

FUS(Find Undetectable Sound): 들을 수 없는 소리의 방향을 찾기 위한 소리 탐지 애플리케이션

임 영 선¹ · 최 용 순^{2*}¹서강대학교 물리학과 학사과정^{2*}서강대학교 아트&테크놀로지학과 교수

FUS(Find Undetectable Sound): An application to indicate sound direction

Young-Sun Lim¹ · Yongsoon Choi^{2*}¹Undergraduate Course, Department of Physics, Sogang University, Seoul 04107, Korea^{2*}Professor, Department of Art&Technology, Sogang University, Seoul 04107, Korea

[요 약]

고주파 소음에 아이가 노출될 경우, 소음성 난청과 같은 문제가 발생할 수 있지만, 일반 성인은 가청주파수보다 높은 주파수의 소음을 인지하기 어렵다. 이를 해결하기 위해, 고주파 소음이 발생하는 위치를 소리 크기 비교 알고리즘을 통해 스마트폰으로 간편하게 찾는 애플리케이션을 개발하였다. 사전 실험을 통해 바 차트와 그래데이션을 활용한 초기 디자인을 선정한 후, 사용자들이 애플리케이션을 활용해 고주파 소리의 발생 위치를 찾는 실험을 진행하였다. 실험 결과, 다수의 사용자가 고주파 소음이 발생하는 위치를 찾을 수 있었다. 또한 설문조사 분석 결과, 이 서비스를 통해 들리지 않는 고주파 소음을 충분히 찾을 수 있다는 긍정적인 반응을 얻었다. 또한, 향후 디자인과 기능을 친숙하고 신뢰성 있게 개선하고 사용법을 단순화하는 방안을 논의하여 더욱 편하게 소리를 찾을 수 있는 향후 버전의 애플리케이션을 개발하고자 한다.

[Abstract]

When a child is exposed to high-frequency noise, problems such as noise-induced hearing loss may occur, but it is difficult for ordinary adults to recognize noise at a higher frequency than the audible frequency. To solve this problem, an application was developed to easily find a location where high-frequency noise occurs with a smartphone through the algorithm that compares the size of the sound wave. After selecting the initial design using bar charts and gradations through preliminary experiments, users conducted an experiment to find the location of high-frequency sound using the application. As a result of the experiment, a plurality of users were able to find a location where high-frequency noise occurred. In addition, as a result of the survey analysis, the service received a positive response that it was possible to find an audible high-frequency noise. In addition, we intend to develop future versions of applications that can find sound more conveniently by discussing ways to improve future designs and functions familiarly and reliably and simplify usage.

색인어 : 컴퓨터 청각, 음성 신호 처리, 모바일 애플리케이션 개발, HCI, UX/UI**Keyword** : Computer audition, Audio signal processing, Mobile application development, HCI, UX/UI<http://dx.doi.org/10.9728/dcs.2022.23.8.1373>

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 11 July 2022; Revised 16 August 2022

Accepted 19 August 2022

***Corresponding Author; Yongsoon Choi**

Tel: +82-2-3274-4893

E-mail: yongsoon@sogang.ac.kr

I. 서론

이철희 (2018)에 따르면, 사람의 가청 주파수는 20Hz~20,000Hz이다 [1]. 이 가청 주파수는 태어난 이후 지속해서 노화에 의해 감소한다. 보통 20대는 17,000Hz, 30대는 15,000Hz까지의 소리를 들을 수 있다. 더욱이 전 연령대에 스마트폰이 널리 보급됨에 따라 미디어를 접하는 빈도가 증가하였고, 이어폰과 헤드셋 등 다양한 요인들로 인해 현대인의 가청주파수 영역은 더욱 감소하고 있다. 이에 아직 청력을 감소시키는 요인에 비교적 많이 노출되지 않은 0~9세, 십대들이 들을 수 있는 소리와 나머지 연령층이 들을 수 있는 소리의 간극이 심화한다. 영유아와 어린이들이 들을 수 있는 고주파 소리를 그들의 부모나 어른들은 듣지 못하는 것이다.

여기서 문제가 발생한다. 고주파 소리는 청각에 악영향을 주고, 같은 데시벨이라 하더라도 저주파 소리보다 더욱 시끄럽게 인식된다 [1]. 따라서 고주파 소음은 아이들에게 난청과 스트레스를 유발한다. 그러나 그들의 부모는 해당 고주파 소음을 들을 수 없으므로 그러한 문제를 인지조차 하지 못할 수 있다. 또한, 영유아뿐 아니라 최근 증가하는 반려동물(강아지, 고양이)에게도 고주파 소음은 큰 스트레스 요인이 될 수 있다.

따라서 고음과 소음의 위치를 대략적으로 찾아 소음이 발생하는 근원지를 파악해야 할 필요가 있다. Fluke ii900 (2019)와 같이 소리를 시각화해 보여주는 음향 카메라가 존재하지만, 공장 등에서 사용하는 산업용 카메라가 대부분이고 가격 역시 수천만 원을 호가하기 때문에 일상에서 쓰기 부적절하다 [2]. 따라서 낮은 비용으로 소리의 방향성을 찾는 방법이 필요하다. 최근 가장 대중적이고 비용이 적은 방법은 바로 스마트폰 애플리케이션을 이용하는 것이다. 따라서 스마트폰 단위에서 소리를 어떻게 입력받아 소리가 나는 방향을 어떻게 보여줄지에 대한 연구와 논의가 필요하다.

이와 같은 필요에 의해 고주파 소리의 방향을 찾기 위한 소리 탐지 애플리케이션인 FUS (Find Undetectable Sound)를 제작하였다. 해당 애플리케이션에서 소리의 입력은 스마트폰에 이미 존재하는 녹음 기능을 활용한다. FUS는 녹음 시 저장되는 mp4 파일을 전처리하여 데이터로 변환하고, 전처리한 수치를 기반으로 소리가 나는 방향이 어디인지 파악하여 그래프로 나타내는 애플리케이션이다. 3차원 공간을 스마트폰을 기준으로 왼쪽, 오른쪽 2개의 구간으로 세분화하여 소리의 근원지를 예측한다. 이를 위해, 소리 데이터 처리 알고리즘과 소리의 방향성을 나타내는 디자인을 제작했고 사용자 테스트를 거쳐 소리를 신뢰성 있고 친숙하게 보여주는 디자인이 가장 효과적임을 확인했다.

