

청각장애인을 위한 사운드 정보 시각화 안경의 개발

이계환, 김인현, 이준호, 이정훈, 황광일
인천대학교 임베디드시스템공학과
e-mail: {oripon7, illyce92, junolee0914, junghoon450}@gmail.com
hkwangil@inu.ac.kr

Development of Sound Information Visualization Glasses for the Hearing Impaired

Gye-hwan Lee, In-hyun Kim, Jun-ho Lee, Jeong-hoon Lee, Kwang-il Hwang
Dept of Embedded System Engineering, Incheon National University

요 약

통계적으로 일반인보다 청각장애인의 교통 사고율이 높은 것으로 나타나는데, 이는 청각 장애로 대표되는 차량을 포함한 수많은 위험 요소를 인식하기 힘든 상태나 조건에서 기인한다. 자동차가 접근하는 등의 소리를 듣지 못한다는 것은 결국 어떠한 위치에 위험요소가 존재하는지 인지하지 못함에 따라 사고로 이어질 가능성이 존재함을 의미하는데 이러한 문제점을 개선함과 동시에 대화중인 사람의 목소리를 시각화하여 정보를 제공함으로써 청각장애인으로 하여금 더 안전하고 쾌적한 삶을 누리게 하는 것이 청각장애인을 위한 사운드 정보 시각화 안경의 개발 목적이다. 위와 같은 배경을 통해 딥러닝 기술에 기반하여 분류 과정을 거친 소리 정보의 판별을 통해 위험요소를 인식한 후 시각화하여 정보를 제공하는 디바이스를 제안한다.

1. 서론

웨어러블 디바이스는 “신체에 부착하여 컴퓨팅 행위를 할 수 있는 모든 전자기기를 지칭하며, 일부 컴퓨팅 기능을 수행할 수 있는 어플리케이션까지 포함함”이라고 정의된다[1]. 2010년대에 이르러 스마트폰과 태블릿 PC 등 스마트 기기의 발전뿐만 아니라 무선통신 인프라의 구축과 배터리 수명향상 등 기술적 한계들이 극복되면서 일상생활에서도 사용이 가능한 수준에 이르렀고, 웨어러블 디바이스에서 수집된 정보를 스마트폰과 같은 전자기기로 M2M(Machine to Machine:사물통신) 방식을 통해 실시간 상호 전송·교환해 서로 연동하는 방식으로 이용되고 있다. 본 연구는 안경 형태의 웨어러블 디바이스의 개발을 통해 청각장애인을 대상으로 하는 사용자에게 일상영역에서 직접적인 도움들을 제공하고자 하며 다음과 같은 목적들을 갖는다.

청각장애인을 대상으로 하는 사용자는 청력의 부재로 인해 차량을 포함한 수많은 위험 요소를 인식하기 힘든 상태나 조건을 갖는다. 이는 결국 사용자가 어떠한 위치에 위험요소가 존재하는지 인지하지 못함에 따라 일상 영역에서 항상 건청인에 비해 높은 사고 발생의 가능성을 안고 살아감을 의미한다. 사용자가 실시간으로 디바이스로부터 제공받는 정보를 통해 다양한 위험 요소로부터 보호받는 것을 첫 번째 목적으로 한다.

청각장애인은 선천적 또는 후천적으로 청력을 상실하여 주로 시각 정보에 의존하여 생활한다. ‘생존’을 위해 구하거나 독순술 또는 수화를 배워 청인(들을 수 있는 사람)과의

의사소통을 하지만 수화의 경우, 일상 영역에서 수화를 구화할 수 없는 청인이 대부분일 뿐더러 그 특성상 정보 전달이 단편적일 수밖에 없으며 실시간 전달되는 정보의 양이 말로 하는 것에 비해 크게 부족하다. 어린 청각장애인이 성장하여 성인이 되어 사회에 나오면 건청인들과 제대로 어울리지 못함은 자명한 일이며 ‘단편 사회 집단’이 형성될 가능성 또한 존재한다. 그러므로 본 연구는 수많은 청각장애인들이 ‘열린 사회’에서 다양한 실시간 정보의 빠른 습득과 원활한 커뮤니케이션 활동을 누릴 수 있도록 도움을 주는 목적 또한 갖는다. 이러한 각각의 목적들은 디바이스 기능을 구현함으로써 실현되며 제품과 서비스의 디자인은 장애인의 요구와 특성을 고려하여 그들의 시각에서 논의하고 개발해야 함이 당연하다[2-4].

