

캡스톤 디자인 I

종합설계 프로젝트


프로젝트 명	수화 人(in) 플레이어
팀 명	전지적 수화 시점 팀
문서 제목	중간보고서

Version	1.3
Date	2019-04-18

팀원	윤지은 (조장)
	정창희
	이희지
	배진영
	축이림
지도교수	임성수 교수님

CONFIDENTIALITY/SECURITY WARNING


이 문서에 포함되어 있는 정보는 국민대학교 전자정보통신대학 컴퓨터공학부 및 컴퓨터공학부 개설 교과목 캡스톤 디자인I 수강 학생 중 프로젝트 “수화 人(in) 플레이어”를 수행하는 팀 “전지적 수화 시점”의 팀원들의 자산입니다. 국민대학교 컴퓨터공학부 및 팀 “전지적 수화 시점”의 팀원들의 서면 허락없이 사용되거나, 재가공 될 수 없습니다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	수화 인 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 1.3	2019-APR-18

문서 정보 / 수정 내역


Filename	중간보고서-수화인플레이어.doc
원안작성자	윤지은, 정창희, 이희지, 배진영, 축이림
수정작업자	윤지은, 정창희, 이희지, 배진영, 축이림

수정날짜	대표수정자	Revision	추가/수정 항목	내 용
2019-04-15	이희지	1.0	최초 작성	최초 작성
2019-04-16	윤지은	1.1	내용 추가	수정된 연구내용 추가
2019-04-17	배진영	1.2	내용 추가	향후 추진 계획 수정
2019-04-17	전체	1.3	내용 추가	전체적인 수정

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	수화 인 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 1.3	2019-APR-18

목 차

1	프로젝트 목표	4
2	수행 내용 및 중간결과	5
2.1	계획서 상의 연구내용	
2.1.1	단어 및 수화 Data Set 구축	
2.1.1.1	Crawling	
2.1.1.2	MongoDB	
2.1.2	동영상의 음성에서 자막파일 추출	
2.1.3	텍스트 수화 번역 - 형태소 분석	
2.1.4	수화 Data Mapping	
2.2	수행내용	
2.2.1	단어 및 수화 Dataset 구축	
2.2.1.1	크롤링	
2.2.2	동영상의 음성에서 자막파일 추출	
2.2.2.1	음성인식(STT:Speech to Text)	
2.2.2.2	자막 생성	
2.2.3	자막 수화 번역 - 형태소 분석	
2.2.4	수화 번역 영상 생성	
3	수정된 연구내용 및 추진 방향	12
3.1	수정사항	
3.1.1	OpenPose 제외	
3.1.2	Backend 개발 환경 변동	
3.1.3	Android Studio 앱에서 web 서비스로 변동	
3.1.4	형태소 분석 언어를 R언어에서 Python으로 대체	
4	향후 추진계획	13
4.1	향후 계획의 세부 내용	
4.1.1	선택적 서비스	
4.1.2	URL을 이용한 업로드 기능 추가	
4.1.3	단어 유사도 검사	
4.1.4	전문가 자문	
5	고충 및 건의사항	14

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	수화 인 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 1.3	2019-APR-18

1 프로젝트 목표

본 프로젝트의 이름은 ‘수화인플레이어’로 사용자가 정보를 얻고자 하는 영상을 등록하면 영상의 내용을 수화로 번역하여 이를 등록된 영상과 함께 사용자에게 제공하는 것이 목표이다.


현재 시중에 나와있는 대부분의 영상매체는 자막이 함께 제공되는 영상이 적을 뿐더러 더욱이 수화 영상을 함께 제공하는 것은 뉴스나 일부 수화통역 영상 이외에 거의 찾아볼 수 없다. IT 기술 발달로 점점 영상 매체의 의존도가 높아짐에 따라 청인과 청각장애인 간의 정보 격차는 심화될 것으로 예상된다. 하여 본 프로젝트의 기대효과는 청각장애인도 인터넷 강의나 시사/교양 프로그램 등 정보를 담고 있는 영상에서 원하는 정보를 제약없이 습득할 수 있도록 하는 것이다.

목표 달성을 위한 기술적 단계는 크게 4가지로 나뉜다.