II. 관련 연구

먼저, 소리의 위치를 인공지능을 활용해 찾는 연구가 존재한다. Karren Yang (2020)에 따르면, 유튜브 자음 감각 퀘

감 반응 (ASMR; Autonomous Sensory Meridian Response) 영상에 등장하는 인물의 위치와 그때의 소리를 학습시켜 소리의 근원지가 왼쪽인지 오른쪽인지를 판단하는 기계학습 알고리즘 연구가 있다. ASMR 영상 촬영에 주로 사용되는 마이크가 스테레오 타입이기 때문에 왼쪽과 오른쪽에서 들어오는 소리를 구분할 수 있는 특징을 활용하였다. 등장 인물이 영상에서 왼쪽 혹은 오른쪽에 위치함에 따라 달라지는 이미지와 소리를 동시에 멀티 모달 기법으로 분석해, 소리가 발생하는 곳이 왼쪽인지 오른쪽인지를 학습한다 [3].

인공지능 기술을 활용해 소리의 위치와 크기를 이미지로 변환하여 소리의 위치를 찾는 연구도 있다. 한국표준과학연구원 (2021)과 포항공과대학교가 공동으로 협력하여 진행한 이 연구는 56개의 스피커를 구 형태로 설치한 실험실에서 소리 데이터를 기반으로 딥러닝 기술을 통해 소리의 위치와 크기를 이미지로 변환하여 음원 지도를 만들고 소리의 위치를 추적한다 [4]. 주로 조난자 위치를 찾는 정찰, 수송, 구조 등의 분야에 상용화될 것으로 전망된다.

소리의 방향성을 시각화하여 위치를 효과적으로 나타내는 연구가 있다. Dhruv Jain (2015)에 따르면, HMD (Head-Mounted Display) 상에서 소리가 나는 방향을 시각화한 연구이다. 소리의 모습을 파동, 화살표 등의 기호로 나타내고, 소리의 크기는 화살표의 크기를 변화시키는 등의 방법으로 표현하였다. 사용자 실험 결과, 소리를 파동 형태로 표현하는 것이 화살표 형태의 표현보다 더 선호도가 높았다. HMD 화면으로 소리를 시각화하는 여러 방법에 대해 사용자의 다양한 선호도를 확인한 연구이다 [5].

소리의 방향을 찾고 인공지능으로 감지하는 애플리케이션이 존재한다. Mosquitone detector (2017)라는 애플리케이션은 현재 ios에서만 구동되며, 고주파 소음을 감지하고 그 위치를 알려주는 애플리케이션이다. 그러나 고주파 소음이 먼저 감지되지 않으면 위치를 찾는 기능을 사용할 수 없기 때문에 범용성이 적다. 또한, 고주파 소음이 감지되어 위치를 찾을 때 사용자가 시계방향으로 직접 움직여야 하는데, 4rpm의 속력에 맞추어 15초 정도 소리를 녹음해야 하는 등 여러 제약 조건이 있어 번거롭고 사용성이 좋지 않다. 2022년 6월 기준 평점은 2.5점으로 평가가 좋지 않다 [6].

구글에서 개발한 실시간 소리 번역 애플리케이션인 Live Transcribe & Notification (2019)은 주변 소리와 소음을 분석해 사용자에게 현재 어떤 소리가 발생하는지 알려준다. 예를 들어 문 두드리는 소리가 나면 ‘문 두드림’, 개가 짖고 있다면 ‘개 짖는 소리’라는 신호를 사용자에게 문자로 전송한다. 알림 기능은 문자뿐만 아니라 스마트폰의 진동이나 플래시 불빛을 통해서도 확인할 수 있다. 2020년 기준 감지할 수 있는 소리의 종류는 연기 및 화재 경보, 사이렌, 고함, 아기 소리, 초인종 소리, 노크, 개 짖는 소리, 가전제품 소리, 물 흐르는 소리, 전화벨 소리로 모두 10가지다. 현재로부터 12시간 전까지의 데이터가 저장되며, 사용자가 기록을 직접 확인할 수 있다 [7].

III. FUS(Find Undetectable Sound) 애플리케이션

3-1 애플리케이션 요약

FUS(Find Undetectable Sound)는 감지할 수 없는 소리를 찾는다는 뜻으로, 인간의 청력으로 감지하기 어려운 소리를 스마트폰 단위에서 감지하여 그 방향성을 알려주는 스마트폰 애플리케이션이다. 소리 데이터를 수집하는 방법은 스마트폰에 존재하는 녹음 기능을 사용한다. 먼저 사용자는 임의로 공간을 두 개로 분할하여(예시: 왼쪽과 오른쪽, 위쪽과 아래쪽) 두 번의 녹음을 진행한다. 그리고 제작된 애플리케이션에 두 개의 소리 파일을 업로드한다. 그러면 애플리케이션과 연결된 서버에 업로드되어있는 Librosa를 활용한 소리 전처리 알고리즘을 거쳐 소리의 크기가 수치로 반환된다. 반환된 두 소리 크기 데이터를 비교해 어느 쪽에서 소리가 더 크게 입력됐는지를 판단한다. 판단 결과에 따라, 서버를 통해 업로드된 그래프 시각화 알고리즘을 실행한 후 그래프를 반환한다. 반환된 그래프는 애플리케이션 상에서 사용자에게 소리의 방향을 알려준다.

그림 1은 애플리케이션의 작동 방식을 나타낸 시스템 다이어그램이다. 그림 1의 a와 b는 소리가 오른쪽에서 발생하지만, 사용자는 그것을 모르는 상황이며 사용자는 스마트폰을 왼쪽과 오른쪽으로 향하게 하여 소리를 두 번 녹음한다. 그림 1의 c에서 사용자는 녹음한 두 소리를 FUS 애플리케이션에 입력한다. 그림 1의 d에서 사용자는 애플리케이션이 예측한 소리의 발생 방향을 그래프로 확인하고 소리의 근원지를 찾는다.

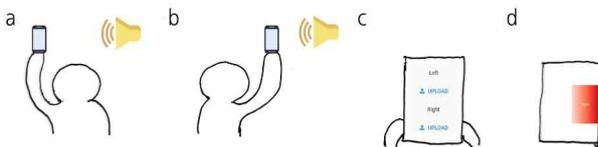


그림 1. FUS 애플리케이션의 시스템 사용 시나리오
 Fig. 1. Usage scenario of FUS Application

3-2 Librosa를 활용하여 소리의 크기로 소리의 방향을 예측하는 알고리즘

Librosa (2015)는 python에서 많이 사용하는 음악, 음성 파일 분석 프로그램이다 [8]. 두 개의 파라미터(y, sampling rate)로 음성 파일을 불러오는데, y는 시간에 따른 소리의 진폭이며 sampling rate는 1초당 분석하는 소리 샘플의 개수이다. 이때 음성 파일의 길이는 음파의 길이/sampling rate이며, 음파의 길이는 y 값들의 개수이다. Librosa에 입력된 음성 파일을 그래프로 나타내면, 그림 2와 같이 x축이 시간이고 y축이 소리의 진폭인 그래프가 나타난다. 이 그래프는 y=0을 기준으로 위아래가 대칭인 2차원 그래프이다.

