



국민대학교  
전자정보통신대학  
컴퓨터공학부

# 캡스톤 디자인 I

## 종합설계 프로젝트


프로젝트 명	Soundview
팀 명	Viewtiful
문서 제목	계획서

Version	3.0.1
Date	2018-04-09

팀원	고가을(조장)
	김예린
	류성호
	정승우

### CONFIDENTIALITY/SECURITY WARNING


이 문서에 포함되어 있는 정보는 국민대학교 전자정보통신대학 컴퓨터공학부 및 컴퓨터공학부 개설 교과목 캡스톤 디자인 수강 학생 중 프로젝트 “Soundview”를 수행하는 팀 “Viewtiful”의 팀원들의 자산입니다. 국민대학교 컴퓨터공학부 및 팀 “Viewtiful”의 팀원들의 서면 허락없이 사용되거나, 재가공 될 수 없습니다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Soundview	
	팀 명	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

## 문서 정보 / 수정 내역


Filename	계획서-Soundview.doc
원안작성자	고가을, 김예린, 류성호, 정승우
수정작업자	

수정날짜	대표수정자	Revision	추가/수정 항목	내 용
2018-03-08	고가을	1.0	최초 작성	계획서 양식 및 파일 작성
2018-03-09	류성호	2.0	내용 수정	용어 통일
2018-04-04	고가을	3.0	내용 수정	계획 변경에 따른 내용 수정
2018-04-09	고가을	3.1	내용 수정	시스템 구조도 변경
2018-04-12	고가을	3.2	내용 수정	Soundview 소개란 내용 추가

 <b>국민대학교</b> <b>컴퓨터공학부</b> <b>캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	Soundview	
	<b>팀 명</b>	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

## 목 차

<b>1</b>	<b>개요</b>	<b>4</b>
1.1	프로젝트 개요	4
1.2	추진 배경 및 필요성	4
1.2.1	추진 배경 및 사전 조사 내용	4
1.2.2	기술 시장 현황	6
1.2.3	개발된 시스템 현황	7
1.2.4	개발된 시스템의 문제점	9
<b>2</b>	<b>개발 목표 및 내용</b>	<b>10</b>
2.1	목표	10
2.2	Soundview 소개	10
2.3	연구/개발 내용	11
2.4	개발 결과	15
2.4.1	시스템 기능 요구 사항	15
2.4.2	시스템 비기능(품질) 요구 사항	15
2.4.3	시스템 구조	16
2.4.4	결과물 목록 및 상세 사양	17
2.5	기대효과 및 활용방안	18
<b>3</b>	<b>배경 기술</b>	<b>19</b>
3.1	기술적 요구사항	19
3.2	현실적 제한 요소 및 그 해결 방안	19
3.2.1	하드웨어	19
3.2.2	소프트웨어	19
3.2.3	기타	19
<b>4</b>	<b>프로젝트 팀 구성 및 역할 분담</b>	<b>20</b>
<b>5</b>	<b>프로젝트 비용</b>	<b>20</b>
<b>6</b>	<b>개발 일정 및 자원 관리</b>	<b>21</b>
6.1	개발 일정	21
6.2	일정별 주요 산출물	22
6.3	인력자원 투입계획	23
6.4	비 인적자원 투입계획	23
<b>7</b>	<b>참고 문헌</b>	<b>24</b>

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트명</b>	Soundview	
	<b>팀명</b>	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

# 1 개요

## 1.1 프로젝트 개요

실시간으로 소리를 분석해 지정된 상황을 진동 및 디스플레이로 송출하여 소리를 듣지 못 하는 청각장애인들이 상황을 인지할 수 있도록 돕는다. 원하는 경우에는 음성을 인식하여 디스플레이에 송출한다. 또한, 지정된 제스처로 예약문장을 재생하는 등 원활한 의사소통을 돕는 투명 디스플레이를 탑재한 안경형 디바이스 및 소프트웨어를 개발하는 것을 목표로 한다.

이 때, 지정된 상황이라 함은 일상 소리와 구별되는 자동차 경적소리, 비정상적인 소음 등 위급 상황을 포함한다. 또한 본인의 이름 혹은 ‘저기’와 같은 지시대명사등의 지정 문장도 포함한다. 이러한 지정된 상황을 청각장애인이 인지할 수 있도록 하여 위급상황을 대비할 수 있게 하고 음성 지시에 대한 적절한 반응을 할 수 있도록 돕는다.

## 1.2 추진 배경 및 필요성

### 1.2.1 추진 배경 및 사전 조사 내용

(단위 : 명)

장애유형별(1)	장애유형별(2)	2016
		계
청각 · 언어	청각	271,843


▲ 국가 통계 포털, 2016, 등록 장애인 수 - 장애유형별 (청각)

국가 통계 포털의 통계 자료에 의하면, 2016년도 기준 대한민국에 있는 청각장애인의 수는 약 27만 명으로 전체 인구 수인 약 5120만 명 중 약 0.5%에 해당한다.

하지만 잇따라 발생하는 청각장애인의 사고는 많은 사람들의 안타까움을 자아낸다.

2017년 1월, 종로3가역 근처 낙원동에 있는 철거 중인 호텔 공사장이 붕괴됨에 따라 공사현장의 사고로 청각장애인 1명을 포함한 노동자2명이 사망하는 일이 있었다. 이 사고에 대해 사람들은 다음과 같은 반응을 보였다.

“선진국처럼 평상시 구조자와 장애인이 함께하는 대피훈련을 통해 신뢰감을 갖도록 하는 대책이

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	계획서		
	프로젝트 명	Soundview	
	팀명	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

필요하지만, 재난상황을 총괄하고 국민의 안전을 담보해야할 국민안전처와 관련 홈페이지 그 어디에서도 재난발생시 장애인 대처방법 관련 자료를 찾을 수 없다.”

‘청각 장애인 인부 사망, 장애인 재난대책 절실’, <에이블뉴스>, 2017/01/09 - 15:15:07,  
<http://www.ablenews.co.kr/News/NewsContent.aspx?CategoryCode=0014&NewsCode=00142017010915061855>  
 9753 (2018/03/07-접속 날짜)

“청력의 손상으로 인해 시각에 의존하고 있어 돌발상황에서 소리를 통한 초기 인지가 어렵고, 사이렌이나 방송, 재난 소음 등으로 정보를 수용할 수 없어 상황을 파악할 수 있는 정보가 부족하다.”

‘대피로 찾을 수 없는 재난 속 청각장애인’, <함께걸음>, 2017/02/06 - 10:18:40, ,  
<http://www.cowalknews.co.kr/news/articleView.html?idxno=15357> (2018/03/07-접속 날짜)

위와 같이, 청각 장애인에게 위험한 상황 발생 시, 상황을 파악할 수 있는 정보가 부족할 뿐 더러 대처방법 관련 자료에 대한 정보는 찾기 힘들다. 뿐만 아니라, 정보가 있음에도 지침 요령은 실제 상황 발생 시 적용하기 힘든 경우가 많다. 행정안전부가 운영하는 별도 사이트인 ‘안전한 TV’ 에서 ‘지진 대응 요령 - 장애인’ 편을 참고하면, 지진 발생시 청각장애인에게 다음과 같이 행동하라고 알려준다.