- 1) 단어 및 수화 DataSet 구축
- 2) 동영상의 음성에서 자막 파일 생성
- 3) 한국어 문법을 수화 문법 체계 형식으로 변형
- 4) 추출한 형태소와 수화 Data Mapping방식 구현

본 프로젝트의 주된 목적은 영상을 자막 및 수화로 번역하는 시스템 자체의 설계 및 구현이며, 모든 문장을 정확히 해석하는 것보다 전체적인 영상의 내용을 파악할 수 있을 정도의 효과를 목표로 한다.

따라서 본 프로젝트의 전제 조건은 다음과 같다. STT나 형태소 분석기의 오인식으로 발생하는 몇몇의 통역 오류는 현실적인 프로젝트 진행 과정을 고려하여 해결하지 않는다. 추후의 이 기술들이 개선되었을 때 자연스럽게 해결 될 문제라고 판단한다. 또한 한국어의 중의적 표현은 고려하지 않고 직역한다. 중의적 표현의 특징은 해학이나 풍자 등에 활용되며 의미의 다양성으로 문학 작품의 예술성을 높이고 싶을 때 사용한다. 하지만 중의적 표현을 전달하지 못하는 것이 전체 영상의 내용을 이해하는데 큰 장애가 되지 않기 때문에 중의적 표현의 처리 기능은 구현하지 않기로 한다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	수화 인 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 1.3	2019-APR-18

2 수행 내용 및 중간결과

2.1 계획서 상의 연구내용

2.1.1 단어 및 수화 Data Set 구축

2.1.1.1 Crawling

크롤링은 웹 페이지를 그대로 가져와 해당 페이지 내의 데이터를 추출해 내는 행위를 말한다. 본 프로젝트에서는 단어마다 단어명, 품사, 의미, 수화 영상 Data를 수집하기 위해 사용한다. 국립국어원한글수어사전 홈페이지

(<http://sldict.korean.go.kr/front/main/main.do>)의 수화 영상을 크롤링한다. 사용되는 파이썬 라이브러리로 다음과 같다.

1. Requests - HTML을 읽어오는 과정을 수행
2. BeautifulSoup - 웹 데이터 크롤링 또는 스크래핑을 할 때 사용하는 Python 라이브러리로 HTML에서 원하는 데이터를 손쉽게 parsing할 수 있도록 한다.
3. re (regular expression) - 웹페이지에서 크롤링을 할 때 불필요한 문자를 제거하기 위해 정규표현식 라이브러리를 사용한다.
4. cv2 (openCV) - 수어사전 홈페이지에서 제공하는 동영상의 경우 1배속과 0.5배속이 더해져 있다. 따라서 openCV로 동영상의 (전체 프레임 수) / 3의 프레임 수를 구해 수화 영상data를 재가공한다.

2.1.1.2 MongoDB


대용량 빅데이터를 처리하기 용이한 NoSQL인 MongoDB를 사용한다. DB는 수어사전 홈페이지에서 제공하는 단어를 저장하는 Dictionary, 사용자가 업로드한 Video 및 음성파일, 자막파일을 저장하는 Video로 나뉜다.

Dictionary

- 수어사전 홈페이지에서 제공하는 단어들 중 일상생활 수어 약 14,000개를 저장한다.
- Collection으로는 일상생활 단어가 담겨 있는 Basic Collection, 숫자 단어가 담겨있는 Number Collection, 지화 단어가 담겨있는 FingerLanguage Collection이 있다.
- 각 Collection의 Documents에는 단어명을 나타내는 word_name, 품사정보가 담겨있는 part, 수화동영상의 url을 갖는 video_url, 프레임 수 정보를 가지고 있는 frame Fields가 있다.

Video

- Collection으로는 사용자가 upload한 동영상의 정보를 저장하는 Upload_Video가 있다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	수화 인 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 1.3	2019-APR-18

- 사용자가 업로드한 동영상에서 음성 추출, 자막 생성을 하기 위해 불러오며 재가공하게 된다.