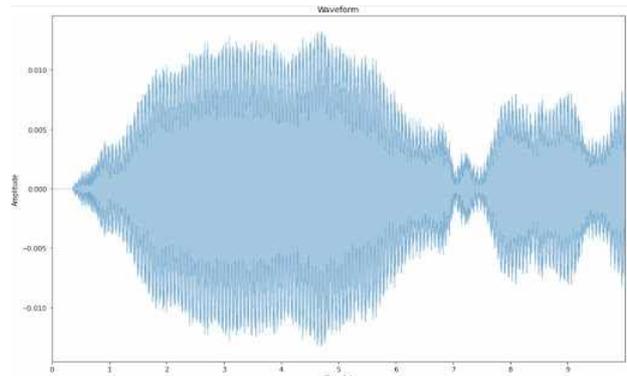


그림 2. 소리 데이터를 Librosa에 넣어 전처리한 그래프. (x축이 시간(초), y축이 소리의 진폭(db))
 Fig. 2. The Graph after Pre-processing a Sound Data. (x axis means time(s), and y axis means decibel of the sound(db))

소리 데이터를 수집할 때는 사용자가 공간을 원하는 만큼 분할하여 수집할 수 있다. 그러나 가장 대표적이고 직관적으로 공간을 분할할 수 있도록 공간을 왼쪽과 오른쪽으로 분할하는 방법을 제시한다. 스마트폰의 마이크를 왼쪽으로 하여 녹음한 음성 파일과 오른쪽으로 하여 녹음한 음성 파일 각각은 스마트폰에 저장된다. 저장된 각 파일을 알고리즘에 입력하면 Librosa를 이용해 음성 파일을 전처리한다. 소리 파일의 길이, 곧 사용자가 녹음한 시간은 관계없고, 오로지 음성 파일에 녹음된 소리의 크기가 중요하다. 다만, 사용자의 소음이 함께 녹음될 경우 부정확한 결과가 나올 수 있으므로 최대한 조용한 환경에서 진행되어야 한다. 특히 녹음을 시작과 끝 지점에서 사용자가 움직여야 하므로 소음이 부득이하게 삽입될 수 있다. 따라서 Librosa를 통해 전처리한 소리 데이터의 양 끝 1초씩을 잘라내어 사용한다. 그림 3은 양 끝 1초씩이 잘린 소리 데이터 그래프이다.

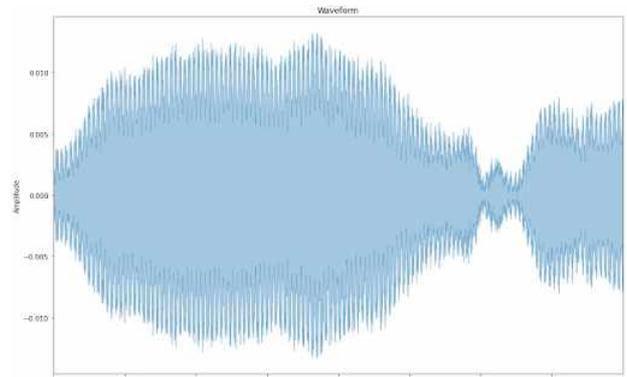


그림 3. 그림 2의 처음과 끝 1초를 잘라낸 소리 데이터 그래프. (x축이 시간(초), y축이 소리의 진폭(db))
 Fig. 3. The Graph after cutting one second at first and in the last part of the sound data in Figure 2. (x axis means time(s), and y axis means decibel of the sound(db))

그림 3을 보면 y=0을 기준으로 대칭인데, 소리의 크기를 측정할 때에 음수 값은 의미가 없으므로 위 데이터에서 y값에 절댓값을 씌워 양수만 존재하도록 한다. 그림 4는 그림 3에 절댓값을 씌운 그래프이다.

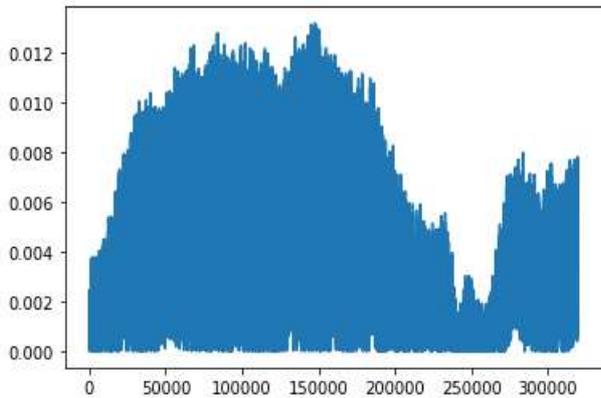


그림 4. 그림 3에 절댓값을 씌운 그래프. (x축이 시간(초), y축이 소리의 진폭(db))

Fig. 4. The Graph which is the absolute values of data in Figure 3. (x axis means time(s), and y axis means decibel of the sound(db))

그 다음, 녹음한 소리의 크기 평균을 수치로 구한다. 소리의 근원지가 마이크에 가까울수록 큰 소리가 입력될 것이라는 가정을 전제로 한다. 그렇다면, 소리의 크기에 해당하는 수치들을 서로 비교하여 더 큰 값을 갖는 방향이 소리의 근원지가 있는 방향이라고 판단할 수 있다. 이때 평균 말고도 중앙값, 최빈값 등 다양한 통계 수치를 활용할 수 있지만, Librosa는 소리 파일로부터 1초에 약 수만 번의 샘플을 추출하므로 그 데이터 중 중앙에 있는 한 값은 해당 소리 데이터를 대표하지 못한다. 또한, 최빈값의 경우도 같은 이유로 대표성을 잃는다. 수십, 수백만 개의 진폭 중 최빈값의 의미가 평균보다 적기 때문이다. 실제로, 평균 대신 최빈값을 추출해 프로토타입을 대상으로 실험한 결과 소리 근원지의 방향에 따른 차이가 평균에 비해 뚜렷하지 않았다. 따라서 평균을 사용하는 것이 가장 적합하다고 판단하였다.

그렇게 구한 두 소리 파일의 크기에 대한 평균을 비교해, 소리 크기의 평균값이 더 큰 방향에서 소리가 발생한 것으로 판단할 수 있다. 이때 고려하는 방향이 두 개이므로, 두 가지 경우의 수에 대해 서로 다른 그래프로 소리의 방향을 나타낸다.

