

자이로 센서를 이용한 가상 음원의 실시간 방향 제어 방법

한 철 수*

A Method for the Real-time Direction Control of Virtual Sound Sources using a Gyro Sensor

Cheolsu Han*

요 약

3차원 음향 기술은 청취자에게 현실감 있는 소리를 제공한다. 기존의 3차원 음향 기술은 원음을 더 정밀하고 충실하게 재현하는 것을 목표로 삼는 경우가 많았다. 본 연구에서는 청취자가 스스로 소리의 정보를 변경하고 들을 수 있도록 하는 방법에 대하여 검토를 진행하였다. 본 논문에서는 청취자가 자이로 센서를 이용하여 음향 신호의 방향을 실시간으로 제어하고, 그 정보에 맞추어 음향 신호를 합성하고 재생하는 방법을 제안하고 시스템을 구현하였다. 구현 시스템의 성능을 평가한 결과, 청취자는 음원 방향을 실시간으로 바꿀 수 있었고, 안정적으로 합성된 소리를 들을 수 있었다. 또한, 새로운 시도를 통해 기존에는 없었던 새롭고 흥미로운 사용자 경험을 청취자에게 제공할 수 있었다.

Abstract

3-D audio technologies present realistic sound to listeners. Previous ones have often aimed to more precisely and faithfully reproduce the original sound. In this study, we review how listeners can interact with and listen to sound information. In particular, it considers the case of a listener using a gyro sensor to control the perceived direction of arrival of a sound signal in real-time. A method to synthesize and reproduce the sound signal according to this information is proposed and implemented. The performance evaluation results for the implemented system shows that the listener is able to change the apparent direction of the sound source in real time and maintain a stable perception of the synthesized sound. In addition, new trials show it is possible to provide listeners with novel and interesting user experiences that were not previously available.

Keywords

3-D audio, head-related transfer function, gyro sensor, real-time, user experience

* 청주대학교 전자공학과 조교수
- ORCID: <https://orcid.org/0000-0001-9515-4392>

• Received: May 26, 2023, Revised: Jul. 07, 2023, Accepted: Jul. 10, 2023
• Corresponding Author: Cheolsu Han
Dept. of Electronic Engineering, Cheongju University, 298
Daeseongro, Cheongwon-gu, Cheongju, Chungbuk, 28503, Rep. of Korea
Tel.: +82-43-229-8446, Email: cheolsuhan@cju.ac.kr

1. 서 론

인간은 소리를 통해 다양한 정보를 얻는다[1]. 음성을 통해 다른 사람의 의견을 들을 수도 있고, 경고음이나 사물의 충돌음으로부터 위험을 감지할 수도 있다. 앞만 볼 수 있는 시각과 달리 청각은 뒤에서 들린 소리도 알아챌 수 있으며, 촉각처럼 가까이 가지 않아도 정보를 얻을 수 있는 것이 특징이다[2].

또한, 인간은 소리의 울림을 통해 소리가 발생한 장소의 공간감도 느낄 수 있고, 두 귀에 들린 소리의 미세한 시간 차이(ITD, Interaural Time Difference)와 레벨 차이(ILD, Interaural Level Difference)를 통해 소리의 방향도 느낄 수 있다[1]-[3].

3차원 음향(3-D audio) 기술을 이용하면 우리가 듣는 소리에 다양한 3차원 음향 정보를 부가할 수 있다[2]-[5]. 예를 들어, 듣고 있는 노래를 마치 콘서트홀에서 듣고 있는 것처럼 울림과 공간감을 조절할 수 있다.

또한, 소리가 모노(Monaural) 방식으로 녹음되었거나 합성된 경우에도 3차원 음향 기술을 이용하면 스테레오(Stereophonic)나 다채널 음향 신호로 바꾸어 실감나게 듣는 것도 가능하다[2]-[4]. 이 때문에 3차원 음향 기술은 최신 영화나 비디오 게임기 등에서 현실감 있는 음향 신호를 제작하는 데 널리 이용되고 있다.

3차원 음향 기술 중 대표적인 방법은 머리전달함수(HRTF, Head-Related Transfer Function)를 이용하는 것이다[1]-[3]. HRTF는 원거리의 음원으로부터 두 귀 입구까지의 전달함수로서, 음원의 위치로부터 두 귀 입구까지의 소리의 반사와 회절 등을 포함한 다양한 정보를 내포한다.

HRTF는 음원의 3차원 방향을 독립 변수로서 갖고, 왼쪽 귀와 오른쪽 귀에 대하여 한 쌍씩 존재한다. 예를 들어, 모노 음원에 전방 1m에서 측정된 HRTF 한 쌍을 합성하고 헤드폰을 통해 듣게 되면 소리는 모노 음원이 아니라 마치 전방 1m에서 발생하고 있는 것처럼 느껴진다[1][2].

3차원 음향 기술의 다수는 녹음되었거나 합성된 소리를 정밀하게 재현하는 데 목적을 둔다[2][4][6]. 예를 들어, 스피커 재생의 경우에는 녹음 당시의 마

이크로폰 배치와 똑같은 형태로 스피커를 배치함으로써 녹음된 소리를 정밀하게 재현시킬 수 있다.

한편, 헤드폰 재생의 경우에는 헤드폰으로 소리를 듣는 중에 청취자가 머리를 움직이면 그에 따라서 소리도 함께 움직여 버리는 느낌을 받게 되고 이것은 소리의 현실감을 떨어뜨린다[4].