2. 구조 및 설계

일반적으로 의사소통은 두 명 이상의 화자와 청자 간의 상호작용을 통해 정보를 전달하고 반응하고 이해하는 과정인 동시에, 물리적 및 사회적 환경에도 크게 영향을 받으면서 시시각각 변화하는 역동성을 지닌 과정이다[5]. 청각장애인을 위한 사운드 정보 시각화 안경이 제공하는 기능은 사용자에게 실시간으로 정보를 제공함으로써 생활안전 및 의사소통 영역에서 직접적인 도움을 주는 것을 목표로 한다. 이를 위해서는 소프트웨어와 하드웨어의 적절하고 효율적인 설계 및 구조가 요구된다.

*이 논문은 2018 한이음 ICT 멘토링 프로젝트의 연구비 지원을 받은 결과물입니다.

2.1 H/W 구조 및 설계

하드웨어는 추상적으로 보았을 때 Nvidia 社의 Jetson TX2 딥 러닝 개발 보드와 사용자가 착용하고 있는 안경으로 구성된다(그림 1). 사용자는 디바이스로부터 도움을 제공받기 위해서 가방을 착용해야 하는데 내부에는 외부 케이스가 장착되어 물리적인 충격으로부터 보호받는 TX2 보드와 전원 공급을 위한 보조배터리가 존재한다. 이를 위해서 ATTA GLOBAL社의 iEnergy 4 모델을 활용하였으며 40000mA의 배터리 용량을 제공함으로써 사용자가 하루 동안 충분히 디바이스를 사용할 수 있도록 한다. 다른 모델을 선택할 경우 TX2 보드에 적합한 19V의 출력 전압을 갖춘 제품이 요구된다.



그림 1. H/W의 구성도

안경은 3D 프린팅 기술로 제작된 디바이스를 안정적으로 고정시키기 위하여 빨데안경이 적합하다고 판단하였다. 안경의 오른쪽 다리에 장착된 디바이스는 Micro OLED 디스플레이 모듈, 거울, 4개의 MIC, 물리적인 버튼, 전원 공급을 위한 배터리 등으로 구성된다. 그림 2에서 확인할 수 있듯이 OLED 디스플레이 모듈의 제어를 통해 원하는 텍스트 정보를 출력하면, 거울의 반사를 통해 안경의 렌즈부에 위치한 정보가 출력될 디스플레이를 향해 빛의 방향을 바꿔준 후 유리를 통과하여 확대시켜줌으로서 시인성을 극대화하여 최종적으로 투명한 플라스틱 소재의 판넬에 투영시켜주는 원리로 동작한다(그림 2).

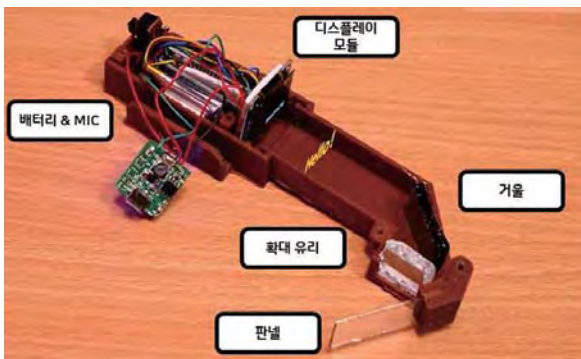


그림 2. 3D 프린팅 기술로 제작된 H/W의 동작 원리

2.2 Google AudioSet

2017년 3월 Youtube에서 YouTube-8M, Youtube-BB에 이어 새로운 데이터셋을 공개하였다. Youtube 동영상에서 추출한 10초 가량의 오디오에 사람이 직접 레이블을 붙였으며 무려 2,084,320개 데이터에 632개의 클래스가 있다. 이 데이터셋은 VGGish 데이터 모델에 학습되어 Tensorflow 환경에서 활용이 가능하다.