3-3 FUS 애플리케이션 프로토타입

애플리케이션 개발은 anvil (2016)이라는, python으로 프론트엔드와 백엔드, 서버와 데이터베이스 연결을 동시에 구성할 수 있는 플랫폼을 사용한다 [9]. anvil을 통해 사용자는 각종 항목을 마우스로 조정하는 직관적인 방법으로 직접 User Interface를 꾸며 웹 애플리케이션을 제작할 수 있다.

또한, 이를 anvil 서버와 연동시켜 많은 클라이언트 측에서 자유롭게 애플리케이션을 활용할 수 있다. 서버뿐 아니라 anvil 데이터베이스를 활용해 데이터를 저장할 수 있고, 다른 애플리케이션과 연동도 가능하다. 무엇보다 Html, JavaScript 등 다른 언어나 프레임워크를 배우지 않더라도 python으로 웹 애플리케이션 제작이 가능한 것이 가장 큰 특징이다.

FUS 웹 애플리케이션은 anvil을 활용하여 제작되었다. FUS는 왼쪽과 오른쪽, 즉 각각 다른 방향으로 스마트폰을 향하여 녹음한 두 소리를 비교해 소리가 어느 쪽에서 발생하는지 그래프로 나타내준다. 작동 방법은 다음과 같다. FUS에서 업로드된 두 소리 파일이 웹 애플리케이션과 연결된 서버를 통해 전처리되고, 그 크기가 수치화되어 서버를 통해 웹 애플리케이션으로 넘어와 FUS 내의 데이터베이스에 저장된다. 저장된 두 수치를 비교한 결과에 따라 서로 다른 그래프를 반환하는 알고리즘을 호출한다. 해당 알고리즘은 결과 그래프를 이미지로 반환하며, 반환된 이미지는 다시 서버를 거쳐 웹 애플리케이션에 표시된다. 그리고 데이터베이스는 매번 첫 번째 소리 파일을 업로드할 때 초기화된다.

1) 인터랙션 디자인

인터페이스에는 스마트폰으로부터 음성 파일을 받는 두 개의 버튼(file uploader)과 그 두 버튼이 각각 왼쪽, 오른쪽 중 어떤 방향의 소리를 입력받는 것인지를 알려주는 표시, 그리고 알고리즘을 거쳐 소리가 어느 방향에서 발생하는지를 보여주는 결과 그래프가 나타날 공간이 존재한다.

두 개의 버튼을 통해 입력된 파일 각각은 anvil 서버에 전송되어 Librosa를 통해 전처리 과정을 거친다. 그다음, 각 소리 크기의 절댓값의 평균을 수치로 변환해 다시 anvil로 전송한다. 그렇게 전송된 두 소리의 크기 값을 anvil과 연결된 데이터베이스에 저장한다. 그리고 anvil 코드 내에서 그 데이터베이스의 수치들을 서로 비교해 왼쪽 소리 크기가 더 크면 왼쪽에서, 오른쪽 소리 크기가 더 크면 오른쪽에서 소리가 발생한다고 판단한다.

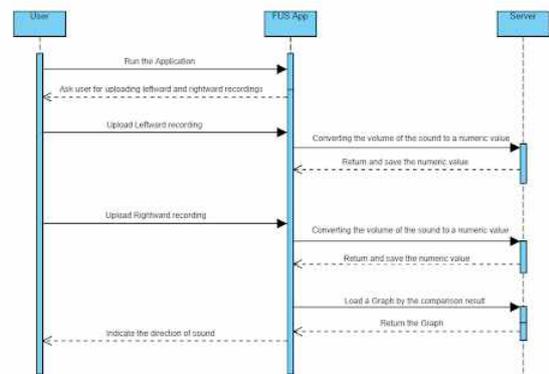


그림 5. FUS 애플리케이션의 사용자 인터랙션에 대한 시퀀스 다이어그램

Fig. 5. Sequence Diagram about user interaction of FUS Application

anvil 서버를 통해 그 결과에 따라 서로 다른 그래프를 보여주는 알고리즘이 동작한다. 알고리즘 결과로 그래프 이미지가 서버를 통해 anvil 웹 애플리케이션으로 반환된다. 반환된 그래프는 이미지 공간에 나타난다. 그림 5는 FUS 애플리케이션의 사용자 인터랙션에 대한 시퀀스 다이어그램이다.

2) FUS 애플리케이션의 스크린 Flow

그림 6의 a는 FUS를 실행한 뒤의 초기 화면이다. 초기 화면은 왼쪽 소리를 입력하는 file uploader 버튼과 오른쪽 소리를 입력하는 file uploader 버튼으로 나누어진다. 각 버튼은 어떤 방향의 소리를 입력해야 하는지 알려줄 수 있도록 Left, Right이라는 이름이 붙어있다. 주의해야 할 점은 왼쪽 소리부터 순서대로 입력해야 한다. 왜냐하면, 왼쪽 소리가 입력됐을 때 사용 중인 데이터베이스가 초기화되어 소리 데이터가 올바르게 입력되도록 알고리즘을 제작했기 때문이다. 오른쪽 소리 데이터부터 입력할 시에는 왼쪽 소리를 입력했을 때 이전에 삽입한 오른쪽 소리 데이터가 초기화되는 문제가 발생한다. 따라서 왼쪽부터 파일을 순서대로 업로드해야 한다. 소리를 업로드하기 위해 upload 버튼을 클릭할 때 파일 업로드 과정은 다음과 같다.

삼성 갤럭시 S21 기종의 경우, ‘작업 선택’ 화면에서 ‘파일’ 선택 - ‘오디오’ 선택 - 특정 방향에서 녹음한 해당 소리 데이터를 선택하면 소리 파일이 업로드된다.

애플리케이션을 실행할 때나, 혹은 file uploader에 소리 파일을 입력시킨 뒤 일정 시간(약 2초) 동안 FUS가 실행되거나 파일이 업로드될 때 로딩 중임을 알려주는 로딩 화면도 존재한다.

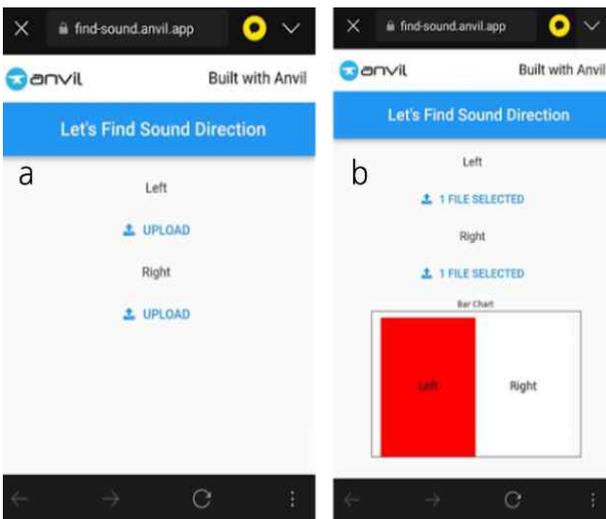


그림 6. a는 FUS 애플리케이션 실행 시 초기 화면. b는 스마트폰 기준 왼쪽에서 소리가 발생할 때 FUS 애플리케이션 실행 결과 화면.