따라서 헤드폰 사용 시에는 사람의 머리 움직임을 감지해 실시간으로 HRTF를 합성하여 소리의 방향이 움직이지 않도록 해주는 방법이 사용되고 있다[6]. 이처럼 음원을 실제와 똑같이 정밀하고 현실감 있게 들려주기 위한 검토는 다수 보고되었지만, 청취자가 스스로 음원의 위치나 방향을 실시간으로 바꾸면서 소리를 듣는 방법에 대해서는 많은 검토가 이루어지지 않았다[7].

본 연구에서는 청취자가 소리의 방향을 자신이 듣고 싶은 대로 실시간으로 조정하면서 들을 수 있는 3차원 음향 시스템에 대하여 검토한다. 이를 위해서는 청취자가 어떠한 방식으로 소리의 방향을 실시간으로 조정할 수 있게 할 것인지 그 방법에 대한 검토가 필요하고, 이를 통해 새로운 방식의 조종 장치의 개발이 요구된다[7]. 또한, 청취자가 조정된 방향 정보를 이용해 소리를 실시간으로 합성하고 재생할 수 있는 정밀한 음향 시스템의 개발이 필요하다[7].

자이로 센서(Gyroscope)는 사물의 회전각을 쉽고 정밀하게 측정할 수 있는 장치이다[8][9]. 특히 최근 개발된 미세 전자 기계 시스템(MEMS, Micro Electro Mechanical Systems) 방식의 자이로 센서는 정밀하게 회전각을 측정할 수 있을 뿐만 아니라 크기도 작고 값도 싸다[9]. 이러한 이유로 소형 장치의 내부에 실장하여 회전각 정보를 취득하는 데 널리 이용되고 있다[7]-[11].

3차원 음향 시스템에서도 사람의 머리 움직임을 감지하기 위해서 자이로 센서가 자주 이용된다[10]. 기존 장치들은 자이로 센서를 머리에 부착해 머리 움직임에 따른 소리의 변화를 없애기 위해 노력하였다[10]. 그러나 반대로 자이로 센서를 손에 쥐고 회전시키면서 얻어진 회전각 정보를 HRTF의 변경에 이용한다면 청취자가 능동적으로 소리의 방향을 조정할 수 있는 조종기를 만들 수 있을 것이다[7].

본 논문에서는 청취자가 능동적으로 가상 음원의 방향을 조정하면서 3차원 음향 신호를 들을 수 있는 실시간 3차원 음향 시스템을 제안한다. 이를 위해 자이로 센서를 이용하여 음향 신호의 방향을 실시간으로 제어하고 그 정보에 맞추어 음향 신호를 실시간으로 합성하고 재생하는 방법을 제안한다. 이와 같은 기술의 개발은 기존 장치들에서는 느낄 수 없었던 새롭고 흥미로운 사용자 경험(UX, User Experience)을 청취자들에게 제공할 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문에서 제안하는 자이로 센서를 이용한 가상 음원의 실시간 방향 제어 방법에 대하여 설명한다. 3장에서는 제안 방법을 이용한 시스템 구현에 대하여 설명한다. 4장에서는 구현된 시스템의 동작을 평가하고 고찰한다. 마지막 5장은 결론 및 향후 과제이다.

II. 자이로 센서를 이용한 가상 음원의 실시간 방향 제어 방법

2.1 자이로 센서를 이용한 회전각 추정 방법

자이로 센서를 이용해 회전각을 추정하는 방법에 대하여 설명한다[9][10]. 회전각의 추정 흐름을 그림 1에 나타낸다.

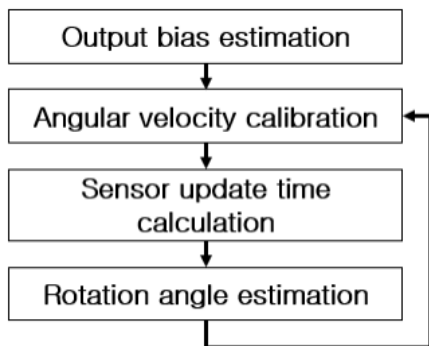


그림 1. 회전각의 추정 흐름

Fig. 1. Flow chart for the rotation angle estimation

자이로 센서는 회전 각속도를 측정할 수 있는 장치이고, 이 값을 이용해 일련의 과정을 통해 회전각을 추정할 수 있다. 이때 주의해야 할 것은 자이로 센서에서 출력되는 미가공 데이터(Raw data)에는 자

이로 센서의 출력 바이어스(Bias)가 포함된다는 것이다. 따라서 가장 먼저 해야 할 것은 출력의 바이어스를 미리 추정해 두고 추후 이값을 이용해 회전 각속도를 교정해야 한다[9][10][12]. 출력 바이어스 추정은 다음과 같이 이루어진다.

$$Z_{bias} = \frac{1}{N} \sum_{i=1}^N Z_{raw}(i) \quad (1)$$

여기에서 $Z_{raw}(i)$ ($i=1,2,\dots,N$)는 자이로 센서가 출력하는 회전 각속도의 미가공 데이터로서 일정 시간 간격으로 수집한 N 개의 표본이다. 표본을 수집할 때 자이로 센서가 움직이지 않도록 하는 것이 매우 중요하다. Z_{bias} 는 추정된 출력 바이어스로서 $Z_{raw}(i)$ 의 평균값이다.