2.1.2 동영상의 음성에서 자막파일 추출


사용자가 등록한 영상의 음성을 텍스트 파일로 변환하기 위해 Google STT API를 사용한다. 카카오나 네이버 같은 우리나라의 소프트웨어 기업들도 개발자들은 위한 음성 API를 제공하고 있지만 자막 제작을 위한 시간 정보를 제공하지 않는다. 따라서 본 프로젝트는 time stamp 기능을 제공하는 Google Cloud Speech-to-Text를 사용한다. 동영상과 해당 자막의 싱크를 맞추기 위하여 구글의 타임 스탬프 기능을 활용한다. 한 문장의 시작점과 끝점의 시간 정보를 가져와 자막 파일에서 사용하는 시간정보로 변환한 뒤 자막파일에 적어주고 그 밑에는 음성인식의 결과를 적어준다. 파일의 확장자는 웹자막 확장자인 .vtt로 저장해준다. 또한 한 문장이 너무 길 경우 자막으로 출력했을 때 가독성이 떨어지므로 적절한 부분에서 잘라주는 과정이 필요하다. 이 과정은 한 문장에 포함되는 단어가 특정 개수를 넘어가면 문장을 분할하는 것으로 구현한다.

2.1.3 텍스트 수화 번역 - 형태소 분석

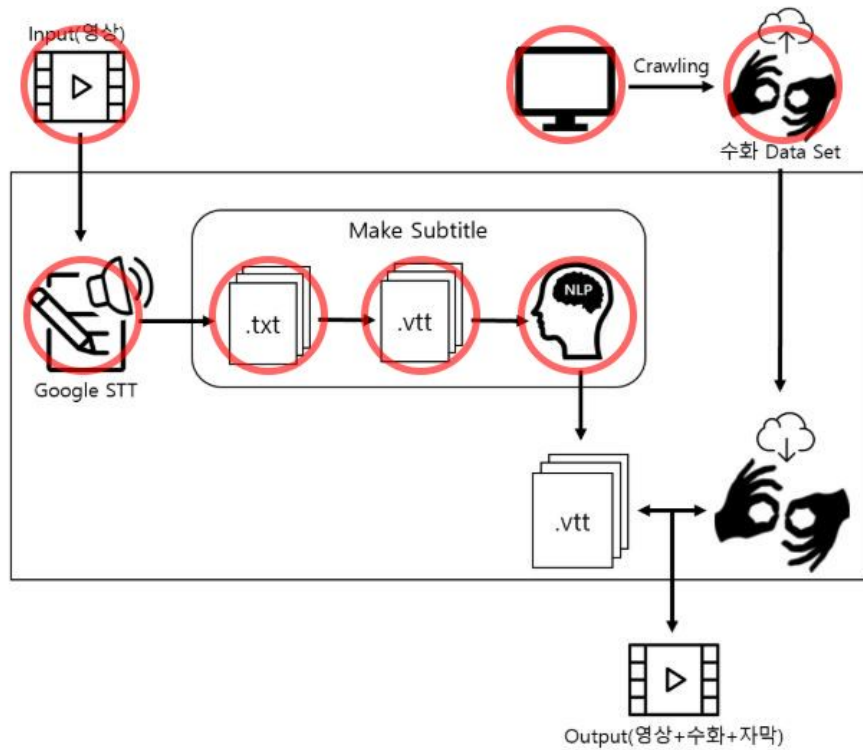
한국어와 외국어의 문법체계가 다르듯이 수화도 하나의 독립된 언어이기 때문에 그 자체만의 문법 체계를 가지고 있다. 한국어를 수화로 번역하기 위해 필요한 단계는 총 3가지 이다. 먼저 한국어를 형태소 분석 패키지 Konlpy를 이용하여 문장의 형태소를 분석한다. 그 다음 조사를 거의 쓰지 않는 수화의 문법 구조를 고려하여 수화에서 사용하는 일부 조사를 제외하고 모두 제거한다. 또한 과거형을 표현하는 시제 선어말 어미(‘았’ 또는 ‘었’)는 ‘끝’이라는 단어로 바꾼 뒤 문장 끝에 배치시킨다. 의문문을 표현하기 위해선 ‘느니까’ 단어를 문장 끝에 배치하여 의문문을 표현한다. Konlpy 시스템 사전에 존재하지 않는 복합명사나 대명사는 사용자 사전에 추가하여 형태소 분석을 더욱 정확하게 만든다.