**‘얼굴을 똑바로 들고, 입을 뿔 수 있는 대로 크게 움직여 상황을 전달합니다.’**


이 지침은 비장애인의 대피 요령과 크게 다를 바가 없으며, 이 지시대로 움직여도 안전하게 재난 대피를 할 수 있을지에 관해서는 의문이다.

청각장애인은 위급 상황에 대한 정보의 습득이 즉각적이지 않아 빠르게 대처할 수 없게 되고 이는 인명 피해로 이어질 수 있다. 이는 재난 상황 뿐만이 아니다. 일상 생활의 전반적인 영역에서 청각장애인들은 다양한 고충을 겪고 있다.

생활 속 의사소통의 단절로 인해, 말로 전달되는 정보들에 대한 소외가 생긴다. 뿐만 아니라 말로 이루어지는 교육에서의 기회 제한이 생기고, 동료 학생들과 소통이 어려워 이해하기 어려운 상황이 발생한다. 이는 학습 장애로 이어지고 경우에 따라 사회에 부적응하는 경우로 이어진다.

또한, 취업기회 및 직종 선택에 대해서도 제한적이다. 청각장애인 근로자 취업직종은 대부분 단순기술 직종이 다수인데, 이러한 직종선택은 직업 선정이 적성이나 흥미보다는 의사소통의 부담이 덜한 직종에 편중됨을 보여준다.

이렇듯, 청각장애인이 일상 생활 속 겪는 불편함들의 원인은 ‘의사소통의 단절’에서 이어진다. 이는 위급한 상황에 대한 정보 습득을 느리게 할 뿐만 아니라, 비장애인과 소통의 어려움 속에서 청각장애인들이 겪게 되는 불편함은 사회참여의 어려움으로 이어질 수 있다.

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	Soundview	
	<b>팀 명</b>	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

## 1.2.2 기술 시장 현황

### ● 스마트 안경(Smart Glasses) 시장 현황 및 전망

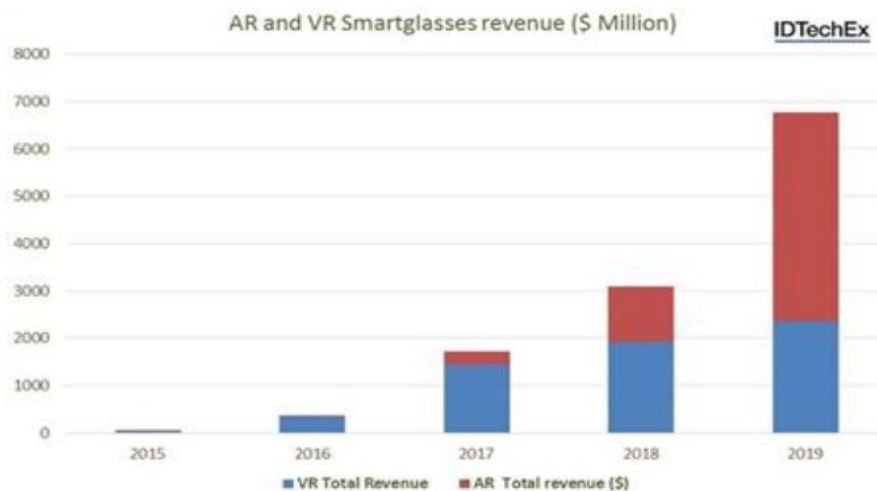
현재 웨어러블 시장은 스마트워치와 스마트 밴드 등의 헬스케어 제품이 주도적이지만, 이들의 성장세는 점차 느려지면서 전체 웨어러블 시장의 75%의 점유율에서 2021년 50% 점유율 수준으로 하락할 것이라 예상된다.

주니퍼 리서치에 따르면, 2016년 약 3억 2700만 달러 규모의 스마트 안경 시장 규모가 2020년에는 90억 달러까지 성장할 것으로 전망한다.


스마트안경은 앞으로 5년간 소비자 웨어러블 시장에서 가장 높은 성장을 보여줄 품목으로, 이런 성장세로 스마트안경은 2021년까지 전체 웨어러블 기기 시장의 11%를 점유할 것으로 예측된다.

글로벌 시장조사업체 스트래티지 애널리틱스(SA)에 따르면, 스마트안경은 2022년까지 스마트워치 다음으로 큰 시장이 될 것으로 전망한다.

이 시장은 10억 명 이상의 안경 착용자 중 5~10%만 대체해도 수천만 시장을 형성한다.

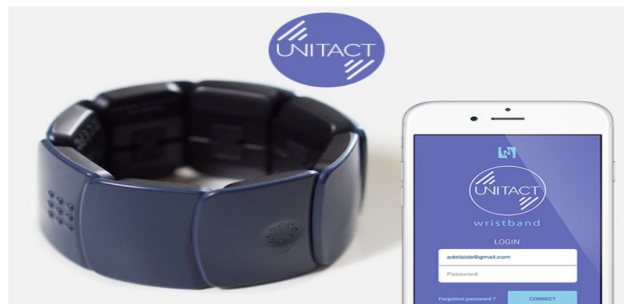


자료원: IDTechEx

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	계획서		
	프로젝트 명	Soundview	
	팀 명	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

### 1.2.3 개발된 시스템 현황

#### 1) 청각장애인들을 위한 커넥티드 팔찌, 'Unitact'



UNITACT, what for ?

“Unitact는 청각 장애인들이 삶의 필수적인 정보를 제공하기 위한 최초의 진동 팔찌입니다. 그들만의 터치감각을 활용하여 일상을 단순화하고 필요한 정보들을 주기 위해서이죠”


Unitact, 'The connected bracelet for deaf people' , <Novitact>, <http://www.novitact.com/fr/unitact/kickstarter>  
(2018/03/07-접속 날짜)

미국의 'Novitact'에서 개발한 'Unitact'는 2016년 10월에 출시한 제품으로, 팔찌의 12가지 다양한 진동을 이용하여 상대방의 스마트폰으로 메시지를 전달할 수 있다. 수화를 모르는 사람들과도 빠르고 간편하게 기본적인 소통이 가능하다. 이 제품은 아이폰과 안드로이드 모두에서 연결할 수 있으며, 한 사람이 아닌 다수에게도 동시에 메시지를 전달할 수 있는 장점을 가지고 있다.

#### 2) 청각장애인들의 소통 돕는 스마트 안경, '쉐어 타이핑 글래스'



▲에이유디와 싸운드룸이 공동 제작한 '쉐어 타이핑 글래스'

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	Soundview	
	<b>팀 명</b>	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

## 실시간 자막 공유+스마트안경=웨어타이핑 글래스, ‘누군가를 돕는 기술을 만들고 싶어요’

‘청각 장애인 소통 돕는 스마트 안경, 웨어 타이핑 글래스’, <성북장애인복지관>, 2015/08/25 - 15:56:37,  
[http://sb.sqwon.or.kr/bj\\_board/bjbrd\\_view.htm?board\\_id=0503&num=466](http://sb.sqwon.or.kr/bj_board/bjbrd_view.htm?board_id=0503&num=466) (2018/03/07-접속 날짜)

웨어타이핑은 ‘에이유디’가 개발한 스마트 안경으로, 웹과 모바일에서 작성한 글을 채팅방에 접속한 사람이 실시간으로 볼 수 있도록 도와준다. ‘구글 문서도구’와 비슷한 서비스이지만 문자가 입력되는 속도가 빠르며 여러 명이 동시 접속해도 속도가 느려지지 않는다.