Fig. 6. a is the initial screen of FUS Application. b is the result screen when the sound occurs left side of the phone.

그림 6의 b는 스마트폰을 기준으로 왼쪽에서 소리를 발생시켰을 때의 FUS 애플리케이션 실행 결과 출력 화면이다. 해당 결과 출력 화면은 왼쪽으로 마이크를 대고 녹음한 소리의 크기가 오른쪽으로 마이크를 대고 녹음한 소리의 크기보다 크다는 것의 의미이다. 즉, 왼쪽에서 소리가 발생하고 있을 것이라는 결과를 나타낸다.

IV. 사용자 실험

4-1 실험 구성

5명(50대 남자 1명, 20대 남자 2명, 20대 여자 2명, 평균 나이 30세)을 상대로 한 사전 실험을 통해 정해진 최종 프로토타입을 본 실험에 결과를 보여주는 그래프 디자인으로 사용했다.

본 실험에 참여한 사용자는 22명(20대 남성 16명, 30대 남성 1명, 20대 여성 4명, 30대 여성 1명, 평균나이 24.5세)이었다. 실험 진행에 있어 사전 촬영에 대한 동의, 개인정보 수집에 대한 동의를 구했다. 실험 전에 실험에 대한 목적 및 방법, 고주파 소음에 대한 유의 사항을 미리 안내하였다.

사용자는 귀마개를 쓰고, 공간 안에 있는 2개의 스피커 중 하나에서 나오는 소리를 듣지 못한 채, 애플리케이션을 통해 소리의 방향을 찾았다. 이것을 3회 반복하고 애플리케이션의 결과를 참고해 사용자가 생각하는 소리가 나는 방향을 설문지에 입력하였다.

그리고 사용성 검증을 위해 피터 모빌 (2004)의 허니콤 모델을 기초로 한 12개의 리커드 5점 척도 문항을 구글 폼을 이용해 설문 조사하였고, 개별 사용자 의견 및 피드백 수집은 주관식 문항을 통해 입력받았다. 설문조사의 시간제한은 없었다 [10]. 실험 참여자들은 실험 중 씹크어라우드(Think Aloud)가 가능하여 실험 중 드는 자신의 생각이나 느낌에 대해 자유롭게 말하도록 요청하였다.

4-2 실험 계획

1) 사전 실험을 거친 최종 프로토타입: Bar Chart

시각화 방법으로 bar chart를 선택한 이유는 Ian Spence (1991)의 논문에 근거한다 [11]. 논문에 의하면, 가장 대표적인 시각화 그래프는 pie chart와 bar chart인데, bar chart가 pie chart보다 대중적으로 더 시각화에 효과적이라고 알려져 있다. 사전 실험을 위한 bar chart는 ‘bar chart’라는 제목을 갖고 배경색이 하얀색인 left, right 두 개의 x값을 갖는 바 그래프의 형태를 띤다. 왼쪽에서 소리가 발생한다고 판단할 경우에는 왼쪽 바가 빨간색으로 채워지고, 오른쪽 바는 하얀색으로 배경색과 동일하여 바의 형태를 찾아볼 수 없다. 반대로 오른쪽에서 소리가 발생한다고 판단할 경우에는 오른쪽 바가 빨간색으로 채워지며, 왼쪽 바가 하얀색으로 채워져 배경과 구분할 수 없게 된다. 소리가 나지 않는다고 판단되는 방향의

바 그래프를 보이지 않도록 하며 방향성을 부여하는 것이다.

사전 실험은 50대 남자 1명, 20대 남자 2명, 20대 여자 2명 총 5명에 대해 진행되었다. 우선 이 애플리케이션의 목적과 실험 방법에 대해 오리엔테이션을 진행하였다. 실험 방법은 다음과 같다. 실험 참가자들은 청력 안전을 위해 모두 귀마개를 사용한다. 16,000Hz의 고주파 소음을 왼쪽 방향 a, 오른쪽 방향 b 중 하나에서 발생시킨다. 참가자들은 16,000Hz의 고주파 소음이 어디서 발생하는지 모르는 상태로 a, b 두 방향에 대해 소리를 녹음하고, 녹음한 소리를 anvil 웹 애플리케이션에 입력시켜 결과를 확인한다. 소리를 녹음할 때는 사용자 기준으로 왼쪽과 오른쪽을 판단하여 핸드폰의 마이크를 왼쪽과 오른쪽으로 향하게 한 뒤 녹음을 진행한다. 확인한 결과를 토대로 현재 고주파 소음이 어디서 발생하고 있는지 결정하고 사용성 평가를 진행한다. 사용성 평가는 인터뷰로 진행되며, 인터뷰 질문은 다음과 같다.

1. 해당 그래프를 보고 소리의 방향성을 파악할 수 있었나?
2. 이 서비스는 유용한가?
3. 이 서비스를 사용하는 과정(소리 녹음, 녹음한 소리 파일을 웹 애플리케이션에 전달, 전달 후 결과를 받기까지의 전 과정)은 어떠한가? 번거롭거나 지루하지 않았나?
4. 디자인에 대해 보완할 점은 무엇인가?

인터뷰 결과,

“소리 방향성을 파악할 수 있었다, 서비스의 목적이 좋다, 아이를 둔 부모들에게 유용할 것 같다, 녹음과 애플리케이션을 사용하는 전반적인 과정은 단순하고 쉬웠다, 왼쪽과 오른쪽을 더 분명히 나타내고 방향이 더 직관적으로 나타났으면 좋겠다.” (50대 남성)

“방향성을 파악할 순 있지만 방향이 더 확실하게 표시되었으면 좋겠다, 화살표나 소리를 나타내는 기호 등을 사용했으면 좋겠다.” (20대 남성)

“왼쪽 혹은 오른쪽 그래프가 나올 때 그래프가 나타내는 방향으로 점점 색이 짙어지도록 그라데이션이 있으면 좋을 듯 하다. 혹은 해당 방향으로 점점 그래프가 올라가는 식으로 디자인해도 좋을 것 같다.” (20대 여성)

이것을 통해 일부 사전 실험 참여자들은 그래프의 방향성에 대해 좀 더 명확하고 직관적인 디자인이 필요하고, 그래프의 높낮이를 변화시키거나 그라데이션을 추가하여 디자인을 개선할 수 있다는 아이디어를 이야기했다.