2.4 수화 Data Mapping

수화 문법 체계로 재구성한 형태소들에 적절한 수화를 매핑시켜주기 위해서 수화 데이터베이스에서 Django 쿼리문을 이용하여 적절한 수화동영상을 가져온다. 각 형태소에 맞는 수화 동영상을 가져와 이어붙인 뒤 동영상 자막의 싱크 정보와 수화통역 영상 길이 간의 상관관계를 계산하여 싱크를 맞춰준다. 만약 수화와 매핑을 할 때 동음이의어가 나올 경우를 대비하여 디비에 참조단어 속성을 추가적으로 저장한다. 문장의 동음이의어 앞 뒤의 단어를 가져오고 디비에 저장되어 있는 참조단어와 비교한다. 만약 참조단어에 비교 단어가 존재하면 그것을 선택하고 없다면 Python의 Gensim 패키지를 이용하여 단어를 벡터화 한 뒤 단어간 유사도를 계산하고 유사도가 더 높은 것으로 선택한다.


 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	수화 인 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 1.3	2019-APR-18

2.2 수행내용



<그림 1. 프로젝트 진행 상황>

<그림 1>은 현재 프로젝트 진행 상황을 나타낸 그림이다. 프로젝트 예상 시나리오에 진행 완료된 작업에 동그라미로 표시하였다. 구현해야하는 기술들이 많아 팀원마다 각자 파트를 분배하여 Agile한 방식으로 프로젝트를 진행하고 있다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	수화 인 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 1.3	2019-APR-18

2.2.1 단어 및 수화 Dataset 구축

2.2.1.1 크롤링

```
[[ '과거시제', 'none', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19583', '99' ]]
[[ '과거완료', 'none', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19584', '118' ]]
[[ '과거진행', 'none', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19585', '120' ]]
[[ '과거진행완료', 'none', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19586', '160' ]]
[[ '관계부사', 'none', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19587', '120' ]]
[[ '관계언', '명사', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19588', '100' ]]
[[ '관계적의미', 'none', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19589', '120' ]]
[[ '관용구', '명사', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19590', '160' ]]
[[ '관용어', '명사', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19591', '120' ]]
[[ '관용어법', 'none', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19592', '160' ]]
[[ '관형격', '명사', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19593', '120' ]]
[[ '관형격조사', 'none', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19594', '180' ]]
[[ '관형구', '명사', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19595', '140' ]]
[[ '관형사', '명사', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19596', '120' ]]
[[ '관형사형전성어미', 'none', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19597', '240' ]]
[[ '관형절', '명사', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19598', '120' ]]
[[ '관형절', '명사', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19599', '140' ]]
[[ '관형절을안은문장', 'none', 'http://sdict.korean.go.kr/front/sign/signContentsView.do?origin_no=19600', '200' ]]
```


<그림 2. 크롤링 결과 출력>

<그림2>에서 보는바와 같이 단어명, 품사, 수화영상의 url, 사용되어질 프레임 수를 한국수어사전 홈페이지에서 Crawling하여 Dictionary DataBase에 저장하였다.

2.2.2 동영상의 음성에서 자막파일 추출

2.2.2.1 음성인식(STT:Speech to Text)

Google Speech to Text API를 사용하여 로컬 파일에서 동기 음성 인식을 수행하도록 했다. 입력한 파일에서 소리를 추출하기 위해 영상 파일을 인코딩하여 주파수와 파일 형식을 변형시킨 뒤 STT를 적용한다. 추출한 데이터에 단어 타임스탬프를 가져와 텍스트로부터 변환되는 오디오의 시차 값을 기록한다. 문장의 구별을 위해 텍스트 변환 결과에 마침표 및 쉼표와 같은 구두점을 삽입했다. 추후에 파일을 GCS에 업로드한 뒤 영상 stt의 효율성을 높이하고자 한다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	수화 인 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 1.3	2019-APR-18

```

1  우리 start_time 0.4 end_time 1.2
2  친구들 start_time 1.2 end_time 1.8
3  본격적으로 start_time 1.8 end_time 2.5
4  1단원 start_time 2.5 end_time 3.4
5  시작합니다 start_time 3.4 end_time 4.2
6  우리 start_time 4.2 end_time 5.4
7  생활과 start_time 5.4 end_time 5.9
8  민주정치 start_time 5.9 end_time 6.9
9  만화로 start_time 6.9 end_time 7.6
10 만나 start_time 7.6 end_time 7.9
11 봤어요 start_time 7.9 end_time 8.5
12 우리가 start_time 8.5 end_time 9.0
13 어떤 start_time 9.0 end_time 9.5
14 부분들을 start_time 9.5 end_time 10.2
15 공부할지 start_time 10.2 end_time 10.8
16 말이에요 start_time 10.8 end_time 11.6
17 일하는 start_time 11.6 end_time 12.9
18 게 start_time 12.9 end_time 13.1
19 없으면 start_time 13.1 end_time 13.8
20 너무 start_time 13.8 end_time 14.1