웨어타이핑은 청각장애인을 연결해주는 역할의 제품으로, 착용한 뒤에 사용자 외 누군가가 타이핑을 해주면 실시간으로 안경에 글자가 입력된다. 현재 강연회, 공청회, 세미나 등에서 활용되고 있으며, 청각장애인이 사회에 참여할 수 있도록 도와주고 있다.

### 1.2.4 개발된 시스템의 문제점

스마트 안경 시장은 스마트 워치 등 다른 IoT 제품들의 시장규모에 비해 틈새 시장에 불과하며, 여전히 높은 가격, 과도한 기능 추가, 접목된 기술력의 한계 등의 우려를 불식시켜야 하는 난제를 가진다.

#### 1) 높은 가격


스마트 안경은 출시된 제품의 수가 아직 적지만, 평균 가격 약 100만원 이상으로 일반 소비자들이 구입하기에는 부담스러운 가격대이다. 일반적인 스마트 제품에는 GPS, 터치패드 디스플레이, 카메라, 모바일 프로세서 등 스마트폰에 들어갈 수 있는 각종 전자부품이 들어가 있다. 대량 생산이 이루어지지 않은 상황에서 ‘소품종 소량 생산’된 스마트 안경의 가격은 부품의 수를 줄이지 않고서는 높을 수밖에 없다.

실제로 ‘구글 글래스’의 경우, 부품 원가가 보통 800~1000달러 선이지만, 실제 판매되는 가격은 1500달러로 추정되었다.

#### 2) 과도한 기능 추가

대부분의 스마트 안경은 청각 장애인이 아닌 비장애인도 사용할 수 있도록 제작되어 과도하게 많은 기능을 넣어 오히려 사용자의 불편을 초래하고 있다. 이는 특정 대상이 아닌 불특정 다수를 대상으로 모두 이로운 기능을 제공하기에 발생한 문제이다. 스마트폰으로도 확인할 수 있는 모든 기능을 스마트 안경으로 확인해야 하는 명확한 이유가 부재하며, 착용 후 외출 시 여러 기능을 담을 만한 안경의 디자인과 무게는 사용자로 하여금 불편함을 불러 일으킨다.



 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	Soundview	
	<b>팀 명</b>	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

### 3) 접목된 기술력의 한계

스마트 안경의 문제는 IoT 기반 디바이스 문제와 일맥상통한다. 또한 안경이라는 제품의 특성상 발생하는 특정 문제점도 있다. 우선, 부족한 배터리 용량과 개인 정보 수집으로 사생활 침해 문제 등이 발생할 수 있는 일반적인 IoT 기반 디바이스의 문제이다.

안경은 사람의 시각과 관련되어 즉각적이고 효과적인 위험 대처에 가장 큰 역할을 하지만, 이는 양날의 검이 되어 시야 확보의 어려움, 눈의 피로도 증가 등의 문제로 변질될 수 있다. 추가적으로, 스마트 안경 내부에 완성도 높은 음성인식 서비스를 내장한 하드웨어는 현재 출시된 바가 없다.

## 2 개발 목표 및 내용

### 2.1 목표

본 프로젝트는 청각 장애인들을 위해 일상 생활에서 필요하다고 판단되어지는 소리를 투명 디스플레이에 띄워주는 스마트 안경 제작을 목표로 한다.


- 1) 상대가 말하는 소리를 텍스트로 바꾸어 투명 디스플레이에 띄워주는 기능
- 2) 일상 생활에서 인지가 필요하다고 판단되는 소리에 대한 알림 기능
- 3) 소리의 크기 및 소리가 들려오는 방향 안내 기능

위 3가지 기능을 통하여 청각 장애인들이 일상 생활에서 겪을 수 있는 어려움을 해소하고자 한다.

### 2.2 SoundView 소개

‘SoundView’는 ‘눈앞에 보여주는 소리’라는 의미를 가지고 있으며, 청각 장애인들의 위험 상황 감지와 그에 따른 그들의 대처 의사를 보다 신속하고 편리하게 일반인들에게 전달할 수 있도록 구현된 스마트 안경이다.

안경처럼 착용하여 주변에서의 소리를 인식하고 인식한 내용을 투명 디스플레이에 보여주어 즉각적으로 소리에 대한 반응을 할 수 있게 한다. 안경에 부착되어 있는 마이크 모듈을 이용하여 소리를 인식하고, 서버에서 소리를 분석한 후 투명 디스플레이를 통해 소리에 대한 정보를 제공해주는 형식이다.

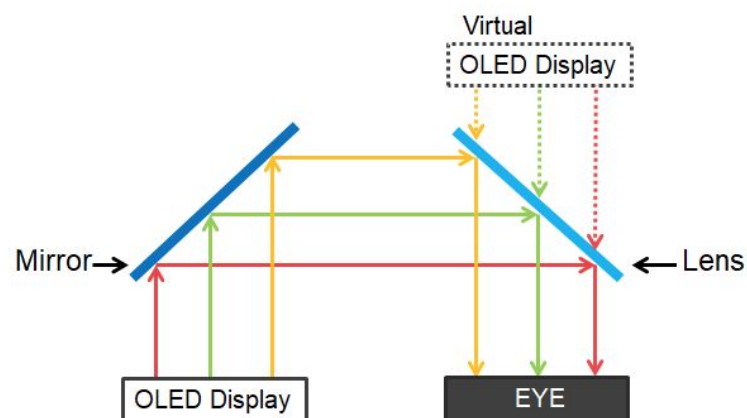
 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	Soundview	
	<b>팀 명</b>	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

추가로 소리의 크기를 판단하여 디스플레이에 띄워주고, 마이크 모듈을 2개 부착하여 소리가 발생한 방향을 알려주는 기능을 추가하고자 한다.




[그림] 스마트 안경 예상 결과물

(출처 : <http://www.instructables.com/id/Uwear-a-DIY-AR-Headset/>)



[그림] OLED 화면이 눈에 보이는 과정

위 그림은 OLED 디스플레이가 사용자의 눈에 보이는 과정을 시각화한 것이다. OLED에서 광원이 나오고, 거울을 통해 그 빛이 반사되면 같은 각도의 투명 렌즈에 비치게 된다. 이는 마치 OLED 디스플레이 광원이 바른 각도 및 형상으로 들어오는 것처럼 보인다.

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	Soundview	
	<b>팀 명</b>	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

## 2.3 연구/개발 내용

안경 형태의 하드웨어에 라즈베리파이를 부착하여 여러 모듈을 사용해 인풋값을 얻도록 한다. 이 때 인풋값을 인식하는 모듈로는 마이크 모듈이 있다. 이 모듈을 통해 사용자 주변의 소리를 인식한 후 소리를 분석하는 작업을 한다. 이 때, 소리의 크기를 분류한 것과 소리가 들려오는 방향을 투명디스플레이에 띄운다.

서버에서 FFT(Fast Fourier transform) 알고리즘을 이용하여 소리를 분석하여 그에 맞는 알람을 투명 디스플레이 또는 스피커를 통해 송출한다. 소리가 음성(사람 목소리)인 경우 음성인식 오픈 API를 이용하여 음성을 분석하고 텍스트로 변환하여 투명디스플레이로 보여준다.