출력 바이어스의 추정치가 얻어지면 이제 회전 각속도를 교정할 수 있다. 회전 각속도의 교정 방법은 다음과 같다[10].

$$Z_{calib}(n) = Z_{raw}(n) - Z_{bias} \quad (2)$$

여기에서 $Z_{raw}(n)$ 은 회전 각속도의 미가공 데이터이고, Z_{bias} 는 식 (1)에서 추정한 자이로 센서의 출력 바이어스이다. 회전 각속도의 미가공 데이터에서 출력 바이어스를 뺀으로써 회전 각속도를 교정할 수 있다. $Z_{calib}(n)$ 은 교정된 회전 각속도이고, 단위는 deg/s이다.

다음으로 자이로 센서의 갱신 시간을 다음과 같이 산출한다[10].

$$\Delta t(n) = t_{now}(n) - t_{prev}(n) \quad (3)$$

여기에서 $t_{now}(n)$ 은 현재 시각을, $t_{prev}(n)$ 은 반복 처리의 직전 단계의 시각을 나타낸다. 즉, $t_{prev}(n)$ 는 $t_{now}(n-1)$ 과 같다. $\Delta t(n)$ 은 센서의 갱신 시간으로서 현재 시각과 반복 처리의 직전 단계의 시각의 차이이고, 단위는 초(s)이다.

마지막으로 회전각을 추정한다. 회전각의 추정에는 앞선 식 (2)와 식 (3)의 결과가 이용된다[10].

$$Yaw(n) = Yaw(n-1) + Z_{calib}(n) \cdot \Delta t(n) \quad (4)$$

여기에서 $Yaw(n)$ 은 최종적으로 추정된 회전각의 크기이고, $Yaw(n-1)$ 은 반복 처리의 직전 단계에서의 회전각의 크기이다. 교정된 회전 각속도인 $Z_{calib}(n)$ 에 센서의 갱신 시간 $\Delta t(n)$ 를 곱하면 갱신 시간 동안 회전한 각도를 계산할 수 있고, 이 값을 이전 회전각인 $Yaw(n-1)$ 에 더해 최종적으로 현재의 회전각을 추정한다. 회전각의 초깃값은 0도이다.

2.2 회전각 정보를 이용한 음원의 실시간 합성 및 재생 방법

추정된 회전각을 이용해 음원의 방향을 실시간으로 바꾸어 가면서 합성하고 재생하는 방법에 대하여 설명한다. 본 논문에서는 모노 음원에 대한 신호처리 방법과 스테레오 음원에 대한 신호처리 방법을 각각 제안한다.

모노 음원에 대한 신호처리 흐름도를 그림 2에 나타낸다. 그림에서 $s(t)$ 는 모노 음원을 나타낸다. $x_L(t)$ 와 $x_R(t)$ 는 각각 헤드폰을 통해 입력되는 왼쪽 귀 입력 신호와 오른쪽 귀 입력 신호이다. HRIR은 머리전달 임펄스응답(Head-Related Impulse Response)으로서 HRTF를 역 푸리에 변환(IFT, Inverse Fourier Transform)하여 얻을 수 있다[13].

a 는 음원의 방향을 나타내고, b 는 청취자가 자이로 센서를 이용한 조종기를 손으로 회전시켜 만든 회전각이다. $HRIR_L(a+b)$ 는 $a+b$ 의 방향각에 해당하는 왼쪽 귀용 HRIR이고, $HRIR_R(a+b)$ 은 오른쪽 귀용 HRIR이다.

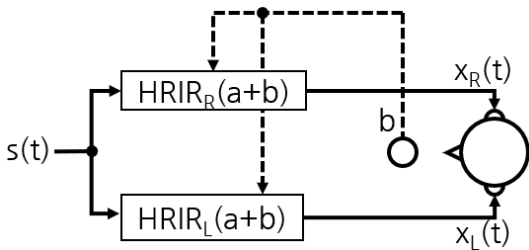


그림 2. 모노 음원에 대한 신호처리 흐름
Fig. 2. Signal processing flow for monaural sound sources

모노 음원에 대하여 두 귀 입력 신호인 $x_L(t)$ 와 $x_R(t)$ 는 다음과 같이 얻어진다.

$$x_L(t) = s(t) * HRIR_L(a+b) \quad (5)$$

$$x_R(t) = s(t) * HRIR_R(a+b) \quad (6)$$

여기에서 $*$ 는 컨볼루션(Convolution) 연산을 나타낸다[13]. HRIR은 청취자의 주위 360도의 모든 방향각에 대해서 5도 간격으로 측정되어 좌우 귀용한 쌍씩 존재한다[3]. 모노 음원의 방향 a 에 더 회전시키고 싶은 회전각 b 를 더하면 음원의 새로운 방향이 얻어진다. 이 새로운 방향에 대한 HRIR을 모노 음원에 컨볼루션 연산하면 두 귀 입력 신호한 쌍이 얻어지게 된다.

이렇게 얻어진 두 귀 입력 신호를 청취자가 헤드폰을 통해 듣게 되면 음원 $s(t)$ 는 마치 $a+b$ 의 방향에서 들리는 것처럼 느껴지게 된다[4][6].

스테레오 음원에 대한 신호처리 흐름도를 그림 3에 나타낸다. 그림에서 $s_L(t)$ 는 스테레오 음원의 L 채널 출력 신호를, $s_R(t)$ 는 R 채널 출력 신호를 각각 나타낸다. a_L 은 L 채널 스피커의 방향을, a_R 은 R 채널 스피커의 방향을 각각 나타낸다.