“색이 구분되므로 방향성을 파악할 수 있다. 좌측과 우측 각각 소리가 어느 정도인지 수치상으로 나타내면 훨씬 좋겠다. 현재 디자인도 괜찮다, 글자에 배경으로 색이 입혀진 게 눈에 잘 들어온다, 그라데이션 있으면 좋을 것 같다, Bar가 Half Pie 차트보다 낫다.” (20대 남성B)

“소리 방향성을 매우 파악하기 좋고, 서비스가 매우 유용하고, 그래프 디자인도 문제없고, 하나 추가하자면 애플리케이션의 UI에서 Left 파일 업로드 부분은 왼쪽에, Right 파일 업로드 부분은 오른쪽에 위치시켜도 좋을 듯 하다. 그라데이션을 넣어도 좋다. Bar와 비교해 Half Pie 차트(원형)가 크게

다르지는 않으나 디자인적으로 좀 더 예쁘다.” (20대 여성 B)

이를 통해 일부 사전실험 참여자들은 현재 디자인으로도 방향성을 파악할 수 있고, Bar 그래프가 Half Pie 그래프 등 다른 그래프 디자인과 비교해 비슷하거나 더 낫다는 의견을 보였다.

인터뷰 내용을 종합적으로 판단하였을 때, 배경색을 활용해 방향성을 줄 수 있는 그라데이션을 활용하고, 불필요한 테두리를 삭제하기로 했다. 또한, 방향을 더욱 직관적으로 나타내기 위해 글자를 하얀색으로 설정하여 소리가 발생하지 않는다고 판단되는 방향의 경우 글자가 보이지 않도록 설정하였다. 따라서 그림 7과 같은 Bar 차트 형태의 그래프를 최종 디자인으로 선정했다.

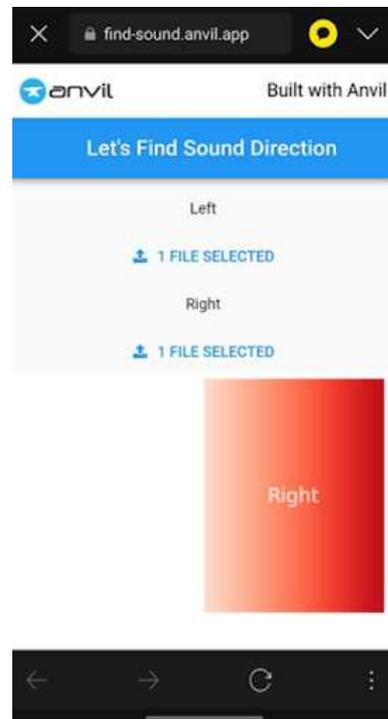


그림 7. 소리가 오른쪽에서 발생할 때 결과 Bar Chart 그래프
Fig. 7. The Result Graph of the FUS Application with Bar Chart When the Sound Occurs at the Right Side of the Phone.

2) 실험 참여자

실험 참여자는 모두 서강대학교 학부와 대학원에 재학 중인 학생들이다. 실험 참여 인원은 22명(남자 17명, 여자 5명)이 모집되었다. 모든 참가자는 스마트폰과 스마트폰 애플리케이션 조작 경험이 있다.

인구 통계학적으로 20대가 전체의 91%, 30대가 전체의 9%를 차지했다. 구체적으로 20대 남성 16명, 30대 남성 1명, 20대 여성 4명, 30대 여성 1명으로 구성된다. 20대 남성은 전체의 약 73%, 30대 남성은 전체의 약 5%, 20대 여성은 전체의 약 18%, 30대 여성은 전체의 약 5%이다. 모든 실험 참가자는 3M사의 귀마개(모델명: 3M 1100)를 착용해 고주파 소리를 듣지 못하도록 하였으며 실험 참가자마다 실험 시작 전 소리가

들리지 않는 것을 확인하였다. 사용 기기는 애플의 아이폰 7과 삼성의 갤럭시S21이나, 아이폰을 처음 조작하는 참가자가 있어 한 명의 사용자를 제외하고는 갤럭시 S21로 진행하였다.

3) 설문지

사용성 테스트를 위한 설문지는 리커트 척도 5점으로 구성되며 피터 모빌 (2004)의 HoneyComb 모델을 참고하였다 [10]. 12개의 질문은 HoneyComb 모델에 따라 총 4개 분야로 분류된다. Useful(유용성)에 대한 질문, Usable(사용성)에 대한 질문, Findable(소리를 찾는 것)에 대한 질문, Credible(신뢰성)에 대한 질문이다. 이 설문조사의 목적은 웹 애플리케이션에 대한 만족도 및 사용성을 조사하기 위함이다.

표 1. 사용성 테스트를 위한 설문지

Table 1. The Questionnaire for Serviceability Test

Question Types	Questions
Questions about 'Useful'	1. Is this service useful?
	2. Does this service coincide with the purpose of finding a sound?
	3. Do you want to find a sound with this service?
Questions about 'Usable'	4. Was this service easy to use?
	5. Was the design of this service familiar to use?
	6. Was this service simple to use?
Questions about 'Findable'	7. Were you able to find the location of the sound through this application?
	8. Did you have any difficulty uploading the file?
	9. Were there any difficulties in recording the sound and using the application to find the direction of the sound?
Questions about 'Credible'	10. Is the directional information of the sound that this service tells you reliable?
	11. Is this service reliable in the design of sound direction?
	12. Can you trust and use this service?

4) 실험 진행

2022년 4월 6일 오후 3시부터 4월 7일 오후 7시까지 서강대학교 하비에르관 신영균 스튜디오에서 실험을 진행했다. 실험 진행에 앞서 실험 참여자들에게 실험 장면에 대한 사진 촬영을 허락받았고, 개인정보 수집(성별, 나이대)에 대한 동의를 읽게 한 후 동의를 구했다. 개인정보는 1개월 이내 폐지됨을 명시하였고, 실험 이후 진행되는 설문조사에 대한 동

의를 구했다. 또한, 실험에 대한 목적 및 방법과 16,000Hz의 고주파 소리로 인한 청력 손상을 방지하기 위해 귀마개를 착용해야 한다는 유의 사항을 안내했다.

실험 참가자는 2개의 스피커 중 고주파 소리(16,000Hz)가 나오고 있는 하나의 스피커를 찾는 과제를 3번 수행한다. 그림 8과 같이 스피커는 사용자를 기준으로 왼쪽에 1개, 오른쪽에 1개를 각각 같은 거리만큼(약 1.8m) 떨어뜨려 배치되어 있으며, 스피커끼리는 약 1m가 떨어져 있다. 사용자는 의자에 앉아서 실험을 진행한다. 또한, 사용자는 귀마개를 착용해 고주파 소음을 들을 수 없는 상태이며 실험 진행 전에 소리가 들리는지에 대한 확인 절차를 거쳤다. 다음 실험 참가 예정자들은 스튜디오 뒤편에 앉아서 대기하며 실험 진행자는 실험 참가자 우측에 앉아서 소리를 재생시키는 환경이다.