자세한 개발 내용은 다음과 같다.


### 2.3.1 라즈베리파이 및 각종 센서를 이용하여 스마트 안경 하드웨어 구현

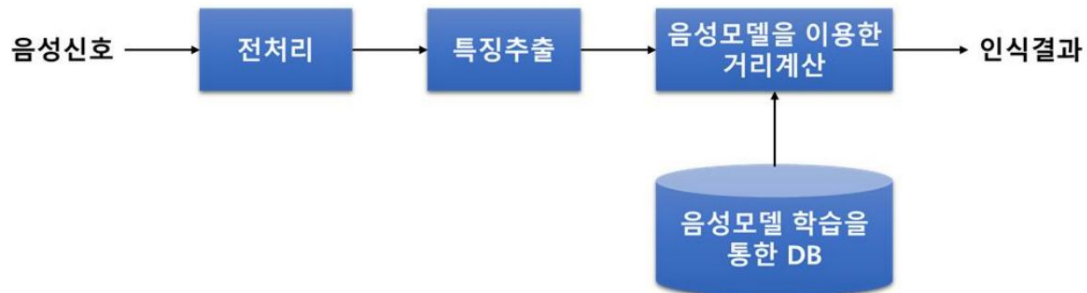
<http://www.instructables.com/id/Uwear-a-DIY-AR-Headset/> 사이트의 내용을 응용하여, Raspberry Pi B+ 모델, 마이크 모듈, 버튼 모듈 등 여러 모듈을 이용하여 스마트 안경을 직접 제작한다.

### 2.3.2 음성인식 및 분석 기능 구현

#### a. 실시간 음성인식


음성인식은 음성 신호로부터 언어적 의미 내용을 식별하는 것으로, 음성 파형을 입력하여 단어나 단어열을 식별하고 의미를 추출하는 처리 과정이다. 크게 음성 분석, 음소 인식, 단어 인식, 문장 해석, 의미 추출의 5 가지로 분류된다. 음성 인식 기술의 원리는 다음과 같다.

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	계획서		
	프로젝트 명	Soundview	
	팀 명	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12



음성 인식 기술의 원리 [출처: 음성 인식 오픈 API의 음성 인식 정확도 비교 분석]

음성 인식 기술에 대한 개발은 활발하게 이루어지고 있다. Google은 인공지능 음성인식 비서인 ‘Google Assistant’를 출시했으며, 자동차, 가전 제품 등에 활용 가능한 음성인식 기능들을 오픈 API로 제공하고 있다. 또한 국내에서도 Naver가 2017년 5월 12일에 음성인식 인공지능인 Naver Clova를 출시하였고, 삼성전자에서도 빅스비를 출시하였다. 세계적인 음성 인식 오픈 API로는 Google의 Cloud Speech API가 있으며, 국내에서는 대표적으로 카카오의 뉴톤(Newtone)과 네이버의 Clova Speech API가 제공되고 있다. [표 1]은 3개의 오픈 API를 여러 특징들로 비교 분석한 표이고, [표 4]는 음성인식 API 종류별 문장 인식결과 표이다.

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	Soundview	
	<b>팀 명</b>	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

[표 1] 음성인식 오픈 API 비교 분석 표[8]

[Table 1] Comparison analysis table of speech recognition open API


	Google Cloud Speech API[9]	카카오 뉴톤(Newtone)[10]	Naver Clova Speech Recognition[11]
지원 자연 언어	한국어 포함 80여개 언어 지원	한국어	한국어, 영어, 일어, 중국어(간체)
지원 컴퓨터 언어	C#, GO, JAVA, NODE.JS, PHP, PYTHON, RUBY	C#, JAVA	C#, JAVA
무료 서비스 제한 범위	한 달에 음성인식 1시간 무료	없음	하루 3,600초
활용 분야	스마트폰, PC, 테블릿, IoT 기기(자동차, TV, 스피커 등)	스마트폰	스마트폰, 웹
기술 지원 범위	SDK 제공, API 문서 제공, 설치 및 설정 가이드, 설정 최적화 예시, 음성인식 개념 설명 제공, 샘플 어플리케이션 제공	SDK 제공, API 문서 제공, 설치 및 설정 가이드, 설정 최적화 예시, SDK 및 API에 대한 일반적인 질문 등	SDK 제공, API 문서 제공, 설치 및 설정 가이드, 설정 최적화 예시, 샘플 APK 제공
커뮤니케이션 채널	공식 커뮤니티 제공, 외부 오픈 소스 커뮤니티 링크 제공	공식 카페 제공	외부 오픈 소스 커뮤니티 링크 제공, 공식 개발자 포럼 제공
인공지능 지원	지원	미지원	지원

[표 4] 음성인식 API 종류별 문장 음성인식 결과

[Table 4] speech recognition result of sentences.

음성인식 API 종류	각 회별 틀린 개수 (각각 전체: 20)										계 (전체:140)
	1	2	3	4	5	6	7	8	9	10	
Google	0	1	0	1	0	0	0	0	0	0	2(20%)
Naver	0	0	0	0	0	0	0	0	0	0	0(0%)
카카오	0	0	0	0	0	0	0	0	0	0	0(0%)

분석 결과, 카카오뉴톤이 한국어 문장 음성인식에 대하여 오인식률이 0%이고, 무료 서비스이지만, 인공지능 지원이 안되며, 안드로이드에 제한적이다. 그래서 Google의 Cloud Speech API를 이용하여 사용자에게 정보를 제공하고자 한다.

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	Soundview	
	<b>팀 명</b>	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

사용자가 버튼을 눌러 음성을 텍스트로 바꿔주는 기능을 실행하고자 하면 Cloud Speech API를 이용하여 소리를 분석하고, 분석된 내용을 바탕으로 텍스트로 변환하여 투명디스플레이에 보여주는 과정을 거친다.