2.3 사운드 정보 판별을 통한 위험 요소 인식

어플리케이션(S/W)은 디바이스에 탑재된 물리적인 버튼을 통해 두 가지 동작모드로 제어된다. 버튼을 누르고 있지 않을 때 사운드 정보 판별 및 위치 인식의 기능을 제공한다. 안경 다리 부분에 부착된 디바이스의 마이크로부터 어플리케이션이 동작하는 동안 계속해서 사운드 데이터를 습득한다. 습득한 데이터를 인자값으로 어플리케이션에 넘겨주어 Tensorflow 라이브러리 환경 기반의 Google AudioSet 데이터 모델을 통해 판별 작업을 수행한다. 내부 어플리케이션에서 예상되는 결과 리스트를 확률과 함께 출력 후 확률에 근거하여 필터링 작업을 진행한다. 이후 가장 높은 확률을 지닌 결과값의 객체명이 위험 요소로 분류될 경우 안경의 렌즈부에 위치한 디스플레이에 출력하여 사용자에게 정보를 제공한다(그림 3).

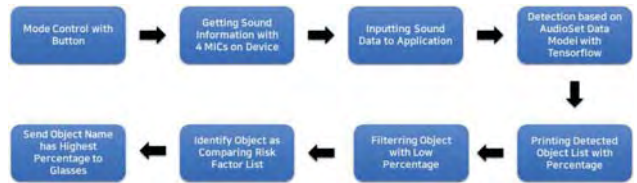


그림 3. 사운드 정보 판별을 통한 위험 요소 인식 동작원리

2.4 실시간 사람 음성 인식

두 번째 동작모드는 실시간 음성인식 기능을 제공하며 그림 4와 같다. 디바이스에 탑재된 물리적인 버튼을 누르고 있는 동안 안경 다리 부분에 부착된 디바이스의 마이크로부터 버튼에서 손을 뗄 때 까지 사운드 데이터를 습득한다. 습득한 데이터를 저장한 후 어플리케이션에서 읽어 들여 Google Speech API와 같은 STT API를 호출하여 인자로서 넘겨준다. 반환받은 인식된 음성 결과를 디스플레이 모듈의 해상도에 맞게 파싱 작업을 진행한 후 파싱된 데이터를 순차적으로 안경의 디스플레이 모듈에 전송한다. 이후 디스플레이 모듈에서 순서대로 결과를 출력하여 안경의 렌즈 부에 위치한 판넬 디스플레이에 투영한다(그림 4).



그림 4. 실시간 사람 음성 인식 동작 원리

3. 실험 결과

실험은 데이터 모델을 기반으로 학습된 어플리케이션의 성능 평가를 위해 진행되었다. 그림 5의 테스트 용 어플리케이션에서 결과를 확인할 수 있으며 마이크로부터 2m 떨어진 거리에서 스피커의 볼륨을 최대 출력의 80%로 맞추어 실제 도로 및 공원 등의 환경에서 들리는 소리의 음량과 유사하게 측정하였다.

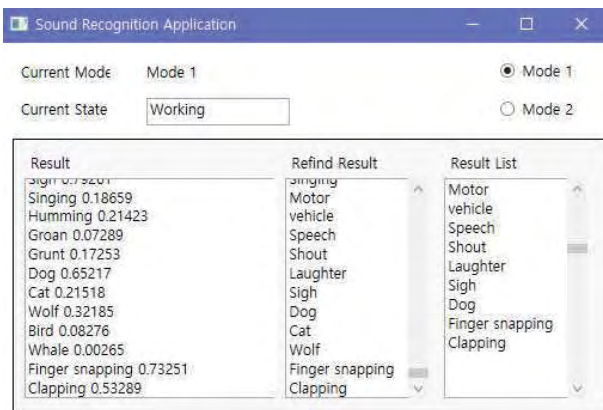


그림 5. 사운드 정보 판별을 위한 테스트 어플리케이션

	Test Source	Detection Result	Measurement (sec)
1	Car Horn	Motor Vehicle	2.3
2	Human Voice	Speech, Shout, Laughter, Sigh	3
3	Lion	Dog	1.8
4	Presentation	Speech, Finger snapping, Clapping	2.4