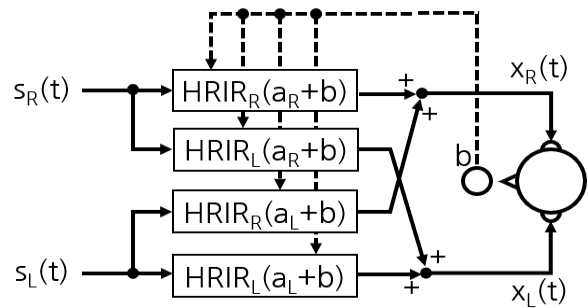


그림 3. 스테레오 음원에 대한 신호처리 흐름
Fig. 3. Signal processing flow for stereophonic sound sources

스테레오 음원에 대하여 두 귀 입력 신호인 $x_L(t)$ 와 $x_R(t)$ 는 아래와 같이 얻어진다.

$$x_L(t) = s_L(t) * HRIR_L(a_L+b) + s_R(t) * HRIR_L(a_R+b) \quad (7)$$

$$x_R(t) = s_R(t) * HRIR_R(a_R + b) + s_L(t) * HRIR_R(a_L + b) \quad (8)$$

위와 같이 얻어진 두 귀 입력 신호를 청취자가 헤드폰을 통해 듣게 되면, a_L 과 a_R 의 방향에 스피커를 각각 두고 소리를 들었을 때의 느낌이 회전각 b 만큼 회전하여 들리는 것처럼 느껴지게 된다[4][6].

III. 시스템 구현

2장에서 논한 자이로 센서를 이용한 가상 음원의 실시간 방향 제어 방법을 이용해 시스템을 구현하였다. 시스템 블록도를 그림 4에 나타낸다. 시스템은 크게 회전각의 추정을 담당하는 부분과 음원의 실시간 합성 및 재생을 담당하는 부분으로 나뉜다. 그림에서 화살표의 방향은 정보의 전달 방향을 나타내고, 단방향과 양방향이 존재한다. 화살표 옆의 단어는 통신에 사용된 통신 방식을 나타낸다.

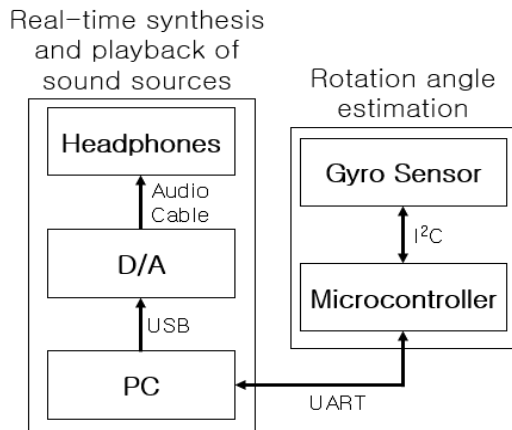


그림 4. 시스템 블록도
Fig. 4. System block diagram

마이크로 컨트롤러는 2.1절의 방법을 이용하여 자이로 센서가 측정한 회전 각속도를 이용해 회전각을 추정한다. PC(Personal Computer)에서는 일정 시간 간격으로 마이크로 컨트롤러에 회전각 정보를 요청하고, 그 정보를 이용해 음원의 실시간 합성 및 재생을 수행한다. 이때 2.2절의 방법이 이용된다. PC에서 합성된 두 귀 입력 신호는 D/A(Digital-to-analog converter)를 통해 아날로그 신호로 바뀌고 오디오 케이블을 통해 헤드폰으로 전달되어 청취자의

귀로 입력된다. 본 논문에서 시스템을 구현하는 데 사용한 시스템 구성은 표 1과 같다.

표 1. 시스템 구성
Table 1. System configuration

PC	Windows PC
D/A	Creative sound BlasterX G6
Headphones	Sony MDR-1A
Microcontroller	Arduino nano
Gyro sensor	TDK MPU-6050

PC에서 동작하는 실시간 처리 프로그램의 그래픽 사용자 인터페이스(GUI, Graphical User Interface)를 그림 5에 나타낸다. 그림에서 왼쪽 위의 두 버튼은 음원 방향의 제어를 위한 자이로 센서 기반의 조종기와 연결하기 위한 버튼이다. 이 버튼을 이용해 PC와 마이크로 컨트롤러는 UART(Universal Asynchronous Receiver/Transmitter) 통신을 초기화하거나 연결 상태를 확인할 수 있다.

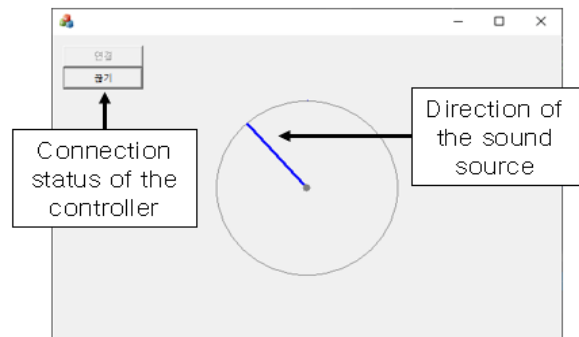


그림 5. 실시간 처리 프로그램의 GUI
Fig. 5. GUI of the real-time processing program

화면 중앙의 점은 청취자의 위치를 나타내고, 큰 원은 청취자의 앞뒤 좌우 360도의 방향을 나타낸다. 중앙의 점과 큰 원을 잇는 굵은 선은 음원의 현재 방향을 나타낸다. 지금 그림은 음원이 전방에서 왼쪽으로 약 40도 정도 회전해 있는 상태를 나타낸다. 음원의 초기 방향은 0도이다.