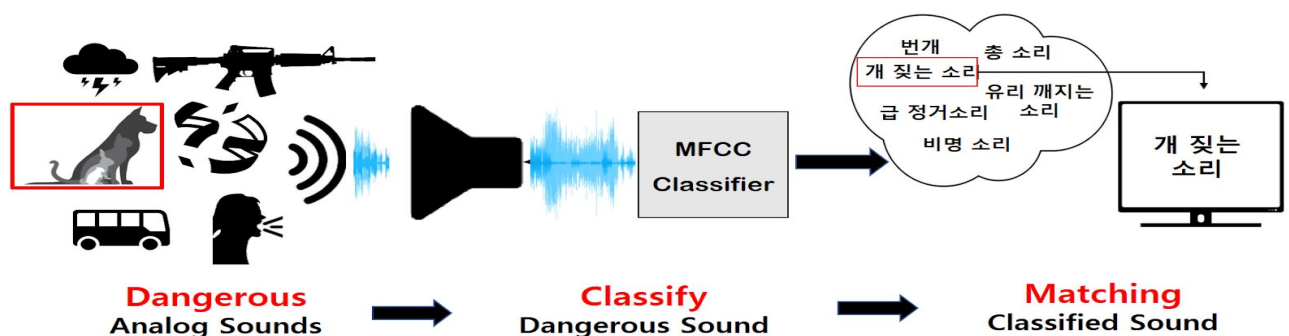
## b. 소리분석

‘갑작스럽게 발생하는 사건들은 일정한 신호 패턴과 특정한 주파수가 있다. 교통의 경우 평소 30~40[dB]와 5,000[Hz] 이하의 주파수대역이다. 하지만 자동차사고 발생 시 스킨드럼은 1[kHz]~3[kHz] 대역에서 소리가 발생한다. 차량 충돌 시 발생하는 음은 주파수 대역 500[Hz] 이하에서 사고가 발생하지 않은 시점보다 30[dB] 이상의 차이를 보인다. 사람들의 비명의 경우에도 구별되는 특징을 찾을 수 있다. 여성의 경우 1,000~2,000[Hz] 구간의 소리를 남성의 경우 500~1,500[Hz] 구간의 소리를 낸다. 또한 전쟁 시 발생하는 위기상황들은 20[kHz]이상의 매우 큰 주파수와 dB 을 보인다.’ (백진환, 기대호, 오하영(2015), "라즈베리파이를 활용한 소리 분석 기반 실시간 사고 탐지 기법", 한국통신학회 2015년도 추계종합학술발표회, p 435)


결과적으로, 주파수에 대한 데이터를 잘 활용한다면 다양한 사고 위험 소리를 감지할 수 있다.

사용자에게 소리에 대한 정보를 제공하기 위해서는 증폭 회로를 사용한 모듈로 전원에 의한 노이즈를 제거하는데 탁월하여 보다 선명한 소리를 만들어주는 ‘Electret Microphone Amplifier – MAX4466 with Adjustable Gain’를 이용한다. 이는 시간에 따라 변화하는 소리의 세기에서 주파수 대역별로 분리된 정보를 추출해내 사용자 주변에서 나는 소리를 감지한다.

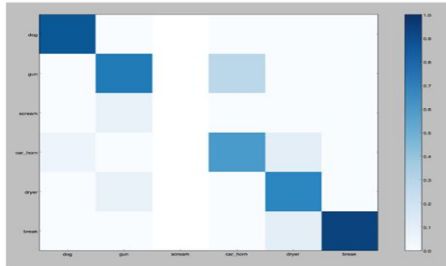
이 때, mfcc 알고리즘을 이용하여 소리의 특징을 추출한다. mfcc는 입력된 소리 전체를 대상으로 하는 것이 아니라, 일정 구간씩 나누어, 이 구간에 대한 스펙트럼을 분석하여 특징을 추출하는 기법이다. 이렇게 추출된 특징 데이터를 .ceps.npy파일로 저장하여 LogisticRegression 분석을 이용하여 분류기를 구축한다. 실시간으로 입력된 소리를 이 분류기를 이용하여 어떤 소리에 해당하는 지 추출해 낼 수 있다.



[표 1] Display Dangerous Sound using MFCC Classifier(분류기)

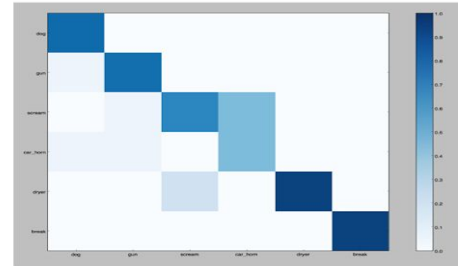
 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	Soundview	
	<b>팀 명</b>	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

<<Confusion matrix of a FFT based classifier>>



[행렬 1]

<<Confusion matrix of a CEPS based classifier>>



[행렬 2]

[표 2] FFT → MFCC (분류기 별 예측 정확도 차이)


처음 소리 데이터 특징 추출을 위해 FFT 파일을 생성하여 로지스틱 회귀 분류기를 만들었다. 분류성능 확인을 위해 시각화한 혼돈 행렬에서는 왼쪽 상단에서 오른쪽 하단까지의 대각선이 검정 사각이며, 나머지 영역은 밝은색일수록 분류 성능이 좋다는 것을 의미한다. 위 표 2의 [행렬 1]에서, FFT 기반 분류기는 높은 성능을 기대하기는 어렵다. 좀 더 정확한 분류를 위해 멜 주파수 캡스트럴 계수인 MFCC를 사용했고, FFT를 이용한 [행렬 1]과 비교하여 MFCC를 이용한 [행렬 2]가 분류 성능이 더욱 높아졌다.

### 2.3.3 소리의 크기 및 소리가 들려오는 방향 안내

#### a. 소리의 크기

위험 상황에 공통적으로 발견되는 소리의 특정 신호 패턴과 주파수 이외에, 소리의 크기는 상황을 인지할 수 있는 또 다른 방법이다. 보통 비장애인들이 감각적으로 느끼는 소리의 크기는 소리의 크기(dB)가 클수록, 소리의 높이(주파수, Hz)가 높을 수록 커진다. 또한 큰 소리가 들릴 수록 더욱 위험한 상황이라 인지하는 경우가 많다. 하지만 실제로 소리의 크기와 소리의 높낮이는 서로 관련이 깊지 않다. 초음파와 같이 인간이 들을 수 없는 영역의 주파수는 위험한 소리가 아닐 뿐더러 야구장, 애견카페와 같은 장소에서 계속 들리는 큰 소리를 위험하다고 생각하는 사람은 없을 것이다. 다음의 표는 사람이 듣는 일상적인 소리의 크기를 db로 표시한다.



 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	Soundview	
	<b>팀 명</b>	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

소리의 종류	dB	느낌
제트기 이륙(60 m 떨어진 곳에서)	120 dB	견디기 어렵다
공사장 소음, 헤비메탈 연주회	110 dB	
고함(1.5m 에서)	100 dB	대단히 시끄럽다
대형 트럭(15m 에서), 굴착기(1m)	90 dB	
대도시 거리 소음	80 dB	꽤 시끄럽다
자동차 실내 소음	70 dB	
보통 대화(1m 떨어져서)	60 dB	보통
교실, 사무실	50 dB	
조용한 거실	40 dB	조용하다
밤중의 침실	30 dB	고요하다
방송국 스튜디오	20 dB	
나뭇잎 스치는 소리	10 dB	거우 무엇인가 들린다
들을 수 있는 가장 작은 소리	0 dB	

[표 1] 사람이 듣는 일상적인 소리의 크기(db)

'소리의크기', <Novitact> <http://www.enjoyaudio.com/study/sound/sound-level.html>

(2018/04/09-접속 날짜)

실시간 소리의 크기 분석은 파이썬 내장 모듈인 pyaudio를 사용해 일상의 소리를 아날로그신호에서 디지털 신호로 수치화 시킨다. 이 때의 수치는 'Audio level', 즉, 전기 파형으로 변환된 디지털 오디오 신호의 전압 레벨로 표현한다.