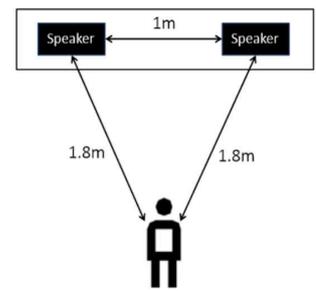


그림 8. 실제 실험환경 사진(왼쪽)과 상공에서 보았을 때의 실험환경 도식화(오른쪽).

Fig. 8. The picture of the actual experimental environment(Left) and Schematic illustration of the experimental environment from the sky(Right).

실험과정은 크게 녹음과 애플리케이션 실행으로 구분된다. 녹음 과정은 총 6번, 즉 2번씩 3 세트로 진행되는데, 3번 수행되는 과제 각각에 있어 왼쪽과 오른쪽에 핸드폰을 향하게 하여 녹음해야 하기 때문이다.

그리고 각 과제에서 FUS 웹 애플리케이션에 접속하고, 각 세트에 포함되어있는 녹음 파일 2개를 업로드하여 현재 고주파 소음이 사용자를 기준으로 왼쪽에서 발생하는지 혹은 오른쪽에서 발생하는지에 대한 결과를 세트별로 확인한다.

그리고 3번의 과제 수행에 대해 정답을 맞힌 횟수를 측정한다. 실험이 끝난 후에는 구글 폼을 통해 5점 리커트 척도를 이용한 설문조사를 실시하고, 주관식 설문조사 문항으로 인터뷰를 대체한다.

실험 진행 동안 실험 참여자들이 씽크어라운드(Think Aloud)를 허용하여 실험 중 드는 생각이나 인상에 대해 자유롭게 말하도록 요청하였다.

4-3 실험 결과 분석

총 22명이 3번의 과제(실험)를 수행하여 구한 66번의 애플리케이션 사용 결과를 토대로, 총 정답률은 약 72.7%이며, 총 오답률은 약 27.3%이다.

1번째 실험에서는 22명 중 17명(77.3%)이, 2번째 실험에서는 22명 중 15명(68.2%)이, 3번째 실험에서는 22명 중 16명(72.7%)이 정답을 맞히었다. 오답률은 각각 22.7%, 31.8%, 27.7%이다. 22명 중 모든 정답을 맞힌 사람은 8명이었고, 1개만 틀린 사람은 10명, 2개를 틀린 사람은 4명이었고 모두 맞추지 못한 사람은 없었다.

5점 리커트 척도 설문조사에 대한 정량적 분석을 진행하였다. 1점은 '전혀 그렇지 않다'이고, 5점은 '매우 그렇다'이다. 모든 설문조사 문항은 FUS 애플리케이션에 대한 긍정성을 묻는 질문으로 구성하였다.

먼저, 12개 문항과 전체 문항의 평균과 표준편차는 다음과 같다.

표 2. 사용성 테스트를 위한 설문지
Table 2. The Questionnaire for Serviceability Test

Question Number	Mean	Standard Deviation
#1	3.818	0.795
#2	4.364	0.658
#3	3.636	1.049
#4	3.500	1.144
#5	3.500	0.913
#6	3.682	1.211
#7	4.364	0.790
#8	3.773	1.020
#9	3.500	1.263
#10	4.046	0.785
#11	3.818	0.795
#12	3.955	0.899
Total Questions	3.830	0.985

표2에 따르면, 열두 문항의 전체 평균은 3.830, 전체 표준편차는 0.985였다. 5점으로 갈수록 긍정적인 반응이므로, 전체적으로 긍정적인 결과를 얻었다. 특히 2번, 7번, 10번, 12번 문항에서 전체 문항의 평균값보다 높은 평균값을 보였다. 또한, 2번, 7번, 10번, 12번 문항에서 전체 문항의 표준편차보다 낮은 표준편차 값을 보였으므로 위 네 문항이 묻는 FUS의 특징에 대해 사용자가 특히 더 긍정적인 신호를 보내고 있다는 것을 확인하였다.

반대로, 1번, 5번, 11번 문항에서 전체 문항의 표준편차보다 낮은 표준편차 값을 가지면서 동시에 전체 문항의 평균값보다 낮은 평균값을 보였다. 이 세 문항이 묻는 FUS의 특징에 대해 사용자가 부정적인 반응을 보인다는 것이다.

V. 논의 및 결론

긍정적인 반응을 보인 설문조사 문항을 살펴보면, 2번 문항은 해당 애플리케이션이 소리를 찾는 목적에 부합하는지를 물어보았고, 7번 문항은 애플리케이션을 통해 소리가 발생하는 위치를 찾을 수 있었는지를 묻는 질문이었다. 10번 문항은 FUS가 알려주는 소리의 방향이 믿을 만한지에 대한 질문이었고, 12번 문항은 FUS를 믿고 사용할 수 있느냐는 질문이었다. 요약하자면, 사용성 테스트를 통해 애플리케이션이 목적에 부합하고, 사용자들이 애플리케이션을 활용해 소리를 찾을 수 있었으며, 그 결과는 신뢰성이 높다는 결론을 얻을 수 있었다.

설문조사의 마지막 문항이었던 주관식 문항의 응답을 살펴보면, 소리를 찾는 애플리케이션이 신기하다는 반응들이 존재했다. 이를 통해, 소리를 찾는 기능과 목적에 대해 사용자가 갖는 긍정적인 반응을 뒷받침할 수 있다.

신뢰성 있고 친숙한 디자인 - 사용자가 부정적인 반응을 보였던 1번과 7번, 11번 문항은 각각 서비스가 유용한지, 디자인이 사용하기에 친숙한지, 소리를 나타내는 디자인에 신뢰가 가는지 묻는 문항이었다. 실험 당시 서비스가 다소 유용하지 않았으며, 애플리케이션의 디자인이 친숙하지 않고 신뢰가 가지 않는다는 결론을 얻었다.

이에 대한 근거를 주관식 문항을 통해 받은 부정적인 반응으로 확인할 수 있었다. 애플리케이션을 제작할 때, anvil 플랫폼에서 제공하는 서버를 활용하였는데 서버와의 연결이 다소 불안정했다. 실험 중간에 그러한 사실을 파악했지만, 서버 자체를 개선할 방법이 없기 때문에 애플리케이션을 껐다가 켜는 방식으로 서버와의 연결을 재설정하였다. 이와 관련하여 서버에 대한 부정적 응답이 다수 존재하였고 이 불편함이 애플리케이션의 유용성을 하락시킨 주요 요인이 되었다.