마이크로 컨트롤러를 제어하는 프로그램과 PC의 실시간 처리 프로그램을 모두 C++ 언어를 이용하여 구현하였다. 실시간 처리 프로그램의 GUI는 MFC(Microsoft Foundation Class) 라이브러리를 이용하여 제작하였고, 오디오 입출력을 위한 사운드카드 제어는 PortAudio 라이브러리를 이용하였다[14].

오디오 처리를 위한 샘플링 주파수는 44.1kHz를 이용하였고, 오디오 버퍼의 크기는 1,024로 하였다. 따라서 방향 정보를 취득하고 음원의 실시간 합성 및 재생을 수행하는 데까지 최대 약 23ms 내에서 처리를 완료해야 한다. 음원의 실시간 합성에는 오버랩-애드(Overlap-Add) 필터링 기법을 이용하였다[13].

또한, 2.2절의 방법과 같이 모노 음원과 스테레오 음원에 대하여 모두 처리할 수 있도록 프로그래밍하였다. HRIR은 KEMAR 더미 헤드(Dummy Head) 마이크로폰으로 측정된 HRIR을 이용하였고[3], 청취자의 앞뒤 좌우의 수평 360도 방향에 대해서 5도 간격으로 구성된 좌우 72쌍의 HRIR을 사용하였다.

IV. 동작 수행 평가 및 고찰

4.1 회전각 추정

구현된 음원의 방향각 제어 장치가 정상적으로 동작하는지를 판별하기 위해서 동작 수행 평가를 진행하였다. 청취자는 조종기를 자신의 앞에 두고 두 손으로 편하게 잡은 후, 시계방향 또는 반시계 방향으로 자유롭게 회전시켜 보도록 지시하였다. 이때 실시간으로 추정되는 회전각의 크기를 PC에서 요청하여 저장하였다.

회전각 추정의 결과를 그림 6에 나타낸다. 그림에서 가로축은 시간을, 세로축은 조종기가 향한 수평 회전각을 나타낸다. 단위는 각각 초(s)와 도(deg)이다. 세로축의 0도는 사용자의 정면을 나타내고, +90도는 정확히 왼쪽, -90도는 오른쪽, ±180도는 뒤를 낸다. +쪽은 조종기의 방향이 +Z축을 기준으로 반시계 방향으로 회전된 것을, -쪽은 시계방향으로 회전된 것을 뜻한다.

그림에서 조종기의 사용자는 처음 약 2초간 조종기를 정면을 향하게 그냥 두었다가, 이후 5초까지 반시계 방향으로 회전시켰고, 다시 10초까지 시계방향으로 회전시켰다는 것을 알 수 있다. 이후 사용자는 자유롭게 두 방향을 번갈아 가면서 회전을 계속했다. 반시계 방향으로는 최대 약 106도까지 회전시켰고, 시계방향으로는 최대 약 125도까지 회전시켰다는 것을 알 수 있다.

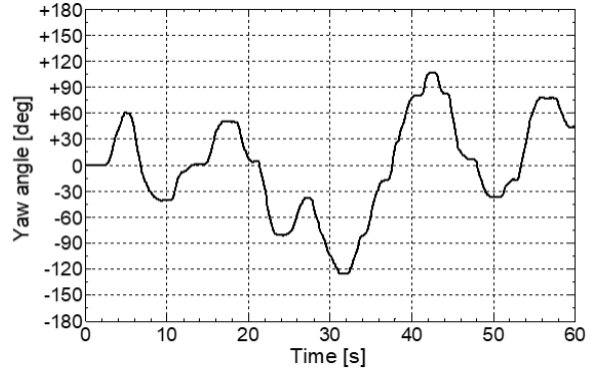


그림 6. 회전각 추정 결과

Fig. 6. Rotation angle estimation results

자이로 센서를 이용한 회전각 추정 장치의 대표적인 성능 지표 중 하나는 드리프트 오차(Drift Error)의 증가율이다[10][12]. 식 (4)에서 회전각을 추정하기 위해 새롭게 회전한 각도를 기존 회전각에 더하게 되는데, 이 과정이 반복될수록 회전각에는 오차가 누적되게 된다. 이처럼 시간이 지날수록 점점 커지는 회전각의 오차를 드리프트 오차라고 한다. 선행 연구에서 보고된 드리프트 오차의 평균 증가율은 약 0.1406 deg/s였다[10].

본 논문에서 구현한 장치에 대하여 드리프트 오차의 평균 증가율을 측정된 결과, 그 값은 약 0.0072 deg/s였다. 이것은 회전각의 측정 결과에 포함되는 오차가 1초당 약 0.0072도씩 증가한다는 것을 뜻한다. 선행 연구의 결과와 비교해서 약 19.53배 개선된 것을 알 수 있다.

이상의 결과로부터 구현된 조종기 시스템은 조종기가 향한 회전각을 실시간으로 추정하고 그 값을 PC로 문제없이 전달할 수 있는 것으로 판단된다.