- Audio (Signal) Level: 가청주파수 범위에 있는 음성 및 모든 자연소리를 전기적인 파형 신호로 변환시키는 것(단위: dBu)


실시간 소리의 수치화 및 분석에 앞서 연속된 소리를 어느 정도 구간의 프레임을 나눠 분석할 지 정하고 이 프레임 안에 저장된 데이터를 사용자에게 전달하기 전까지 임시로 가지고 있을 버퍼의 크기를 정해야 한다. 프레임은 소리의 지속 구간을 나타내는 진동수(Hz)로 일정하게 나누었고, 약 44Kbps로 고정했다. 이 수치는 일반적으로 Audio Level을 구하기 위해 사용되는 값으로 samples(Hz)/second를 나타낸다. 또한 버퍼의 크기는 1024bytes로 지정해 프레임 안의 Audio level 배열을 약 1024byte의 용량만큼 잘라서 수치화에 사용한다. 버퍼의 크기 역시 일반적으로 많이 쓰이는 크기로 지정하였으며, 버퍼는 임계점 이상으로 커질수록 Audio level로의 데이터 변환률이 낮아진다는 특징을 가져 임계점을 찾는 과정이 중요하다.

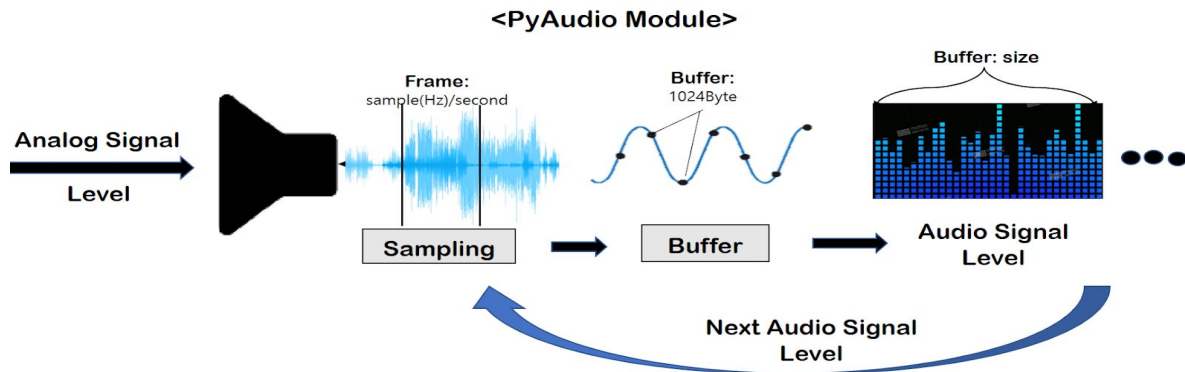
데이터 수치화 작업을 위한 전처리 작업을 마친 뒤 이어질 작업은 Audio Level을 정량화하는 것이다. Audio Level 정량화 표현과 오디오 레벨 미터 방법은 다음과 같다

- 표현: Peak(극값), 실효값(RMS) 등
- 측정: Audio Level Meter(VU Meter, PPM Meter, Average 등)

실시간으로 측정된 일정 구간의 Audio Level의 평균값은 데이터의 최대값과 최소값의 차(약 극값에서의 차)를 이용해서 구한다. 이를 통해 최종 변환된 Audio Level dBu를 무한 loop를 통해 실시간으로 사용자에게 보여지도록 한다.



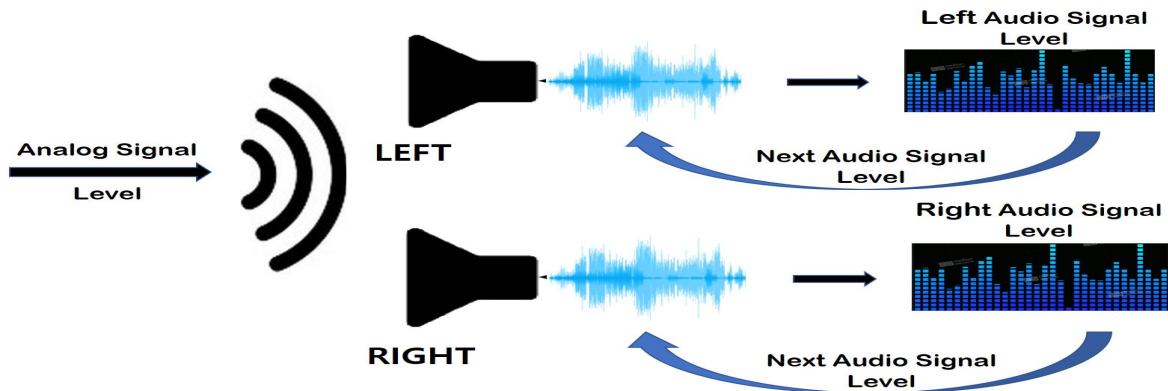
 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Soundview	
	팀 명	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12




[표 1] Real-time audio processing(수치화)

#### b. 소리가 들려오는 방향

위험 소리를 인지하여 사용자에게 알려주는 작업도 중요하지만, 위험 소리가 발생하는 지점도 파악하는 것이 중요하다. 라즈베리 파이 내의 마이크 설정을 통해 좌,우에서 들리는 소리의 크기를 수치화할 수 있다.

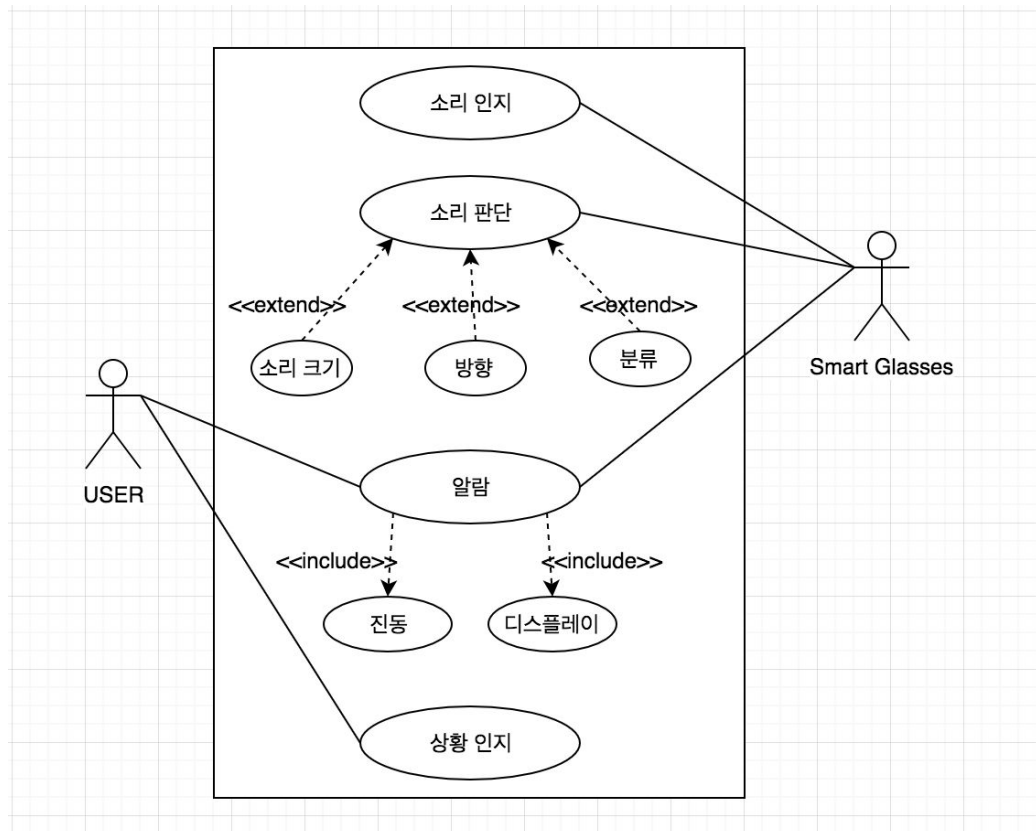


[표 2]Real-time two-side audio processing(수치화)

