

# 캡스톤 디자인 I 종합설계 프로젝트


프로젝트 명	Soundview
팀 명	Viewtiful
문서 제목	중간보고서

Version	1.3
Date	2018-04-09

팀원	고가을 (조장)
	김예린
	류성호
	정승우
지도교수	우중우 교수

## 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	중간보고서		
	프로젝트 명		
	팀 명		
	Confidential Restricted	Version 1.3	2018-APR-09


## 문서 정보 / 수정 내역

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

수정날짜	대표수정자	Revision	추가/수정 항목	내 용
2018-04-09	고가을	1.0	최초 작성	
2018-04-10	정승우	1.1	내용 수정	향후 계획 내용 추가
2018-04-11	류성호	1.2	내용 수정	소리의 크기 및 방향 내용 추가
2018-04-12	김예린	1.3	내용 수정	시스템 그림 추가
2018-04-12	고가을	2.0	내용 수정	목차 정리

## 목 차

1	프로젝트 목표	3
2	수행 내용 및 중간결과	4
2.1	계획서 상의 연구내용	4
2.2	수행내용	8
3	수정된 연구내용 및 추진 방향	11
3.1	수정사항	11
4	향후 추진계획	12
4.1	향후 계획의 세부 내용	12
5	고충 및 건의사항	13


 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명		
	팀 명		
	Confidential Restricted	Version 1.3	2018-APR-09

## 1 프로젝트 목표

본 프로젝트는 청각 장애인들을 위해 일상 생활에서 인지가 요구되는 소리를 투명 디스플레이상의 적절한 출력을 통해 인지할 수 있도록 돕는 스마트 글라스 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.

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>중간보고서</b>		
	<b>프로젝트 명</b>		
	<b>팀 명</b>		
	Confidential Restricted	Version 1.3	2018-APR-09

## 2 수행 내용 및 중간결과

### 2.1 계획서 상의 연구내용

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 출력한다.

서버에서 MFCC 알고리즘을 이용하여 소리의 특징을 추출하여, 추출한 데이터를 이용해 회귀분석으로 구축한 분류기를 통해 소리를 분류한다. 이후 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.


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

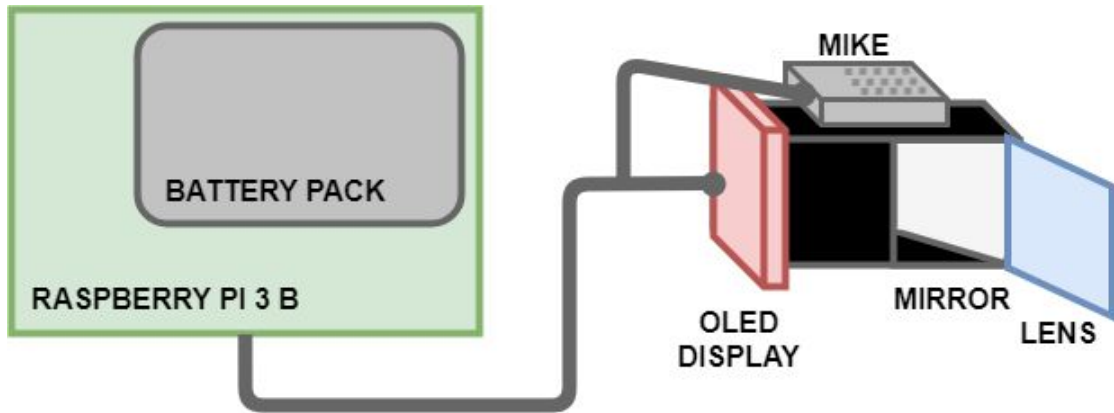
Raspberry Pi B+ 모델, 마이크 모듈, 버튼 모듈 등 여러 모듈을 이용하여 스마트 안경을 직접 제작한다.