4.2 음원의 실시간 합성 및 재생

음원의 실시간 합성 및 재생의 정밀도를 확인하기 위해 측정 실험을 진행하였다. 측정 실험에는 음향계(Acoustical system)의 임펄스응답을 빠르고 정확하게 측정할 수 있도록 고안된 OATSP(Optimized Aoshima's Time-Stretched Pulse signal) 신호를 이용하였다[15]. 그림 7에 OATSP 신호를 나타낸다. (a)는 OATSP 신호이고, (b)는 역OATSP 신호이다. 가로축은 시간 인덱스를 나타내고, 세로축은 상대적 크기

를 나타낸다. OATSP 신호는 역OATSP 신호와 콘벌루션 연산을 하게 되면 다음과 같이 지연된 임펄스 응답이 된다[15].

$$\delta[n-(N-1)] = h[n] * h^{-1}[n] \quad (9)$$

여기에서 $h[n]$ 은 OATSP 신호를, $h^{-1}[n]$ 는 역 OATSP 신호를 나타낸다. n 은 시간 인덱스를 나타내고, N 은 OATSP 신호의 길이이다. $\delta[n-(N-1)]$ 은 콘벌루션 연산을 통해 얻은 지연된 임펄스 응답이고, 임펄스 응답은 $N-1$ 만큼 지연된다.

그림 7의 두 신호를 식 (9)와 같이 콘벌루션 연산하여 얻은 지연된 임펄스 응답을 그림 8에 나타낸다.

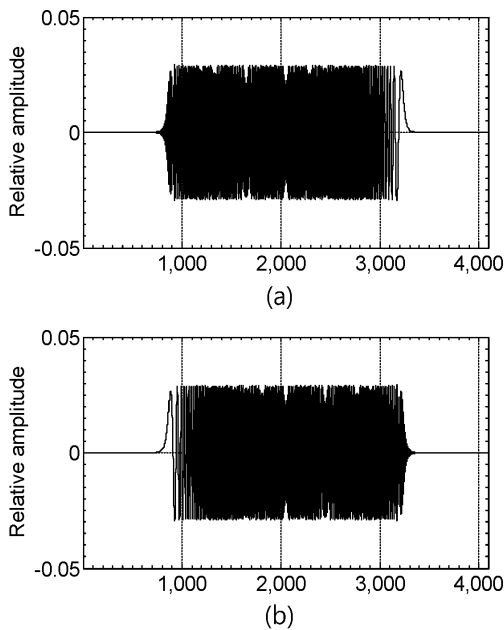


그림 7. OATSP 신호 (a) OATSP (b) 역OATSP
Fig. 7. OATSP signals (a) OATSP (b) Inverse OATSP

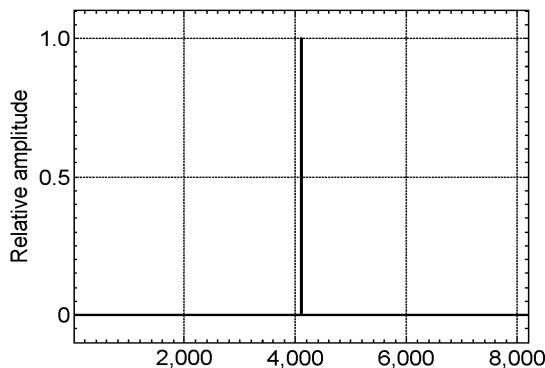


그림 8. OATSP 신호로부터 얻어진 임펄스 응답
Fig. 8. Impulse responses measured with OATSP signals

2.2절의 방법을 이용해 구현된 음원의 실시간 합성 및 재생 시스템의 임펄스 응답을 측정하였다. 그림 2의 HRIR을 임펄스 신호로 교체한 후, 모노 음원 $s(t)$ 로 OATSP 신호를 사용하고 출력하여 두 귀 입력 신호 $x_L(t)$ 와 $x_R(t)$ 를 녹음하게 되면, 녹음된 두 신호에는 OATSP의 특성뿐만 아니라 구현 시스템의 음향 특성이 포함된다[15]. 따라서 식 (9)와 같이 녹음된 두 귀 입력 신호에 역OATSP 신호를 콘벌루션 연산하여 구현 시스템의 임펄스 응답을 얻을 수 있다.

측정 실험에서는 오디오 입출력의 샘플링 주파수를 48kHz로 설정하였고, OATSP 신호의 길이를 65,536으로 설정하였다. OATSP 신호를 12번 연속해서 출력하고 동시에 녹음을 진행하였다. 녹음된 신호는 앞뒤 1회분씩 제거한 10회분의 신호에 대하여 가산 평균을 취하고, 역OATSP 신호와 콘벌루션 연산하여 시스템의 임펄스 응답을 산출하였다. 임펄스 응답의 측정 결과를 그림 9에 나타낸다.

그림 9에서 가로축은 시간 인덱스를 나타내고, 세로축은 상대적 크기를 나타낸다. 임펄스 응답은 임펄스가 가운데 위치하도록 조정하고 1023의 길이를 갖는 구형 창(Rectangular window)을 이용해 추출한 것이다[13].

