성은 느려지는 반면 FPGA 기반 설계의 경우 응답성에 변함이 없다. 이런 특징을 통해 신호처리부분의 많은 부하가 걸리는 부분에 대해 하드웨어 가속기로서의 역할을 담당할 수 있다.

V. 결론

십자 형태로 구성된 네 개의 마이크로폰 어레이를 이용하여 2차원 상에서 실시간으로 비교적 정확한 음원의 방향을 탐지할 수 있는 시스템에 관하여 연구하였다. 개발 시스템은 음원의 방향을 예측할 수 있으며, 가장 많은 연산이 요구되는 GCC 부분을 FPGA를 이용해 구현하여 다른 프로세서를 사용한 시스템에 비해 실시간 처리속도를 높였다. 그리고 음원 존재 가능 구역을 구분하여 더 정확하게 전방향 음원 방향 탐지가 가능하도록 하였다.

본 연구는 작은 면적에서 하드웨어 구현이 가능하여 DAQ가 요구되는 기존 연구의 제약을 극복하였다. 그리고 실제 실험에 마이크로폰이 저가형이며 네 개만을 이용하여 실내 환경에서 어느 정도 신뢰성 있고 안정적인 음원 방향 추정이 가능하다. 본 연구의 방식을 이용한 로봇의 HRI 시스템이나 스마트 감시카메라와 같은 많은 응용 분야에 활용이 가능할 것으로 기대된다.

참고 문헌

1. 구자현, "Rayleigh의 소리의 방향 지각연구에 대한 과학사적 고찰," *한국음향학회지*, 21권, 8호, 695-702쪽, 2002.
2. 구자현, "미국 음향학자 앨프레드 메이어의 음향학 연구," *한국음향학회지*, 28권, 1호, 44-50쪽, 2009.
3. Azizul H. Quazi, "An Overview on the Time Delay Estimate in Active and Passive System for Target Localization," *IEEE Transactions on Acoustic, Speech, and Signal Processing*, vol. ASSP-29, no. 3, pp. 527-533, 1981.
4. Kim, Young Ik, "Spatial Hearing Algorithms Based on Binaural

Zero-Crossings: Sound Source Localization, Segregation and Dereverberation," Doctoral thesis, KAIST, 2007.

5. Jindong Chen, Jacob Benesty, and Yiteng Huang, "Time Delay Estimation in Room Acoustic Environments: An Overview," *EURASIP J. Appl. Signal Processing*, vol. 2006, pp. 1-19, 2006.
6. Charles H. Knapp, "The Generalized Correlation Method for Estimation of Time Delay," *IEEE Transactions on Acoustic, Speech, and Signal Processing*, vol. ASSP-24, no. 4, pp. 320-327, 1976.
7. Richard M. Stern, "Models of Binaural Interaction," Chapter in *Handbook of Perception and Cognition*, volume 6: Hearing.
8. Byounggho Kwon, Youngjin Park and Youn-sik Park, "Sound Source Localization Method Using Spatially Mapped GCC Functions," *한국소음진동공학회논문집*, 19권, 4호, 355-362쪽, 2009.
9. H. Atmoko, D. C. Tan, G. Y. Yian and Bruno Fazenda, "Accurate Sound Source Localization in a Reverberant Environment using Multiple Acoustic Sensors," *Meas. Sci. Technol.*, 19, 024001, 2008.
10. Jean-Marc Valin, Francois Michaud, Jean Rouant and Dominic Letourneau, "Robust Sound Source Localization Using a Microphone Array on a Mobile Robot," *Proceedings of the 2003 IEEE/RSJ Intl. Conference on Intelligent Robots and Systems*, vol. 2, pp. 1228-1233, 2003.
11. Russ Duren, Jeremy Stevenson and Mike Thompson, "A Comparison of FPGA and DSP Development Environments and Performance for Acoustic Array Processing," *Circuits and Systems, MWSCAS 2007. 50th Midwest Symposium*, pp. 1177-1180, 2007.

저자 약력

•김 태 완 (Taewan Kim)



2007년 경희대학교 전자·전파공학과 (공학사)
 2009년 경희대학교 전자·전파공학과 (공학석사)
 2011년 현재 경희대학교 전자·전파공학과 박사수료
 ※ 주관심 분야 : 디지털 시스템 설계, 얼굴인식, 음향신호처리, SOC 설계

•김 동 훈 (Donghoon Kim)



2009년 경희대학교 전자·전파공학과 (공학사)
 2011년 경희대학교 전자·전파공학과 (공학석사)
 2011년 현재 LG넥스원(주) 연구개발본부 Maritime 연구소
 ※ 주관심 분야 : 음향신호처리, 디지털 시스템 설계, SOC 설계, 표적탐지 및 추적

•정 연 모 (Yunmo Chung)



1980년 경북대학교 학사
 1982년 KAIST 공학석사
 1987년 경제기획원 전산처리관
 1992년 미시간주립대학교 공학박사
 2011년 현재 경희대학교 전자정보대학 교수
 ※ 주관심 분야 : SoC 설계, 임베디드 시스템, RTOS