[그림] 스마트 글라스 예상 결과물

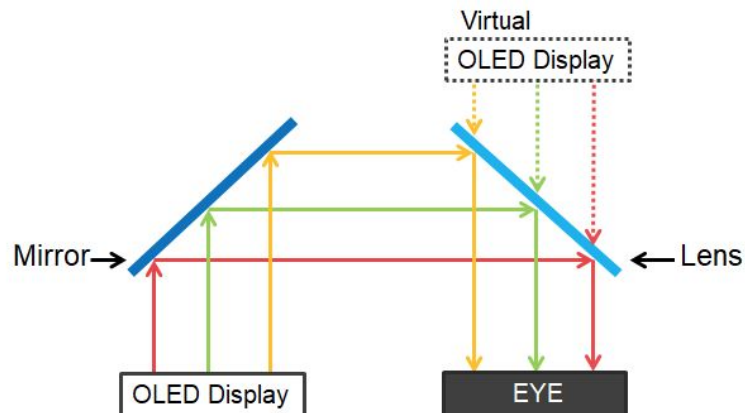
(출처 : <http://www.instructables.com/id/Uwear-a-DIY-AR-Headset/>)

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명		
	팀 명		
	Confidential Restricted	Version 1.3	2018-APR-09




[그림] 예상 하드웨어 구조도

먼저, 주요 기능을 수행하는 라즈베리파이는 항상 전원 공급이 되어야 한다. 따라서 라즈베리파이 크기와 비슷한 크기의 배터리 팩을 이용하여 전원 공급을 한다. 그리고 케이블을 길게 하여 OLED 디스플레이와 마이크 모듈을 연결하면 기본 기능을 모두 구현할 수 있게 된다. 실제 OLED 디스플레이 화면에 띄워지는 자세한 과정은 다음과 같다.



[그림] OLED 광원이 반사각을 이용하여 사용자 눈에 들어오는 과정

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>중간보고서</b>		
	<b>프로젝트 명</b>		
	<b>팀 명</b>		
	Confidential Restricted	Version 1.3	2018-APR-09

위 그림은 OLED 디스플레이가 사용자의 눈에 보이는 과정을 시각화한 것이다. OLED에서 광원이 나오고, 거울을 통해 광원이 반사되면 같은 각도의 투명 렌즈에 비치게 된다. 이는 마치 OLED 디스플레이의 광원이 전방에서 들어오는 것처럼 보인다.

## 2.1.2 음성인식 및 분석 기능 구현

### a. 실시간 음성인식

본 프로젝트에서는 Google의 Cloud Speech API를 이용하여 사용자에게 정보를 제공하고자 한다.


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

	Google Cloud Speech API[9]
지원 자연 언어	한국어 포함 80여개 언어 지원
지원 컴퓨터 언어	C#, GO, JAVA, NODEJS, PHP, PYTHON, RUBY
무료 서비스 제한 범위	한 달에 음성인식 1시간 무료
활용 분야	스마트폰, PC, 태블릿, IoT 기기(자동차, TV, 스피커 등)
기술 지원 범위	SDK 제공, API 문서 제공, 설치 및 설정 가이드, 설정 최적화 예시, 음성인식 개념 설명 제공, 샘플 어플리케이션 제공
커뮤니케이션 채널	공식 커뮤니티 제공, 외부 오픈 소스 커뮤니티 링크 제공
인공지능 지원	지원

[표] Google Cloud Speech API 분석

### b. 소리분석

주파수에 대한 데이터를 잘 활용한다면 다양한 사고 위험 소리를 감지할 수 있다. 사전에 정제된 위험소리데이터를 기반으로 구축해놓은 로지스틱회귀 분석 분류기를 통해 실시간으로 주변의 소리를 분류하여 디스플레이에 소리에 대한 정보를 출력한다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명		
	팀 명		
	Confidential Restricted	Version 1.3	2018-APR-09


### 2.1.3 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 높이는 관련이 깊지 않다. 초음파와 같이 인간이 들을 수 없는 영역의 주파수는 위험한 소리가 아닐 뿐더러 또는 야구장, 애견카페와 같이 장소에서 계속 들리는 큰 소리를 위험하다고 생각하는 사람은 없을 것이다.

