

3주차 과제

제출기한: 2020년 4월 9일 (목) 23:59:59

팀 번호: 13

1. 프로젝트 주제

팀원들의 2주차 과제를 바탕으로, 팀의 최종 프로젝트 주제를 선정한다. 진행되는 프로젝트는 Github에 지속적으로 코드를 업로드하여 지도교수 및 조교가 진행상황 및 팀원의 기여도를 확인할 수 있도록 한다. 따라서 모든 팀원이 Github 계정을 만들고 팀원 중 한 명(팀장)이 저장소(repository)를 프로젝트 이름으로 만든 뒤, collaborator로 나머지 팀원을 추가하도록 한다. (파일 마지막 장 참고자료 참고) 또는 프로젝트 이름의 계정을 만들어서 그 안에 저장소를 만들고, 해당 저장소에 모든 팀원을 collaborator로 추가해도 됨.

프로젝트 주제: SoW;Sound of Writing

프로젝트 Github 저장소 주소: <https://github.com/KyuhyoJeon/SoW>

팀원1(팀장)	이름: 전규효	Github ID: KyuhyoJeon
팀원2	이름: 이수연	Github ID: dltndus5244
팀원3	이름: 이지현	Github ID: nuri0311
팀원4	이름:	Github ID:

2. 프로젝트 계획

확정된 프로젝트 주제를 구현하기 위한 계획을 작성한다. 아래 항목은 글씨 크기 9인 자유 포맷으로 작성 하되, 분량은 전체 2장을 넘어야 함. 2주차 과제에서 작성한 내용을 활용 가능함.

- **필요성** : 해당 프로젝트를 활용할 수 있는 상황 및 대상 등 실제 사례를 바탕으로 한 필요성 서술.
- **독창성** : 비슷한 주제로 기존에 개발된 프로그램 등을 조사하여 작성하고, 본 프로젝트에서 개발할 내용은 어떤 점이 다른지 서술.
- **필요 기술** : (1) 팀원들이 사용 가능한 프로그래밍 언어에 기반하여, 본 프로젝트에 사용할 프로그래밍 언어를 1개 이상 작성. (2) 프로젝트에 활용 가능한 오픈소스 라이브러리를 조사하고 본 프로젝트의 어떤 부분에 사용할 수 있는지에 대한 설명과 출처(링크 등)를 작성.
- **월별 계획 설정** : 4월~10월에 대해 월별 계획을 기술. 참고: 6월 중순에는 이번 학기 중간 결과 발표 예정이며, 최근 4년간 캡스톤디자인 전시회는 11월 7-14일 경 진행되었음(올해 일정은 미정). 따라서 최종적으로 10월까지의 마무리되는 것을 목표로 할 것.
- **주차별 계획 설정** : 남은 학기 (4~6월)에 대해 주차별 계획을 기술하고, 각 팀원의 예상 역할을 나눠볼 것. 추후 팀원 간 평가도 성적에 반영될 예정이니 각자 작업 분배와 참여를 확실히 할 것.

1) 프로젝트 내용

SoW는 시각장애인이거나 노인, 어린아이와 같은 시각적인 불편함을 가진 사람들을 대상으로 책 등의 출판

및 문서화된 문자열 집합을 대상으로 촬영한 뒤 이미지로부터 텍스트를 추출하고 음성으로 변환하여 스피커로 출력하는 단일 디바이스를 제작하는 것을 목표로 하는 프로젝트다.