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Soundview	
	팀 명	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

## 2.4 개발 결과


### 2.4.1 시스템 기능 요구사항



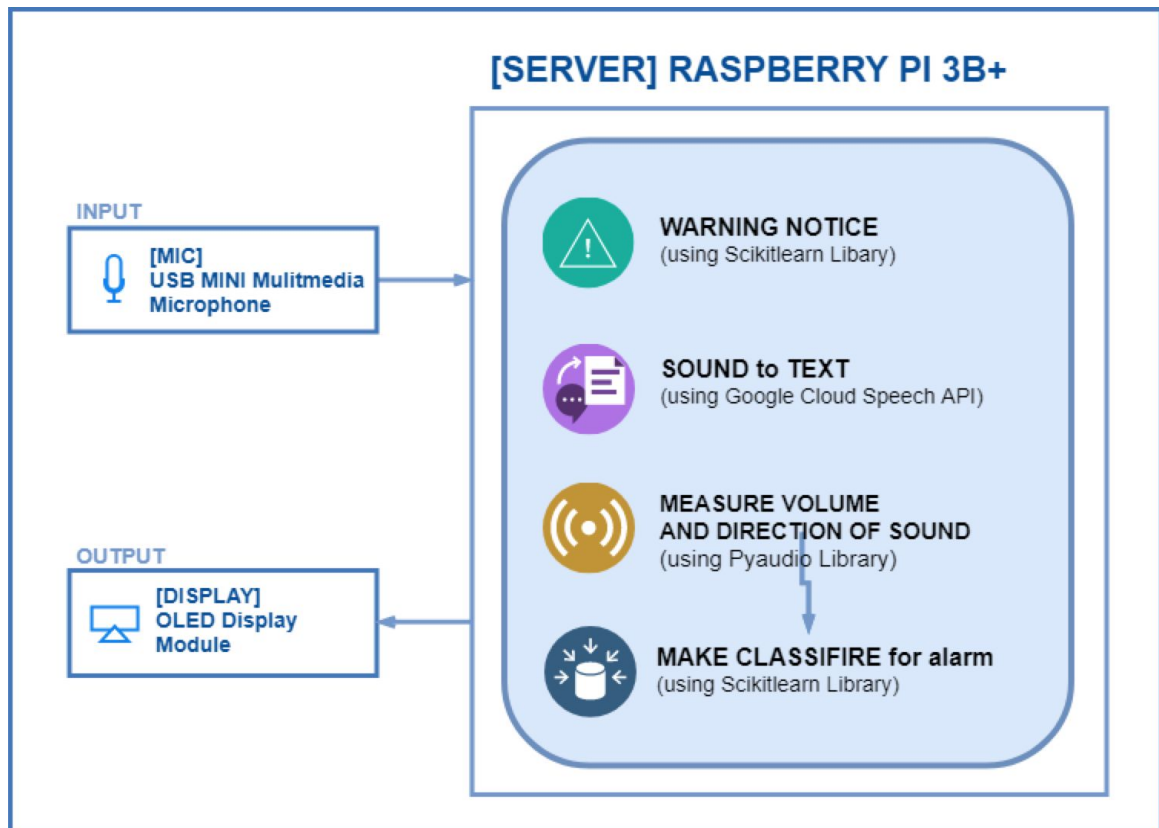
- smart glasses는 마이크 모듈을 통해 소리를 인지하고, 들려오는 소리를 분석한다.
- 소리의 크기, 소리가 들려오는 방향, 어떤 소리인지 분류를 한다.
- 소리 분석을 한 뒤 진동으로 알림을 주거나 투명 디스플레이에 알림을 띄우는 방식으로 사용자에게 알린다.
- 사용자에게 어떤 상황인지, 주변 상황에 대한 정보를 알려 준다.

### 2.4.2 시스템 비기능(품질) 요구사항

특수소리를 인지하고 어떤 유형의 소리인지 판단하여 디바이스로 알림을 주는 것이 1초 이내에 수행되어야 한다.


 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Soundview	
	팀 명	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

## 2.4.3 시스템 구조



## 2.4.4 결과물 목록 및 상세 사양

대분류	소분류	기능	형식	비고
라즈베리 파이(서버)	소리데이터 MFCC변환	마이크에서 받아온 소리데이터를 MFCC를 이용하여 특징데이터를 추출한다.		
	특수상황 분석	인지하고자 하는 특수상황을 사전에 구현해 놓은 분류기를 통해 인지하여 출력한다.	함수	
	예약상황 소리 데이터화	특수상황으로 취급할 새로운 데이터를 등록한다.	함수	
	음성 텍스트화	구글 API를 이용하여 마이크에서 받아온 음성데이터를 텍스트로 변환하여 출력한다.	함수	
	소리 크기 및 방향 분석	마이크에서 받아온 소리데이터의 소리크기 및 방향을 분석하여 출력한다.	함수	

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	Soundview	
	<b>팀 명</b>	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

## 2.5 기대효과 및 활용방안

스마트 안경 산업은 구글이 지난 2014년 ‘구글 글래스’ 출시를 통해 전 세계에 스마트 안경을 처음 알린 뒤, 2018년 현재에도 꾸준히 시장 규모가 성장하고 있는 산업이다.


그러나 시장에서 통용되는 스마트 안경은 대중들을 대상으로 스마트폰의 기능을 입체적으로 보여주는 데에 그쳤다. 구글 글래스의 경우, 비싼 가격과 시야를 가려 오히려 사용자의 불편을 초래했다는 의견들로 시장에서 큰 인기를 얻지 못했으며, 구글의 명성에 맞지 않는 완성도로 이미지에 큰 타격을 입혔다.

기술적, 실용적 한계에 따라 스마트 안경의 최대 수혜자는 일반인이 아닌 장애인이다. 그 중 소리를 듣지 못해 위험 감지에 더디고, 대처 의사가 수화를 모르는 일반인에게 전달 되지 못하는 청각 장애인들에게는 스마트 안경의 필요성이 절실하다.

실제로 일반인들에게 대수롭지 않은 위험이 청각장애인들에게는 위험이 되고, 비록 위험 상황을 알지만 그에 따른 대처를 신속하게 하지 못해 더 큰 위험에 빠진 사례들이 여럿 발견된다.

SoundView는 또한 그 활용을 비장애인들에게도 넓힐 수 있다. 수 많은 위험이 도사리는 현실에 스마트 안경은 아이들에게 안전 지킴이가 될 수 있으며, 위험에 대한 대처가 비교적 느린 노인들에게는 누구보다 가까운 곳에서 그들을 지키는 구조대가 될 수 있다.

SoundView가 대량 제작되어 유비쿼터스를 선도하는 디바이스로 발전된다면, 그 가치와 의미는 새로운 기술의 변화에 맞게 그 빛을 더욱 발할 것이다.