#### b. 소리가 들려오는 방향

뒤에서 어떤 소리가 들리는 경우 비장애인은 소리가 난 방향으로 뒤를 돌아보고 몸을 피해 위험 상황을 모면한다. 하지만 청각장애인들은 소리를 들을 수 없기 때문에 소리의 위치를 정확히 찾는데 어려움을 느껴 사고의 위험에 쉽게 노출 될 수 있다. 따라서 위험 소리를 인지하여 사용자에게 알려주는 작업도 중요하지만, 위험 소리가 발생하는 지점도 파악하는 것이 중요하다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명		
	팀 명		
	Confidential Restricted	Version 1.3	2018-APR-09

## 2.2 수행내용

### 2.2.1 라즈베리파이 및 각종 센서를 이용하여 스마트 안경 하드웨어 구현의 수행 내용

하드웨어를 구현하는 데에 필요한 모듈 및 장비에는 ‘OLED Display Module’, ‘스위치 모듈’, ‘마이크 모듈’이 있다. ‘OLED Display Module’, ‘스위치 모듈’의 배송 기간으로 인해 스마트 안경 하드웨어 구현이 불가능하므로 소프트웨어 개발과 순서를 바꾸어 마이크 모듈을 사용한 기능을 먼저 구현하도록 한다.

### 2.2.2 소리인식 및 분석 기능 구현의 수행 내용

#### a. 실시간 음성인식

Google cloud speech API를 이용하여 라즈베리파이의 마이크 모듈에서 받아오는 음성데이터를 실시간으로 텍스트변환하는 것을 구현하였다. 추후 디스플레이에 배치할 적정 길이 및 시간 등을 최적화하고자 한다.


#### b. 소리분석

라즈베리파이에 마이크 모듈을 연결하여 받아온 소리를 일정한 길이로 잘라 wav파일 형식으로 저장한다. 저장한 wav 파일에서 MFCC 알고리즘을 이용하여 특징데이터를 추출해 npy 파일로 저장하는 기능을 구현하였다.

로지스틱회귀알고리즘으로 정제된 소리의 특징을 뽑아낸 집단 데이터를 통해 일정 수준이상의 분류성능을 내는 분류기를 구현하였다.

현재 구현된 분류기는 정제된 소리 데이터를 기반으로 구현되어 실시간으로 잡음과 함께 들어오는 소리데이터에 대하여 분류하는 기능이 추가적으로 구현되어야 하며, 이때 분류 성능이 일정 이상 보장되도록 구현하여야 한다.

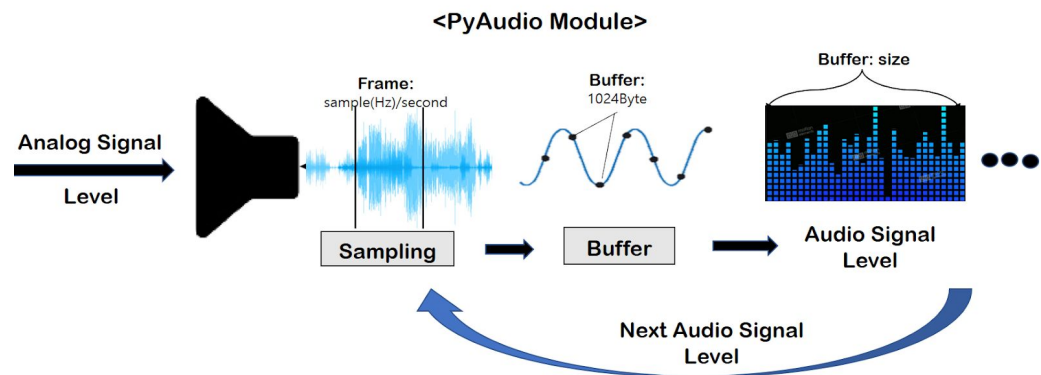


 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명		
	팀 명		
	Confidential Restricted	Version 1.3	2018-APR-09

### 2.2.3 소리의 크기 및 소리가 들려오는 방향 안내

#### a. 소리의 크기

일상생활에서의 아날로그 신호를 분석이 가능한 Audio Level로 실시간 변환한다.  
Audio level로 변환된 디지털 신호의 크기는 OLED화면을 통해 사용자에게 제공된다.  
실시간 소리의 크기는 파이썬의 모듈 pyaudio를 사용해 수치화되고, 이를 사용자에게 잘 보여지도록 Audio Level Meter 중 Peak값을 사용한 변환 방법을 사용한다.




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.