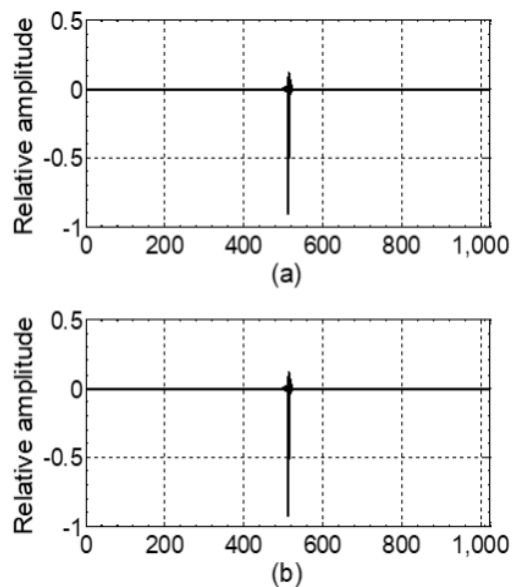


그림 9. 측정된 임펄스 응답 (a) 왼쪽 신호 (b) 오른쪽 신호

Fig. 9. Measured impulse responses (a) Left signal (b) Right signal

그림의 (a)는 왼쪽 귀 입력까지의 임펄스응답을, (b)는 오른쪽 귀 입력까지의 임펄스응답을 나타낸다.

시간 축 정보인 임펄스응답을 푸리에변환을 통해 주파수 축 정보인 전달함수로 변환하였다[13]. 전달함수의 진폭 특성을 그림 10에 나타낸다. 그림에서 가로축은 주파수를 나타내고, 세로축은 상대적 크기를 나타낸다. 단위는 각각 Hz와 dB이다. 그림에서 점선은 왼쪽 귀 입력까지의 진폭 특성이고, 실선은 오른쪽 귀 입력까지의 진폭 특성이다.

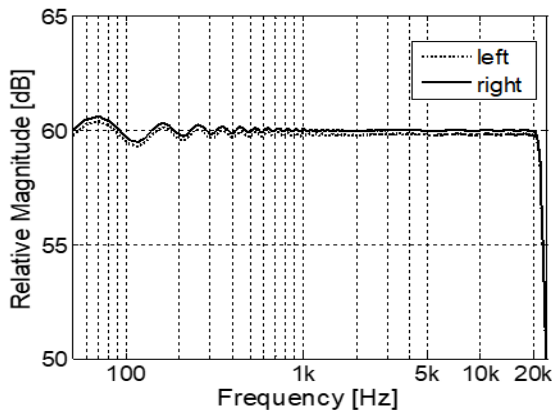


그림 10. 진폭 특성

Fig. 10. Amplitude characteristics

그림에서 약 300Hz부터 20kHz까지의 진폭 특성은 평탄하고 우수한 특성을 보인다. 20kHz 이상에서 보이는 골짜기는 샘플링 주파수를 48kHz로 설정한 것으로부터, D/A가 샘플링 주파수에 맞춰 창 함수를 씌운 결과라고 생각된다. 약 300Hz 이하에서 눈에 띄는 물결과 전대역에 걸친 두 선분 사이의 약 0.2dB의 차이는 보급형 D/A의 성능적 한계로 생각된다.

3장에서 설명한 것처럼 방향 정보를 취득하고 음원의 실시간 합성 및 재생을 수행하는 데까지 최대 약 23ms 이내에 처리를 완료해야 한다. 만약 그 사이에 오디오 프레임(Audio frame)의 합성 및 재생이 완료되지 않는다면 TSP 신호는 온전하게 모두 출력되지 못하고 따라서 정상적인 임펄스응답을 얻을 수 없게 된다.

그러나 그림 9와 10의 결과로부터 일반적인 임펄스응답이 얻어진 것을 알 수 있다. 이것은 구현된 시스템이 약 23ms 이내에 방향 정보를 취득하고 음원의 실시간 합성 및 재생을 모두 정상적으로 수행하고 있음을 뜻한다.

이상의 결과로부터 구현된 시스템은 음원을 실시간으로 합성하고 재생하는데 눈에 띄는 성능적 결함은 없는 것으로 판단된다.

본 논문에서는 자이로 센서 기반의 조종기를 이용해 청취자가 원하는 대로 음원의 방향을 실시간으로 움직이며 청취할 수 있는 3차원 음향 시스템을 제안하고 구현하였다. 앞선 실험 결과로부터 청취자는 음원의 방향을 실시간으로 바꿀 수 있고, 안정적으로 소리를 들을 수 있다는 것을 확인하였다.

또한, 새로운 시도를 통해 기존에는 없었던 새롭고 흥미로운 사용자 경험을 청취자에게 제공할 수 있게 되었다. 그러나 장치의 정밀도를 보다 높이기 위한 추가적인 검토가 필요하다고 생각된다. 또한, 방향뿐만 아니라 거리의 조정과 관련된 검토나 다수의 조정기를 이용한 새로운 시도 등도 필요하다고 생각된다.

V. 결론 및 향후 과제

3차원 음향 기술을 이용하면 소리에 다양한 음향 정보를 부가할 수 있다[2]-[4]. 이 때문에 최신 영화나 비디오 게임기 등에서 현실감 있는 음향 신호를 제작하는 데 3차원 음향 기술이 널리 이용되고 있다. 기존의 3차원 음향 기술은 원음을 더 정밀하고 충실하게 재현하는 데 초점을 맞춘 경우가 많았다 [2][4][6].

따라서 음원을 실제와 똑같이 정밀하고 현실감 있게 들려주기 위한 많은 연구가 진행되었다. 반면 청취자의 의도에 의해 청취자가 스스로 음원의 위치나 방향을 실시간으로 바꾸면서 소리를 듣는 방법에 대해서는 많은 검토가 이루어지지 않았다[7].