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	Soundview	
	<b>팀 명</b>	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

## 3 배경 기술

### 3.1 기술적 요구사항

#### 3.1.1 프로젝트 개발 환경

- 1) 운영체제  
Raspbian OS / Ubuntu MATE
- 2) 개발언어  
C++ / Python / Java / C#
- 3) 라이브러리  
Scikit-learn Lib / Google Cloud Speech API
- 4) 컴파일러  
g++ / JDK

#### 3.1.2 프로젝트 결과물 확인 환경

- 1) 디바이스  
Raspberry Pi 3B, 각 종 센서 및 모듈을 연동한 안경형 디바이스


### 3.2 현실적 제한 요소 및 그 해결 방안

#### 3.2.1 하드웨어

라즈베리파이의 전원 공급 측면에서 여러 기능을 포함한 프로그램을 계속 실행시킨다면 급격한 배터리 소모가 이루어질 수 있다. 이 경우, 버튼 모듈을 사용하여 구동한 상황에만 각 기능의 프로그램이 돌아가도록 한다.

#### 3.2.2 소프트웨어

상대방이 한 말을 텍스트로 바꿔주는 자막 기능의 경우 실시간으로 처리가 되어야한다. 따라서 기존 사용하던 Raspberry Pi 3 B 모델에 비해 프로세서 처리 속도가 약 1.2배 빨라진 Raspberry Pi 3 B + 모델을 사용하여 성능을 높이고자 한다.

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	Soundview	
	<b>팀 명</b>	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

### 3.2.3 기타

착용 소프트웨어 개발 중심으로 이루어지는 프로젝트에서 성능만을 고려한 스마트 안경의 디자인이 착용자들로 주변의 이목을 너무 집중시켜 착용에 반발감이 들 수 있음을 염두해두어야 한다.


따라서 안경 형태가 아닌 캡모자의 형태로 만드는 것도 고려한다. 캡모자에 OLED 디스플레이를 다는 경우, 눈과 OLED 디스플레이 사이의 거리를 고려하여 제작하도록 한다.

## 4 프로젝트 팀 구성 및 역할 분담

이름	역할
고가을	<ul style="list-style-type: none"> <li>- Team Leader</li> <li>- 라즈베리파이 환경 설정</li> <li>- speech to text</li> </ul>
김예린	<ul style="list-style-type: none"> <li>- 소리 데이터 수집 및 정제</li> <li>- 실시간 소리 분류기 구현</li> </ul>
류성호	<ul style="list-style-type: none"> <li>- 소리 크기 분류 및 출력</li> <li>- 소리 들려오는 방향 판단</li> </ul>
정승우	<ul style="list-style-type: none"> <li>- 소리 데이터 수집 및 정제</li> <li>- 실시간 소리 분류기 구현</li> </ul>

## 5 프로젝트 비용

항목	예산치 (MD)
아이디어 구상	60
개발 환경 구축	5
하드웨어 제작	30
DB 구축	55
음성 데이터 FFT 변환	40
예약상황 주파수 데이터 등록 기능	10


 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	Soundview	
	<b>팀 명</b>	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

DB에 삽입된 특수 상황 주파수와 일치 확인	50
제스처 등록 및 인식 후 음성 송출	60
합	290

## 6 개발 일정 및 자원 관리

### 6.1 개발 일정


항목	세부내용	1월	2월	3월	4월	5월	6월	비고
요구사항분석	요구 분석							
	아이디어 구상							
관련분야연구	소리 분석 및 mfcc, ceps 연구							
	관련 시스템 분석							
설계	시스템 설계							
	디바이스 설계							
구현	소리 유형 인식							
	음성 텍스트화							
	소리 크기 판단							
	소리 방향 판단							
테스트	OLED 디스플레이 및 디바이스 성능 테스트							
	소리 유형 인식 테스트							
	소리 크기 판단 테스트							
	소리 방향 판단 테스트							
최종 발표	발표 준비 및 발표							

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	Soundview	
	<b>팀 명</b>	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

## 6.2 일정별 주요 산출물

마일스톤	개요	시작일	종료일
계획서 발표	시스템 설계 완료 <b>산출물 :</b> 1. 프로젝트 수행 계획서	~	2018-03-09
중간 평가	음성을 텍스트로 바꾸는 기능 구현 완료 소리 크기 판단 및 출력 구현 완료 <b>산출물 :</b> 1. 프로젝트 중간 보고서 2. 프로젝트 중간 소스 코드	2018-03-10	2018-04-13
구현 완료	소리를 통한 특수 상황 학습 및 인지 기능 구현 완료 소리가 발생하는 방향 판단 구현 완료 <b>산출물</b> 1. 각 기능 소스코드	2018-04-14	2018-05-22
최종 보고서	최종 발표 <b>산출물:</b> 1. 최종 보고서 2. 프로젝트 최종 소스코드	2018-05-23	2018-05-29




 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Soundview	
	팀 명	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

### 6.3 인력자원 투입계획

이름	개발항목	시작일	종료일	총개발일(MD)
전원	요구분석/아이디어 구상	2018-01-01	2018-03-09	60
김예린 / 정승우	소리 분석 및 특징추출 연구	2018-03-08	2018-04-30	40
전원	시스템 설계	2018-03-08	2018-04-30	5
전원	디바이스 설계	2018-04-10	2018-05-01	30
김예린 / 정승우	소리데이터 수집 및 정제	2018-03-08	2018-04-01	30
김예린 / 정승우	소리 유형 인식 구현 및 성능 향상	2018-03-08	2018-05-29	40
고가을 / 류성호	음성 텍스트화 구현	2018-03-08	2018-05-29	20
고가을 / 류성호	소리 크기 분류	2018-03-08	2018-05-29	20
고가을 / 류성호	소리가 발생하는 방향 판단 구현	2018-04-09	2018-05-29	20
전원	디스플레이 및 디바이스 성능 테스트	2018-05-01	2018-05-29	15
전원	소프트웨어 성능 테스트	2018-05-01	2018-05-29	15
전원	발표 준비 및 발표	2018-05-20	2018-06-01	10

### 6.4 비 인적자원 투입계획

항목	Provider	시작일	종료일	Required Options
Raspberry Pi 3B	Raspberry	2018-03-12	2018-06-01	
개발용 노트북 4대	LG, Apple, Lenovo	2018-03-12	2018-06-01	

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	Soundview	
	<b>팀 명</b>	Viewtiful	
	Confidential Restricted	Version 3.2	2018-APR-12

## 7 참고 문헌

번호	종류	제목	출처	발행년도	저자	기타
1	보고서	청각·언어장애인(농아인)의 의사소통 접근성 강화방안 연구	강남대학교	2013	강남대학교 산학협력단	
2	논문	라즈베리파이를 활용한 소리 분석 기반 실시간 사고 탐지 기법	한국통신학회	2015	백진환, 기대호, 오하영	
3	보고서	스마트안경 및 음성인식 기술을 활용한 청각장애인의 다자간 대화 지원 시스템 연구	장애의 재해석	2017	이상국, 김유정, 박정현	
4	발표자료	청각장애인과 함께 행복한 소통 그리고 사회문제해결	에이유디 사회적협동조합	2016	aud	
5	신문	청각 장애인 인부 사망, 장애인 재난대책 절실	에이블뉴스	2017	이슬기	
6	기사	대피로 찾을 수 없는 재난 속 청각장애인	함께걸음	2017	조은지	
7	포스팅	청각장애인 소통 돕는 스마트안경 '웨어타이핑 글래스'	성북장애인복지관	2015	김대원	