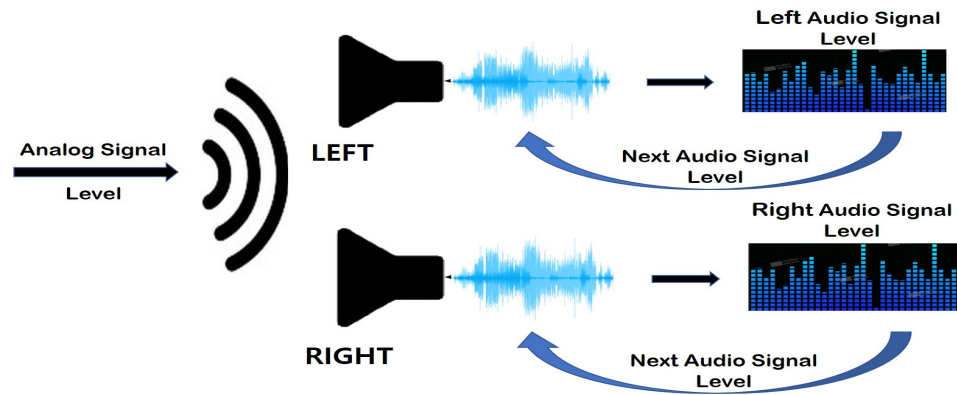
#### b. 소리가 들려오는 방향

라즈베리 파이 내의 마이크 모듈을 2개 연결하여 소리의 방향을 판단할 수 있도록 설정한다. 위 소리의 크기를 변환하는 방법을 통해 일상의 소리를 받아오고, 구간 별로 나눠진 소리의 구간은 또 다시 구간 별 분석을 이용해 좌/우 구분이 가능하게 한다.

사용자는 변환된 방향별 소리의 크기를 실시간으로 OLED화면에 전달받고, 이를 통해

 <div> <p>국민대학교 컴퓨터공학부 캡스톤 디자인 I</p> </div>	중간보고서		
	프로젝트 명		
	팀 명		
	Confidential Restricted	Version 1.3	2018-APR-09

상황에 대한 위험 판단을 내리도록 한다. 위험 신호별 주파수 분석과 같이 사용자에게 계속적으로 제공되며 사용자의 임의로 서비스를 켜거나 끌 수 있다.



### 3 수정된 연구내용 및 추진 방향

#### 3.1 수정사항

##### 3.1.1 목표 변경

본 프로젝트의 기존 목표는 다음과 같았다.

- 1) 상대가 말하는 소리를 텍스트로 바꾸어 투명 디스플레이에 띄워주는 기능
- 2) 일상 생활에서 인지가 필요하다고 판단되는 소리에 대한 기능
- 3) 영상 처리를 통해 사용자의 제스처를 인식하여 등록된 음성을 나오게 하는 기능

하지만 3) 번 기능을 구현하기 위해서는 사용자의 제스처를 인식하는 영상 처리와 음성이 등록된 Database 를 구축하는 과정이 필요하다. 뿐만 아니라 스피커모듈을 추가로 연결해 음성이 나오게 해야하므로 스마트 안경 제작에도 영향을 준다. 따라서 3) 번 기능을 다음과 같이 바꾼다.

##### 소리의 크기 및 소리가 들려오는 방향 안내 기능

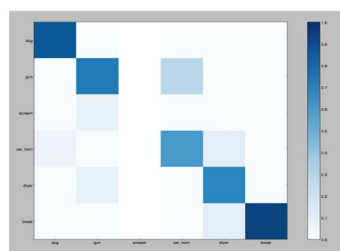
예상 결과물에는 '2) 일상 생활에서 인지가 필요하다고 판단되는 소리에 대한 기능'을 포함한다. 하지만 구현 과정에서 일상 생활의 소리를 모두 구현하지 못한 경우에는 소리의 볼륨을 통해 사용자에게 판단을 할 수 있도록 한다.

뿐만 아니라, 소리가 들려오는 장소가 어디쯤에 있는지 알려주는 것도 청각 장애인에게 중요한 사항이다. 따라서 마이크 모듈의 개수를 2개로 늘림으로써 사용자에게 소리가 들려오는 방향을 안내해주는 기능을 구현한다.

