





2022 애저톤 (2022 Azure Virtual Hackathon)

- 1차 : 아이디어 기획안 서식 -

팀 명: TAVE

팀 원: 임예람, 이문기, 허주희, 박제윤

0. 애저톤 클라우드 스킬 챌린지 (CSC)

참가자 52명		모듈 완료
	임예람	21/21
	이문기	21/21
	Juhui Heo	21/21
	박제윤[대학원석사과정졸업 / 컴퓨터학과]	21/21

- 이번 CSC는 전체 52명의 참가자 중 9명만 완주하였음 ($9/52 = 17.3\%$, 2분야)

- TAVE 팀 4명 전원 완주

1. 아이디어 명칭

- Hi Azure!: Helping video creators who are visually-impaired edit videos using Azure Cloud AI!

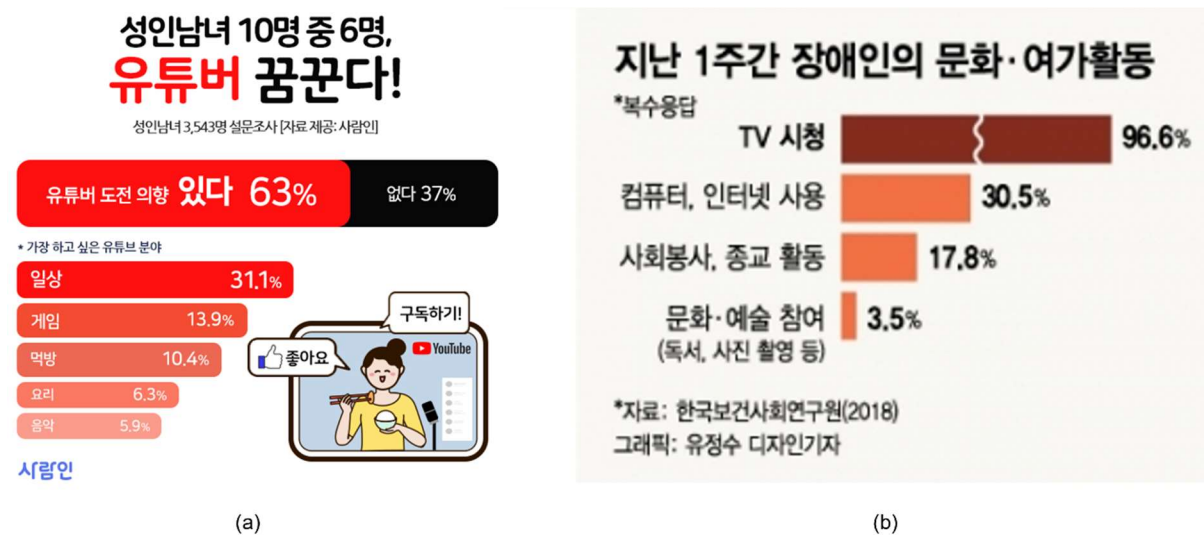
2. 핵심 내용

- 기존에는 시각 장애를 가진 영상 크리에이터들은 주로 청각정보에만 의존하여 영상 편집을 해왔음
- 하지만 조용한 산에서 풍경을 촬영하거나 시끄러운 노래를 틀고 춤을 추는 모습을 담은 영상 등, 청각 정보에 의존하기 어려운 영상들은 혼자 편집하기에 매우 어려울 수 있음
- 이러한 점은 시각 장애인이 크리에이터가 되는 것과 생산하는 영상 콘텐츠 장르에 큰 장벽이 될 수 있음
- 이번 프로젝트에서, 우리는 Azure Cloud AI로 영상의 visual feature를 추출한 뒤 음성으로 변환하여 frame-level description으로 제공하므로써, 시각 장애 영상 크리에이터들이 영상의 시각 정보도 같이 활용할 수

있는 서비스 를 제안하려함

3. 제안배경 및 목적

(1) 제안 배경1: 통계 조사



- (a) 시각장애가 없는 사람들은 관심사에 대해 영상 크리에이터 활동을 자유롭게 꿈꿀 수 있고 문화 활동으로써 크리에이터 활동을 하는데에 큰 제약이 없음
- (b) 하지만, 시각장애인들은 취미의 절대 다수는 tv 시청등의 정적인 활동이고, 이는 자신이 가진 끼와 재능을 더 다채롭게 펼치기에는 많은 현실적인 제약이 있다고 볼 수 있음

(2) 제안배경 2: 실제 시각장애 크리에이터들의 인터뷰



“편집은 혼자하기 어려워 친구가 도와줍니다. 누군가 편집을 어떻게 하냐고 물어보면 ‘좋은 친구를 사귀면 된다.’라고 말하죠. 그 친구에게 고마운 점이 많아요.”