본 논문에서는 청취자가 스스로 소리의 정보를 변경하고 들을 수 있도록 하는 방법에 대하여 검토를 진행하였다. 청취자가 자이로 센서를 이용하여 음향 신호의 방향을 실시간으로 제어하고, 그 정보에 맞추어 음향 신호를 합성하고 재생하는 방법을 제안하였고 시스템을 구현하였다.

구현 시스템의 동작을 평가한 결과, 청취자는 음원 방향을 실시간으로 바꿀 수 있었고, 안정적으로 합성된 소리를 들을 수 있었다.

이러한 장치는 기존에는 없었던 새롭고 흥미로운 사용자 경험을 청취자에게 제공할 수 있다는 점에서 중요하다.

본 논문에서는 구현 시스템의 동작을 일부 평가하였으나 정밀도에 대한 추가적인 평가가 필요하다. 또한, 기술적으로 소리의 방향뿐만 아니라 소리의 거리감 조절을 위한 수단을 추가한다면 청취자에게 더 흥미로운 사용자 경험을 제공할 수 있을 것이다.

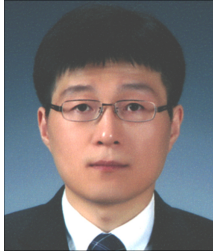
또한, 현재는 하나의 조종기만을 사용했지만 여러 개의 조종기를 이용하기 위해서는 처리 시간의 최적화를 위해 콘벌루션 연산보다는 고속 푸리에 변환을 이용한 주파수 영역에서의 처리 방법을 검토할 필요가 있다. 이상의 추가적인 검토를 본 논문의 향후 과제로 한다.

References

- [1] B. C. J. Moore, "An Introduction to the Psychology of Hearing", BRILL, Apr. 2013.
- [2] W. G. Gardner, "3-D Audio Using Loudspeakers", Ph.D. Dissertation, MIT Media Lab, Aug. 1997.
- [3] W. G. Gardner and K. D. Martin, "HRTF measurements of a KEMAR", *Journal of the Acoustical Society of America*, Vol. 97, No. 6, pp. 3907-3908, Jun. 1995. <https://doi.org/10.1121/1.412407>.
- [4] T. Lubeck, J. M. Arend, and C. Porschmann, "Binaural reproduction of dummy head and spherical microphone array data", *Journal of the Acoustical Society of America*, Vol. 151, No. 1, pp. 467-483, Jan. 2022. <https://doi.org/10.1121/10.0009277>.
- [5] J. Oh, "A Study on Relative Positioning based on Acoustic Signal of Smart phone", *Journal of IIBC*, Vol. 13, No. 1, pp. 229-234, Feb. 2013. <https://doi.org/10.7236/jiibc.2013.13.1.229>.
- [6] H. Kim, "A Study of 3-Dimensional Audio Reproduction Method based on Head Moving", *Journal of KIIT*, Vol. 12, No. 2, pp. 173-179, Feb. 2014. <https://doi.org/10.14801/kiitr.2014.12.2.173>.
- [7] C. Han and H. Kim, "Real-Time Control of the Positions of Virtual Sound Sources using a Gyroscope Sensor", *Proc. of KIIT Conference*, Daejeon, Korea, pp. 90-91, Jun. 2019.
- [8] K. Jeong and K. Lee, "Artificial Neural Network-Based Abnormal Gait Pattern Classification Using Smart Shoes with a Gyro Sensor", *Electronics*, Vol. 11, No. 21, pp. 1-11, Nov. 2022. <https://doi.org/10.3390/electronics11213614>.
- [9] M. Seo, "Build an Arduino drone and code it yourself", *Answerbook*, Apr. 2016.
- [10] C. Han and H. Kim, "A Real-time Yaw Rotation Angle Measurement System for Human Heads Using a Gyro Sensor", *Journal of KIIT*, Vol. 15, No. 8, pp. 59-66, Aug. 2017. <https://doi.org/10.14801/jkiit.2017.15.8.59>.
- [11] D. Cho and S. Park, "Design and Implementation of Frontal-View Algorithm for Smartphone Gyroscopes", *Journal of IIBC*, Vol. 12, No. 6, pp. 199-206, Dec. 2012. <https://doi.org/10.7236/JIWIT.2012.12.6.199>.
- [12] C. Han and H. Kim, "Comparison of Drift Errors Before and After Rotation of a Gyro Sensor", *Proc. of KIIT Conference*, Cheongju, Korea, pp. 94-95, Oct. 2020.
- [13] A. V. Oppenheim and R. W. Schaffer, "Digital Signal Processing", Pearson, Jan. 1975.
- [14] PortAudio, <http://www.portaudio.com>. [accessed: Jun. 20, 2023]
- [15] Y. Suzuki, F. Asano, H. Kim, and T. Sone, "An optimum computer-generated pulse signal suitable for the measurement of very long impulse responses", *Journal of the Acoustical Society of America*, Vol. 97, No. 2, pp. 1119-1123, Feb. 1995. <https://doi.org/10.1121/1.412224>.

저자소개

한 철 수 (Cheolsu Han)



2006년 2월 : 청주대학교

전자공학과(공학사)

2008년 2월 : 청주대학교 대학원

전자공학과(공학석사)

2013년 9월 : 일본 도호쿠대학

대학원 정보과학연구과(공학박사)

2013년 9월 ~ 2015년 8월 :

청주대학교, 한국교통대학교 시간강사

2015년 9월 ~ 현재 : 청주대학교 전자공학과 교육전담

조교수

관심분야 : 신호처리, 알고리즘, 프로그래밍