2) 필요성

2018년 기준 한국의 전국 등록 장애인은 약 전국민의 5%, 그 중에서 시각장애인은 10%가 있다. 정확한 수치로 나타내면 252,957명의 시각장애인이 등록 되어있다고 보건복지부의 통계자료에 나타난다. 그런 시각장애인을 위한 점자형식의 문서들은 다양성 측면에서 많이 부족하다. 비장애인은 읽고 싶은 책이 생기면 서점에 가서 읽거나 구매할 수 있지만, 시각장애인은 종이 점자책으로 출판된 책을 구입하거나, 파일 제작을 복지관이나 점자 도서관에 신청해야 한다. 문제는 점자책으로 출간되는 책이 전체 출판도서의 5%도 안된다는 것이다. 생산도 쉽지 않고 원하는 분야의 문서를 찾기도 어려울 뿐만 아니라 찾더라도 점자에도 다양한 언어가 있어서 한글 점자로 되어 있을 거라는 보장도 없다. 또한 파일 제작을 신청한다 하더라도 짧게는 3개월, 길게는 1년이 걸린다는 단점이 있다. 시각장애인들을 위한 점자 도서관이 있지만 2017년 기준 우리나라 전국의 시각장애인 수가 25만 3천명에 이르는 데에도 불구하고 전국의 시각장애인 도서관은 39개 밖에 되지 않는다. 또한 점자 도서를 제작하는 점역교정사가 턱없이 부족한 실정이라 시각장애인들은 점자책의 종류가 부족한 것에 대해 큰 불편함을 느끼고 있다. 심지어 시각장애인의 약 90%가 선천적 원인이 아닌 후천적 원인으로 인해 시력을 잃은 경우라고 한다. 이러한 경우 점자방식의 독서방법을 새로 익히는 것을 포기하게 되는 경우가 많다고 한다.

이러한 상황에서 일반 문자로 된 문서나 책을 다른 사람의 도움을 받아 읽을 수 있다면 더할 나위 없겠지만, 그 또한 만만치 않은 상황에서 SoW는 책이나 문서를 음성으로 읽어주는 기능을 제공한다. 이런 기능은 시각장애인 뿐만 아니라 독서에 취미를 가지고 있지만 시력이 점차 저하되는 노인들에게도 큰 도움이 될 수 있다. 그리고 말은 알아도 문자를 읽을 줄 모르는 어린 아이들을 대상으로도 큰 도움이 되고, 아이들의 부모님을 대상으로 아이들에게 책을 읽어주는 대신 본 프로젝트의 제품을 사용함으로 개인의 휴식시간을 더 가질 수 있다.

3) 독창성

현재 미국 M.I.T에서 개발중인 "ii핑거리더(FingerReader)"라는 제품이 본 주제와 아주 비슷하다. "핑거리더" 제품의 경우 반지 형태의 디바이스로, 손가락에 끼워 손으로 가리키며 문자열을 따라 이동하면 음성으로 문자를 읽어주는 기능을 가진다. 사용방식은 점자를 읽는 방식과 매우 흡사하다. 하지만 현재 문자 읽을 때 손으로 문자열을 따라 이동해야 인식하는 특성상 터치스크린에서의 사용에 문제점을 가지고 아직 개발 중이다. "핑거리더"는 휴대의 편리함에 있어서 앞서지만, 한치 앞도 볼 수 없는 경우에 촉감을 이용해서 읽는 점자방식과는 달리 주 사용자가 눈이 불편한 사람들이 대상인 만큼, 인쇄된 문자들을 따라 손가락을 가리키는 데는 한계가 있을 것으로 예상된다.

그리고 탐투스의 복리더 "iiiMagic VT-500POT"라는 제품도 있는데 책의 페이지를 스캔하여 이미지를 얻어낸 다음 텍스트를 추출하고 추출된 텍스트를 음성으로 변환하여 스피커로 출력하는 방식으로 책을 스캐너에 올리면 스캐너로 펼쳐진 페이지를 스캔한 뒤 이미지의 텍스트들을 OCR을 이용해 추출하고 TTS를 이용해 음성으로 변환한 뒤 출력하는 디바이스로 본 주제와 제일 흡사한 제품이다. 하지만 스캐너와 같은 설치형 디바이스로 이동과 휴대가 불편하다는 단점과, 제품의 가격이 무척 비싸다.

다음으로 휴대형 스캐너 형식의 "ivVoice Stick"라는 아이디어제품은 휴대형 스캐너 형식으로 책을 스틱 형태의 스캐너로 훑어주면 책의 텍스트를 인식해서 음성으로 출력해주는 제품이다. 다만 이 제품은 손으로 직

접 책을 훑어서 스캔해야 하는데 이 또한 주 사용자가 눈이 불편한 사람들이 대상인 만큼 손으로 직접 문자들을 스캔하는데 한계가 있을 것으로 예상된다.