```

<그림 3. STT음성인식 단어별 Time Stamp >

Transcript: 우리 친구들 본격적으로 1단원 시작합니다. 우리 생활과 민주정치 만화로 만나 봤어요. 우리가 어떤 부분들을 공부할지 말이에요 일하는게 없으면 너무 혼란스럽습니다. 우리 사회도 마찬가지예요. 청취 민주정치 여기에서 우리 친구들 중요한 개념 나옵니다. 정치의 의미 정치라고 하는게 뭐야 국회의원들이 하는게 정치 아니야. 물론 국회의원들이 하는 것도 정치 맞습니다. 그런데 우리 생활 속에서도 우리도 정치를 하고 있어요. 의미입니다. 사람들 사이에서 일어나는 갈등이 대립을 조정하고 많은 사람들이 한 공간에 살면서 공동체 사회를 만들었습니다.

<그림 4. STT음성인식 Transcript 실제 결과 >


Transcript: 우리 친구들 본격적으로 1단원 시작합니다. 우리 생활과 민주정치 만화로 만나 봤어요. 우리가 어떤 (어떠한) 부분들을 공부할지 말이에요. 일하 (규칙이라) 는게 없으면 너무 혼란스럽습니다. 우리 사회도 마찬가지예요. 청취 (와) 민주정치 여기에서 우리 친구들 (정말) 중요한 개념 나옵니다. 정치의 의미 정치라고 하는게 뭐야 국회의원들이 하는게 정치 아니야. 물론 국회의원들이 하는 것도 정치 맞습니다. 그런데 우리 생활 속에서도 우리도 정치를 하고 있어요. (정치외) 의 의미입니다. 사람들 사이에서 일어나는 갈등이 (나) 대립을 조정하고 많은 사람들이 한 공간에 살면서 공동체 사회를 만들었습니다.

<그림 5. STT음성인식 Transcript 오류 표시 >

총 단어 수	인식하지 못한 단어	오류난 단어	정확하게 인식한 단어	정확도
75	2	3	70	약 93.33%

<표 1. STT음성인식 결과 분석 >

TEST 영상을 Google STT를 통해 인식한 단어의 수와 인식하지 못한 단어, 오류난 단어, 정확하게 인식한 단어를 비교분석하였고 정확도를 확인할 수 있다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	수화 인 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 1.3	2019-APR-18

2.2.2.2 자막 생성

STT로 음성인식한 텍스트 파일을 통해 VTT형식의 자막파일로 변환해주는 기능을 구현하였다. 자세한 수행 내용은 다음과 같다. 음성을 인식한 transcript와 word들의 time stamp들을 텍스트파일로 받아온다. transcript를 문장단위로 나누고, 각 문장들을 word들과 비교하며 문장의 time stamp를 지정해준다. 그리고 VTT의 파일 형식에 맞게 문장들과 타임코드를 변환하여 작성하여 자막파일을 생성한다. 사용자의 가독성을 높이기 위해 문장의 길이가 너무 긴 경우 문장을 나누는 기능도 추가해주었다.

```


1 WEBVTT
2
3 1
4 00:00:00.400 --> 00:00:04.200
5 우리 친구들 본격적으로 1단원 시작합니다
6
7 2
8 00:00:04.200 --> 00:00:08.500
9 우리 생활과 민주 정치 만화로 만나 봤어요
10
11 3
12 00:00:08.500 --> 00:00:11.600
13 우리가 어떤 부분들을 공부할지 말이에요
14
15 4
16 00:00:11.600 --> 00:00:14.900
17 일하는게 없으면 너무 혼란스럽습니다