이 밖에도 소리의 크기를 그래픽으로 파악할 수 있었지만, 애플리케이션의 디자인이 다소 딱딱한 것 같다는 의견 등 디자인에 대한 긍정과 부정 피드백이 동시에 존재하였다. 애플리케이션으로 소리를 찾은 후 나타나는 결과 그래프뿐만 아니라, 애플리케이션의 시작부터 끝까지 나타나는 모든 디자인을 조금 더 친숙하게 바꾸어야 할 필요성이 있다. 애플리케이션 실행 중에 녹음을 진행하고 녹음 파일을 업로드하기 위해 애플리케이션과 스마트폰 기본 기능을 번갈아 가며 사용하는 과정 중에 발생하는 디자인 측면에서의 이질감이 존재할 수 있다. 이는 디자인을 곡선을 많이 사용하는 등 사용하는 스마트폰의 UI와 유사하게 수정한다면 해결될 것이다.

쉬운 사용 방법 -사용 방법에 대해서도 쉽다는 의견과 복잡하다는 의견이 혼재했다. 이는 사람에 따라 스마트폰 사용 경험 등에 대해 차이가 존재하기 때문이다. 이는 정답률을 낮추는 원인 중에 하나로 작용했을 것이다. 실험이 각자 익숙한 스마트폰 환경에서 진행되지 않았기에 스마트폰 조작의 어려움이 발생하였고, 결국 애플리케이션 사용의 어려움으로 이어졌을 가능성이 크다. 그리고 아직 애플리케이션 자체에서 녹음할 수 있는 기능이 없고 기본으로 제공되는 녹음 기능과 FUS 애플리케이션을 교대하며 사용해야 하므로 애플리케이션 사용이 더욱 복잡하게 여겨졌을 수 있다.

자체 녹음 기능 및 설명서 - 실제로 애플리케이션 자체에 녹음 기능이 있다면 더욱 이용하기 편할 것 같다는 응답도 존재했다. 해당 애플리케이션은 안드로이드와 ios 환경 모두에서 작동되므로 만약 각자가 주로 사용하는 스마트폰에서 FUS를 사용하고 애플리케이션에서 바로 녹음할 수 있는 기능을 추가한다면, FUS의 사용성이 증가할 것이다. 추가로, 사용법에 대한 쉬운 설명서를 함께 제공한다면 애플리케이션 활용에 도움이 될 것이다.

VI. 결 론

일반 성인이 들을 수 없는 고주파 소음을 스마트폰 애플리케이션으로 간편하게 찾아 난청과 스트레스 등을 예방하는 목적의 FUS 애플리케이션은 사용자 실험을 통해 주된 목적을 달성했다는 것을 확인하였다. 향후 애플리케이션 디자인을 신뢰성 있고 친숙한 디자인 방향으로 수정하고, 자체 녹음 기능과 쉬운 설명서를 추가하는 등의 사용자를 고려한 기능들이 추가된다면 더욱 정확하고 간편하게 고주파 소음을 찾을 수 있을 것이다. 또한, 소리를 찾는 작업을 반복하면 다양한 위치에서 동시에 발생하는 소음을 찾을 수 있다. 추가로, 소리를 찾는 방향을 왼쪽과 오른쪽뿐만 아니라 여러 갈래로 분할한다면 더욱 다양한 상황에서 소음의 발생지를 파악할 수 있을 것이다.

향후 FUS 애플리케이션의 추가 연구를 통해 일상생활 속에서 어른들은 들을 수 없는 고주파 소음으로 아이들이 난청 및 스트레스 환경에 노출되는 문제를 쉽게 발견하고 이를 개선하는 데 도움이 되기를 기대한다.

참고문헌

- [1] C. H. Lee, S. K. Hong, "A study on the hearing characteristic based equalizer design for the elderly," *The Journal of Digital Contents Society*, Vol. 19, No. 4, pp. 779-787, April 2018. <https://doi.org/10.9728/dcs.2018.19.4.779>
- [2] Fluke ii900 Industrial Acoustic Camera (Ultrasonic Camera)

[Internet]. Available:

<https://www.fluke.com/ko-kr/product/industrial-imaging/sonic-industrial-imager-ii900>

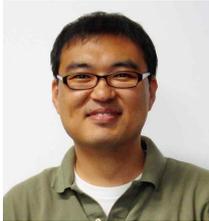
- [3] K. Yang, B. Russell, J. Salamon "Telling Left from Right: Learning Spatial Correspondence of Sight and Sound," *Computer Vision and Pattern Recognition*, pp. 9929-9938, 2020. <https://doi.org/10.48550/arXiv.2006.06175>
- [4] S. Y. Lee, J. Changl, S. Lee "Deep learning-based method for multiple sound source localization with high resolution and accuracy," *Mechanical Systems and Signal Processing*, Vol 161, 2021. <https://doi.org/10.1016/j.ymsp.2021.107959>
- [5] D. Jain, Leah Findlater, J. Gilkeson, Benjamin Holland, R. Duraiswami, D. Zotkin, Christian Vogler, Jon Froehlich "Head-Mounted Display Visualizations to Support Sound Awareness for the Deaf and Hard of Hearing," CHI '15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp. 241-250, April 2015. <https://doi.org/10.1145/2702123.2702393>
- [6] App Store Preview. Mosquitone Detector [Internet]. Available: <https://apps.apple.com/us/app/mosquitone-detector/id1289567807>
- [7] Google Play. Apps on Google Play [Internet]. Available: https://play.google.com/store/apps/details?id=com.google.audio.hearing.visualization.accessibility.scribe&hl=en_US&gl=US
- [8] Librosa. Librosa 0.9.2 documentation [Internet]. Available: <https://librosa.org/doc/latest/index.html>
- [9] anvil. Introducing Anvil [Internet]. Available: <https://anvil.works/blog/introducing-anvil>
- [10] Peter Morville, Ambient Findability: What We Find Changes Who We Become, 1st ed. Sebastopol, CA: O'Reilly, 2005.
- [11] Ian Spence, Stephan Lewandowsky "Displaying proportions and percentages," *Applied Cognitive Psychology*, 5(1), pp. 61-77, 1991. <https://doi.org/10.1002/acp.2350050106>



임영선(Young-Sun Lim)

2017년~현 재: 서강대학교 학부 (이학, 공학학사)

※ 관심분야 : 인공지능(Artificial Intelligence), 인간-컴퓨터 상호작용(HCI)



최용순(Yongsoon Choi)

2013년~현 재: 서강대학교 교수

※ 관심분야 : 웨어러블 인터랙션, 멀티 모델, 멀티 센서리 인터랙션, 미디어 디자인 등