또한 “책마루2”라는 제품은 시각장애인과 난독증을 겪고 있는 독서 장애인을 위해 음성독서, 인쇄물 글자 판독 등을 제공하는 멀티플레이어다. OCR기능을 이용하여 빠르고 정확한 문자 인식, 연속 촬영을 제공하고, OCR 기능으로 판독된 문자를 문단, 문장, 줄, 단어 등 이동 단위를 선택하여 읽기를 지원한다. 내장 스테레오 스피커와 헤드폰 잭을 통하여 음성을 출력하는 기능도 제공한다. 또한 FM 라디오, 알람, 계산기 등 다양한 무가 기능도 제공하는 멀티플레이어이다. 하지만 80만원 이상의 고가 제품이고, 독서 보조기기 또는 별도의 조작장치에 대한 보급이 미미한 상황이며, 60세 이상 시각장애 그룹에서는 사용방법에서 어려움을 느끼고 있다고 한다.

본 프로젝트인 SoW는 위의 세 제품과 OCR을 활용해 이미지에서 텍스트를 추출, TTS를 활용해 텍스트를 음성으로 변환하여 스피커로 출력하는 공통점이 있지만 다음과 같은 차이들을 가진다.

- 쉽게 이동 가능한 독서대 형태의 디바이스로 이동과 설치가 편리하다.
- 쉬운 설치와 더불어 간단한 작동 방식으로 남녀노소 누구나 사용하기 쉽게 제작한다.
- 고정된 각도의 이미지인식 센서(렌즈, 카메라 등)를 이용해 페이지 인식으로 눈이 불편한 사용자들이 글을 찾지 않아도 디바이스 자체적으로 글자를 페이지에서 인식하여 편리하다.
- 하드웨어 기반이 아닌 소프트웨어 중심으로 저렴한 가격으로 구현한다.

4) 필요 기술

(1) 프로그래밍 언어: Python

(2) 활용 가능한 오픈소스 라이브러리 설명 및 출처

- ^vOCR오픈소스, OCR은 광학식 문자 판독으로 이미지화 되어있는 문자들을 실제 2진 데이터인 텍스트로 추출한다.

GitHub: “<https://github.com/kba/awesome-ocr>”

- ^{vi}TTS오픈소스, TTS는 Text-to-Speech의 약자로, 텍스트를 음성으로 변환한다.

GitHub: “<https://github.com/search?q=TTS&type=Repositories>”

- ^{vii}Tesseract: 태서렉트는 다양한 운영체제에서 사용할 수 있는 광학 문자 인식 엔진이다. 휴렛팩커드가 1980년대에 개발했으며 2005년 오픈소스로 개방되었다. 또한 통용되는 OCR 오픈소스 중 가장 정확도가 높다고 알려져 있고, 딥러닝 학습 기능이 내장되어 있는 강력한 OCR 도구이다. 이번 프로젝트에서 이미지로부터 텍스트를 추출하기 위한 부분에서 사용할 수 있다.

링크: <https://joyhong.tistory.com/79>

- ^{viii}Balabolka: 발라볼카는 MS에서 제공하는 TTS의 모듈을 사용한 프로그램이다. 화면 상의 텍스트를 WAV, MP3, MP4, OGG 또는 WMA 파일로 저장할 수 있다. 발라볼카는 무료 프로그램이며 조작법이 간단하고 텍스트 파일이 길어도 음성으로 변환이 가능하다는 장점이 있다. 발라볼카는 이번 프로젝트에서 OCR 기능으로 추출한 텍스트를 음성으로 변환하는 부분에서 사용할 수 있다.

링크: <https://needjarvis.tistory.com/138>

- ixVision API는 이미지에서 텍스트를 감지하고 추출할 수 있다.

링크: <https://cloud.google.com/vision/docs/ocr>

- xClova Speech Synthesis(CSS) - 입력한 텍스트를 자연스러운 목소리로 재생해주는 음성 합성 API다.