표 1. 사운드 정보 판별을 통한 위험 요소 인식 결과

데스크탑에서 테스트용 동영상 샘플 4개를 재생하여 실험을 진행한 결과이다. 첫 번째 샘플은 도로에서 자동차의 경적소리가 울리는 영상이며 첫 번째 경적음이 들린 2.3초 후 Motor Vehicle이라는 결과가 출력된 것을 확인하였다. 두 번째로 예능 프로그램을 재생한 결과 표 1의 4행에 해당하는 결과가 나타났다. 이는 모델에 단순히 Human Voice라는 추상적인 단위가 아닌 Speech, Shout 등으로 세분화되어 샘플들이 학습되었기 때문으로 분석된다. 사자

또는 호랑이와 같은 맹수의 울음소리를 스피커를 통해 출력했을 때 Dog에 해당하는 결과가 출력되었는데 이는 Lion 및 Tiger에 해당하는 데이터 셋의 샘플 수 결핍을 원인으로 분석하였으며 개선될 필요가 존재한다. 마지막으로 프레젠테이션 영상을 바탕으로 테스트를 진행한 결과는 두 번째 테스트 결과와 유사하게 나타났다.

실험의 결과를 통해 청각장애인을 대상으로 하는 사용자가 디바이스를 착용함으로써 일상에서 다양한 사운드 정보들을 읽어 들여 빠른 시간 내에 정보를 제공받을 수 있음을 증명하였다.

4. 결론

본 연구를 통해 청각장애인을 대상으로 하는 사용자가 안경형태의 웨어러블 디바이스를 착용 및 기능을 활용함으로써 일상영역에서 실질적인 도움을 제공할 수 있는 방안을 제시하였다. 판넬 디스플레이가 안경의 렌즈 부에 위치하기에 안경을 쓴 상태에서 사용자는 어떠한 행동을 취하지 않아도 쉽게 정보를 제공받을 수 있었다.

다만 안경을 착용한 상태에서 주위 사물을 바라보던 중 정보를 습득하기 위해 투명한 플라스틱 판넬의 텍스트로 눈의 초점을 맞출 때 디바이스 내부의 디스플레이 모듈 및 구성요소 설계상의 성능적인 문제로 가독성 부분에서 취약한 모습을 보이곤 했다. 또한 상용화를 위해서는 사용자가 기능을 사용하지 않을 때 플라스틱 판넬을 디바이스 내부로 숨길 수 있도록 설계하는 것도 사용자의 시야를 방해받지 않도록 하는 방법의 예가 될 수 있겠다.

현재 시장에 출시되어 청각 장애인이 주로 사용하는 디바이스 중 널리 보급화 된 제품은 보청기 뿐인 현실이다. 어느 정도의 청력을 보유한 청각장애인에게는 효과적일 수 있으나 완전히 청력을 상실한 사람에게는 청각장애인을 위한 사운드 정보 시각화 안경이 더욱 도움이 될 것이다. 더욱이 아직까지 Google 社의 Google Glass 외에 보편적이며 유명한 안경 형태의 웨어러블 디바이스가 시장에 보급되지 않은 상황[6]에서 다양하고 효과적인 안경 형태 웨어러블 디바이스의 보급을 촉진시킬 수 있는 계기가 될 것으로 기대한다.

참고문헌

[1] Wearable device definition from MIT Media LAB, <https://www.media.mit.edu/>
 [2] 송지원, 양승호. "청각장애인을 위한 의사소통 시스템의 디자인 제안." Archives of Design Research, 22.1 (2009.02): 197-206.
 [3] 김유리, 송원경, 김성필. "안경형 스마트 기기를 활용한 장애인용 서비스 기획." 대한인간공학회 학술대회논문집, (2015.4): 96-102.
 [4] Rosson M. B. and Carroll J. M., Usability Engineering: Scenario-based Development of Human-Computer Interaction. San Francisco et al.:

Morgan-Kaufman, 2002

[5] Sharp, H., Rogers, Y., & Preece, J. (2007).
Interaction design: beyond human-computer interaction.

[6] 이상국, 김유정, 박정현. "스마트안경 및 음성인식 기술을 활용한 청각장애인의 다자간 대화 지원 시스템 연구." 장애의 재해석, (2017.11): 1-57.