##### 3.1.2 fft -> mfcc 소리특징 추출 알고리즘 변경

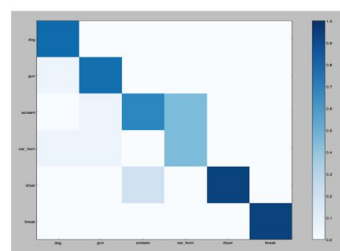
처음 소리 데이터 특징 추출을 위해 FFT 파일을 생성하여 로지스틱 회귀 분류기를 만들었다. 분류성능 확인을 위해 시각화한 혼돈 행렬에서는 왼쪽 상단에서 오른쪽 하단까지의 대각선이 검정 사각이며, 나머지 영역은 밝은색일수록 분류 성능이 좋다는 것을 의미한다. 위 표 2의 [행렬 1]에서, FFT 기반 분류기는 높은 성능을 기대하기는 어렵다. 좀 더 정확한 분류를 위해 멜 주파수 캡스트럴 계수인 MFCC를 사용했고, FFT를 이용한 [행렬 1]과 비교하여 MFCC를 이용한 [행렬 2]가 분류 성능이 더욱 높아졌다.

<<Confusion matrix of a FFT based classifier>>




[행렬 1]

<<Confusion matrix of a CEPS based classifier>>



[행렬 2]

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명		
	팀 명		
	Confidential Restricted	Version 1.3	2018-APR-09

## 4 향후 추진계획

### 4.1 향후 계획의 세부 내용

#### 4.1.1 하드웨어 제작

라즈베리파이와 연결한 마이크 모듈, OLED 디스플레이를 안경 혹은 머리에 착용할 수 있는 형태로 제작하여야 한다. 이 때 디자인보다는 착용시 OLED 디스플레이를 투영하는 투명 아크릴 판의 내용이 사용자가 인지하기 쉽도록 배치를 하는 것에 초점을 둔다. 최적화 중 판단에 따라 안경형태의 디바이스 외관 구성보다 적합하다고 판단되는 구성이 있을 시 유동적으로 바꿀 수 있다. 현재 예상하고 있는 디바이스 외관 구성은 안경형과 모자형으로 총 두가지의 형태에 대하여 고안중이다.

#### 4.1.2 OLED 디스플레이 화면 구성

소리인식 특징 추출 후 분류가 끝난 뒤에 각각의 소리에 대한 알림이 필요하다. 위험소리의 경우 해당 소리의 종류에 대한 정보가 텍스트로 보여지고, 사람의 음성인 경우 그대로 텍스트로 변환되어 디스플레이에 출력된다.. 이 때 시야에 방해를 주지 않고, 사용자의 눈이 편한 상태에서 인지 할 수 있도록 텍스트를 배치해야 할 적합한 위치를 테스트를 통해 최적화 할 예정이다.

#### 4.1.3 실시간 소리 분류


현재 정제된 위험소리데이터에 대해서는 일정 수준이상의 분류 성능을 내는 분류기를 구현하였으나, 실시간으로 들어오는 위험소리에 대한 분류에 대한 구현은 진행되지 않았다. 잡음과 함께 들어오는 소리에 대한 일정 수준이상의 분류 성능을 보장하기 위해 소리를 자르는 구간의 길이, 적절한 전처리, 잡음과 함께 녹음된 데이터를 추가로 분류기에 학습시키는 등 다양한 방법의로의 최적화 및 코드 구현이 요구되는 것으로 보여진다.

#### 4.1.4 소리 크기 및 방향 인식

4.1.3에서 분류되는 소리 이외에도 소리의 크기 및 방향을 통해 사용자가 판단을 할 수 있도록 돕는다. 현재 마이크 모듈로 소리가 인식이 되면 소리의 볼륨이 출력물로 나온다. 이후 양측에 마이크 2개를 이용하여 소리 방향을 감지하는 기능을 추가적으로 구현하는 것을 목표로 한다.

#### 4.1.5 구현된 소스 코드 병합

현재 팀원들이 각자 다른 개발환경에서 개발을 진행 중 이다. 따라서 실제 구현한 프로그램을 실행시킬 라즈베리파이 3B+ 모델이 배송되는 대로 개발환경을 확정 시키고 그동안 구현한 코드를 실행시킬 수 있도록 수정하여야 한다. 또한 동시에 처리되어야하는 기능들이 있어 멀티쓰레드를 돌리는 등 추가적인 작업이 요구된다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명		
	팀 명		
	Confidential Restricted	Version 1.3	2018-APR-09

## 5 고충 및 건의사항

- 졸업 프로젝트 진행 할 수 있는 마땅한 공간이 없어 불편을 겪고 있습니다. 특히 저희 프로젝트는 하드웨어를 요구하여 하드웨어를 보관할 장소를 찾는 것에서도 어려움을 느끼고 있습니다.