링크: <https://www.ncloud.com/product/aiService/css>

- xiCLOUD TEST-TO-SPEECH - 고음질 음성 합성 Google Cloud Text-to-Speech는 텍스트를 30여 개의 언어 및 방언이 지원되는 180여 개가 넘는 자연스러운 음성으로 변환해준다. 음성 합성(WaveNet)의 획기적인 연구 성과와 Google의 강력한 신경망을 적용하여 높은 품질의 음질을 제공한다. 사용이 간편한 API로 사용자와 실제 대화하는 듯한 상호작용이 가능해 고객 서비스, 기기 상호작용, 기타 애플리케이션의 혁신적인 변화가 가능해진다.

링크: <https://cloud.google.com/text-to-speech>

이 외에도 라즈베리파이와 각종 센서를 이용한 하드웨어 구축기술.

5) 월별 계획 설정

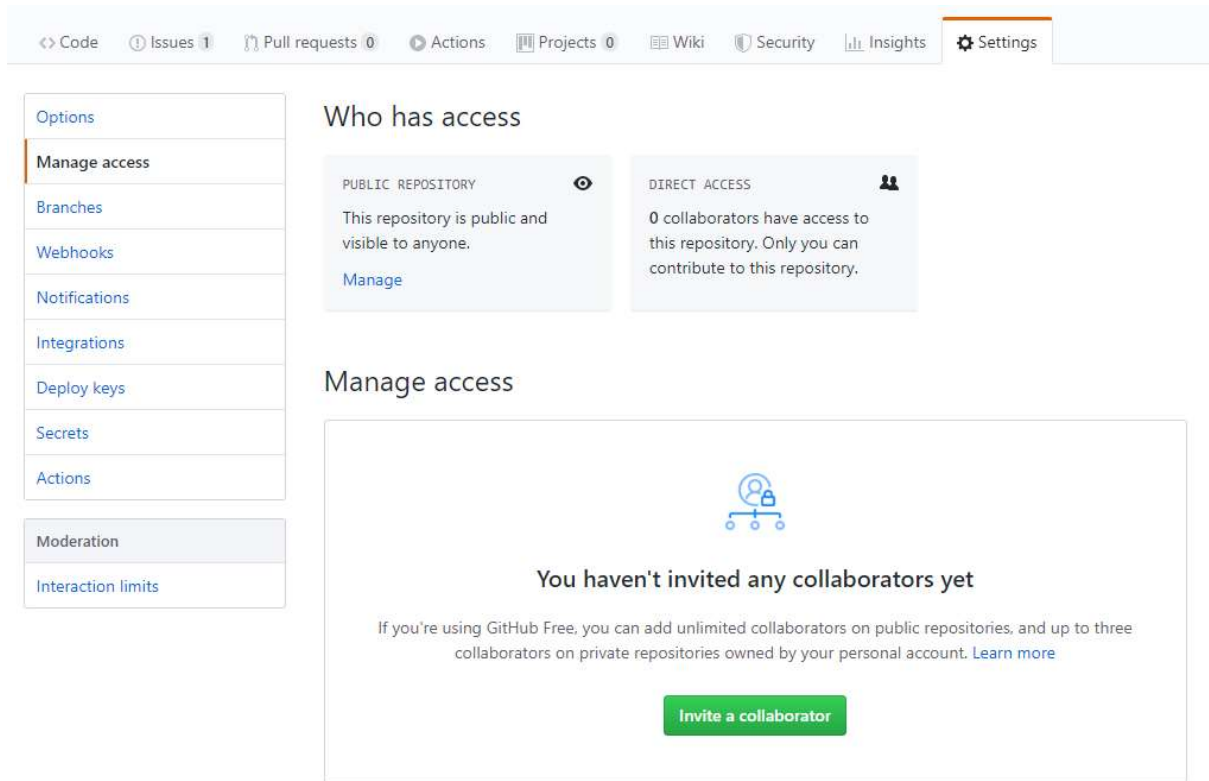
월	계획
4월	명확한 주제 설정 및 구체적인 기능 정리
5월	주요기능 구현 기술(프로그래밍 언어/오픈소스 라이브러리) 학습
6월	개발환경 구축 및 개발 시작, 중간발표 준비
7월	DB 및 SW구축, HW설계 및 구축에 필요한 필요부품 파악 및 준비
8월	시제품 완성
9월	제품 테스트 및 보완, 최적화
10월	제품 완성 및 개요준비