영상 크리에이터 ‘원샷한술’ 한국시각장애인복지관 인터뷰 中



“아무래도 시각장애가 있어서 혼자 컴퓨터나 이런 걸로 편집프로그램을 사용하기는 어려움이 있어요. 그래서 저는 도와주는 PD 언니한테 스크립트를 전달해서 영상을 만들고 있습니다.”

영상 크리에이터 ‘우렁의 유디오’ TBS 인터뷰 中



편집은 어떻게 찍혔는지 기억해가며 진행한다.
이렇다 보니 6 ~ 7분가량의 영상을 만드는 데 3일이 넘는 시간이 걸릴 때도 있다.

영상 크리에이터 ‘브레드박’ 한국일보 기사 中

- 기존의 영상 크리에이터들은 콘텐츠를 만들때 주변사람의 도움을 받거나 (원샷한술, 우렁의 유디오), 그렇지 못할경우 청각정보에만 의존하며 편집을 하고 있음 (브레드박)

(3) 제안 목적

- (1), (2)에서 볼 수있듯이, 자신의 관심사나 특기, 취미등을 영상으로 자유롭게 풀어내고 싶어도 진입장벽이 너무 높아서 **주변의 도움 없이는 청각정보에만 의존하여 콘텐츠를 제작할 수 밖에 없음**

- 따라서, 이번 프로젝트 Azure Cognitive Service를 통해, **영상의 청각정보 뿐만 아니라 시각정보도 활용하는 편집 도움 서비스를 제안**하여 시각장애 크리에이터 영상 편집 및 콘텐츠 제작에 대한 접근성을 보다 높이려고 함