6) 주차별 목표 설정

주차	주차별 계획 및 담당자
04.13-04.19	- 기존제품 분석 및 기능 아이디어 조사(개인별 조사/정리 담당자: 전규효) - 소비자 선호 기능 조사 및 설문(다 함께 토의 후 설문문항 결정 후 개인별 조사)
04.20-04.26	- 1주차 설문 취합 및 미흡 설문 문항 개선(다 함께 토의 후 개선/정리 담당자: 이지현)
04.27-05.03	- 주제의 명확한 컨셉 결정(다 함께 토의 후 결정/정리 담당자: 이지현)
05.04-05.10	- Python 언어 학습(개인별 학습, 기초 예제 풀이 결과 제시 GitHub 개인 폴더에 게시)
05.11-05.17	- OCR 오픈소스 기능 학습(개인별 학습, 프로그래밍 결과 GitHub 개인 폴더에 게시)
05.18-05.24	- TTS 오픈소스 기능 학습(개인별 학습, 프로그래밍 결과 GitHub 개인 폴더에 게시)
05.25-05.31	- 프로젝트 구현 시 사용할 오픈소스 선정(다 함께 토의 후 결정) - 선정된 오픈소스 라이브러리 수집 및 사용방법 정리(담당자: 이수연)
06.01-06.07	- 학습결과를 바탕으로 메인 개발자 선정 및 개발환경 구축 및 시작(담당자: 메인 개발자)
06.08-06.14	- 개발 경과 취합 및 중간발표 준비(담당자: 전규효)
06.19, 06.26	프로젝트 중간결과 발표

<참고자료>

1. Sourcetree를 사용하는 Github 튜토리얼 : <https://milooy.wordpress.com/2017/06/21/working-together-with-github-tutorial/>
2. 저장소에서 collaborator 추가 메뉴 : Settings > Manage access > Invite a collaborator



3. Git/Github의 개념 및 Command line 환경에서의 Git/Github 사용 : e-campus의 3주차 파일 참고

[참고자료]

- i "장애인 현황", e-나라지표, 2019년 04월 25일 수정, 2020년 03월 27일 접속,
"http://www.index.go.kr/potal/main/EachDtlPageDetail.do?idx_cd=2768"
- ii "책 읽어주는 반지, 핑거 리더", Early Adopter, 2014년 07월 16일 수정, 2020, 03월 27일 접속,
"http://www.earlyadopter.co.kr/3080"
- iii "탐투스 북리더 Magic VT-500POT", TAMTUS 교육기자재 전문 탐투스 쇼핑몰, 2020, 03월 27일 접속,
"http://www.tamtus.co.kr/shop/item.php?it_id=1390200959"
- iv "Voice Stick", YD; YANKO DESIGN, 2008년 08월 21일 수정, 2020년 03월 27일 접속,
"https://www.yankodesign.com/2008/08/21/voice-stick/"
- v "Awesome OCR", GitHub, 2020년 03월 13일 수정, 2020년 03월 27일 접속,

"<https://github.com/kba/awesome-ocr>"

vi "mozilla / TTS", GitHub, 2020년 03월 12일 수정, 2020년 03월 27일 접속, "<https://github.com/mozilla/TTS>"

vii "Tesseract로 OCR 하기", joyHong, tistory, 2019년 11월 14일 수정, 2020년 04월 05일 접속, "<https://joyhong.tistory.com/79>"

viii "무료 TTS(TEXT TO SPEECH) 프로그램 추천, BALABOLKA", 자비스가필요해, tistory, 2016년 10월 10일 수정, 2020년 04월 05일 접속, "<https://needjarvis.tistory.com/138>"

ix "Vision API, 이미지의 텍스트 감지", Google cloud AI, Google cloud, 2020년 02월 24일 수정, 2020년 04월 05일 접속, "<https://cloud.google.com/vision/docs/ocr>"

x "Clova Speech Synthesis(CSS)", Naver Cloud platform, 2020년 04월 05일 접속, "<https://www.ncloud.com/product/aiService/css>"

xi "CLOUD TEXT-TO-SPEECH 머신러닝 기반의 텍스트 음성 변환", Google cloud AI, Google cloud, 2020년 04월 05일 접속, "<https://cloud.google.com/text-to-speech>"