4. 아이디어(기술) 세부 내용

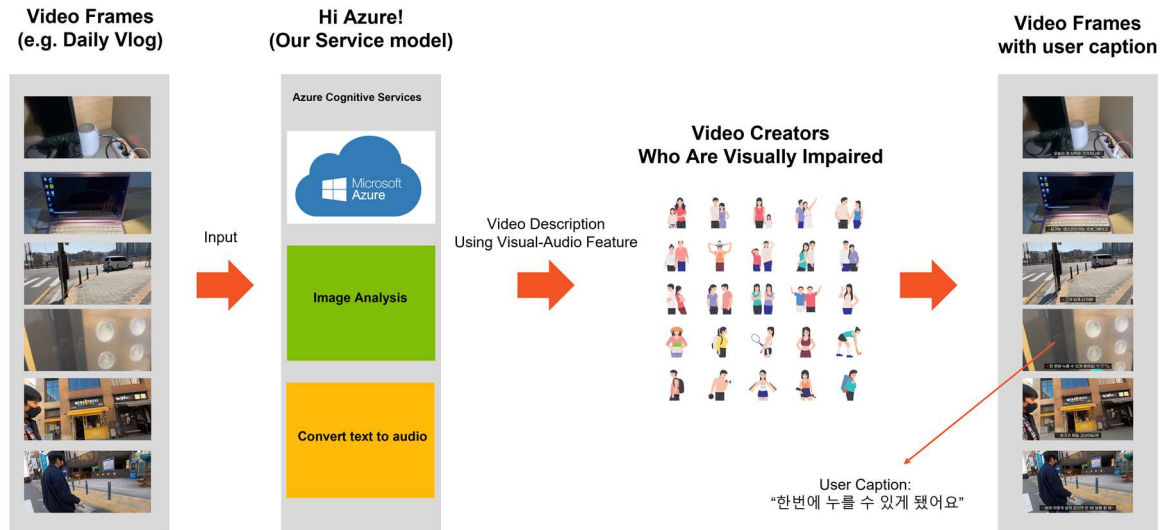
(1) 해결하고자 하는 AI for Good 테마

- Ai for Accessibility: For video creators who are visually impaired

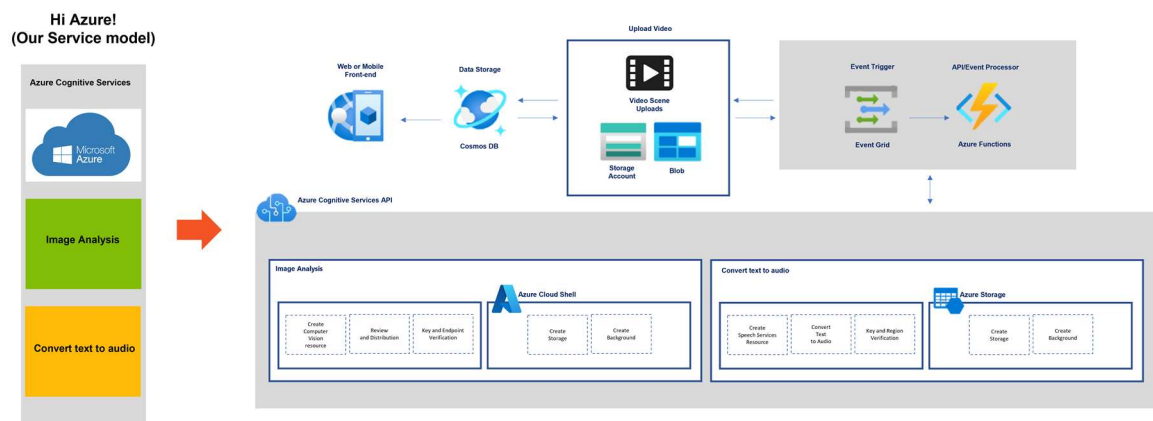


(2) 아이디어에 대한 설명

- 아래 그림의 Hi Azure! (제안한 서비스)는 사용자가 촬영한 영상을 입력으로 사용함
- 영상을 프레임 단위로 나누고 (e.g. 30 fps) 나눠진 프레임 단위마다 Azure를 활용한 Image Analysis module로 영상을 분석함
- 분석 결과 Azure를 사용한 Convert Text to Audio module로 음성으로 변환함
- 사용자는 영상 분석 결과를 소리로 듣고 원하는 부분에 자막을 다는등의 창작활동을 진행할 수 있음



(3) Azure AI를 활용한 아키텍처



- Azure Cognitive Services API를 활용
- Image Analysis
- Convert Text to Audio

(4) 사용하는 데이터셋에 대한 설명 및 활용 방안

- TartanAir dataset (Azure open dataset): 연속프레임 데이터셋, 모델 성능 분석 및 검증에 활용
- Youcook2: 비디오-캡션 데이터셋, 모델 성능 분석 및 검증에 활용

(5) Microsoft Azure 활용 방안

- Cloud shell 가상 공간에 코드와 이미지가 있는 깃 레포지토리를 clone 해와서 사용함
- Git repository안에 azure Cognitive Services에서 제공하는 코드를 활용
- Key값, Endpoint, Running code 세가지를 불러와서 사용

```

PS /home/ramye/ai-900> ./analyze-image.ps1
Analyzing image...

Description:
a woman holding a shopping cart in a grocery store

Objects in this image:
- person

Tags relevant to this image:
- text
- person
- woman
- marketplace
- shop

Synthesizing speech...
Response saved in output2.wav

```

5. 기대효과

(1) 시각장애인의 사회 참여 및 삶의 질 향상

기 준

- ① 취미로서 오디오 청취, 점자책 독서, 사색, 명상 등 혼자만의 단방향 취미 영위
- ② 외부활동이 극히 적어, 인간 관계에 부족함을 느낌
- ③ 주변의 시선에 괜히 움츠러듬

Hi Azure!
→
쉬운 영상 제작

적용 후

- ① 본인의 이야기, 전문지식을 담은 영상을 외부와 공유, 소통하여 독립, 사회 참여율 상승
- ② 영상 작업을 통해 자연스럽게 외부로 신체적 확장 가능 및 경험할 수 있는 공간 증가



시각장애인 크리에이터 브래드박 주민센터 강연활동

출처 : 장애인 유튜브 크리에이터들의 영상 제작 활동과 정동체험에 대한 사례 연구 (한국언론학보, 2022년)

(2) 일반인의 장애인에 대한 인식 변화 유도

- 최근 특히 타 매체 대비 사용율이 높은 영상매체에서의 장애인의 삶에 대한 자연스러운 노출 상승으로 장애인에 대한 배려의식 및 함께 사는 삶에 대한 인식 변화 유도



<유튜브 영상 '시각장애인이 화장을 해?'>

(3) 다른 후천적 시각장애인의 적응을 도와주는 역할 수행

- 시각장애인 관련 영상이 많아짐으로서, 갑작스럽게 생긴 시각장애를 겪는 분들에게 사회 적응, 일반 생활 영위에 도움을 줄 수 있도록 함



저는 18살 때 갑자기 시력을 잃게 되었어요

<유튜브 영상 '갑자기 시각장애인이 된 나의 이야기'>

(4) 'Microsoft clipchamp' 등 Microsoft 영상 편집 프로그램 적용 가능성

- 현재 널리 쓰이고 있는 전문가용 영상 편집 프로그램 'clipchamp' 등에 보편적 접근성 강화를 위한 기능으로서 사용 가능