```

<그림 6. VTT파일 변환 결과>

2.2.3 자막 수화 번역 - 형태소 분석

자막으로 변환한 파일에서 문장들만 추출한 뒤 Konlpy API를 사용하여 형태소 분석을 수행한다. 각 문장의 형태소를 분석한 뒤 수화 문법 체계에 맞도록 재구성한다. 수화 문법에서는 조사를 거의 사용하지 않기 때문에 필요한 몇몇의 조사를 제외하고 모두 제거한다. 과거형이나 의문문을 표현하기 위해 문장의 형태소들을 재배치한다. 의문사는 문장의 끝부분에 위치시키고 동사의 과거형은 동사의 현재형을 배치하고 그 옆에 '끝'이라는 수화 동작을 연결시켜 과거형을 만든다. 기존 한국어의 숫자 표현 형식도 수화 문법에 맞도록 변형한다. 예를들면 한국어에서 432라는 숫자는 수화 문법에선 400 + 30 + 2로 표현 방식을 바꿔줘야 한다. 이 밖의 Konlpy가 미처 신경써주지 못하는 신조어나 대명사는 사용자 사전에 직접 추가하여 형태소 분석 기능을 더욱 유연하게 만든다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	수화 인 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 1.3	2019-APR-18

1 1	1 1
2 00:00:00.400 --> 00:00:04.200	2 00:00:00.400 --> 00:00:04.200
3 우리 친구 들 본격적 으로 1 단원 시작 하다	3 우리 친구 들 본격적 으로 1 단원 시작 하다
4 2	4 2
5 00:00:04.200 --> 00:00:08.500	5 00:00:04.200 --> 00:00:08.500
6 우리 생활 과 민주 정치 만화 로 만나 보다 끝	6 우리 생활 과 민주 정치 만화 로 만나 보다 끝
7 3	7 3
8 00:00:08.500 --> 00:00:11.600	8 00:00:08.500 --> 00:00:11.600
9 우리 어떤 부분 들 공부 하다 말 이다	9 우리 어떤 부분 들 공부 하다 말 이다
10 4	10 4
11 00:00:11.600 --> 00:00:14.900	11 00:00:11.600 --> 00:00:14.900
12 일하다 것 없다 너무 혼란 하다	12 일하다 것 없다 너무 혼란 하다
13 5	13 5
14 00:00:14.900 --> 00:00:17.700	14 00:00:14.900 --> 00:00:17.700
15 우리 사회 도 마찬가지로 예	15 우리 사회 도 마찬가지로 예요
16 6	16 6
17 00:00:17.700 --> 00:00:24.300	17 00:00:17.700 --> 00:00:24.300
18 청취 민주 정치 여기 에서 우리 친구 들 중요 하다 개념 나오다	18 청취 민주 정치 여기 에서 우리 친구 들 중요 하다 개념 나오다
19 7	19 7
20 00:00:24.300 --> 00:00:31.200	20 00:00:24.300 --> 00:00:31.200
21 정치 의 의미 정치 하다 것 무엇 국회 의원 들 하다 것 정치 아니다	21 정치 의 의미 정치 하다 것 무엇 국회 의원 들 하다 것 정치 아니다
22 8	22 8
23 00:00:31.200 --> 00:00:35.300	23 00:00:31.200 --> 00:00:35.300
24 물론 국회의원 들 하다 것 도 정치 맞다	24 물론 국회의원 들 하다 것 도 정치 맞다
25 9	25 9
26 00:00:35.300 --> 00:00:41.100	26 00:00:35.300 --> 00:00:41.100
27 그런데 우리 생활 속 에서 도 우리 도 정치 하다 있다	27 그런데 우리 생활 속 에서 도 우리 도 정치 하다 있다
28 10	28 10
29 00:00:41.100 --> 00:00:42.700	29 00:00:41.100 --> 00:00:42.700
30 의미 이다	30 의미 이다
31 11	31 11
32 00:00:42.700 --> 00:00:49.400	32 00:00:42.700 --> 00:00:49.400
33 사람 들 사이 에서 일어나다 갈등 대립 조정 하다	33 사람 들 사이 에서 일어나다 갈등 대립 조정 하다
34 12	34 12
35 00:00:49.400 --> 00:00:56.500	35 00:00:49.400 --> 00:00:56.500
36 많다 사람 들 한 공간 에 살다 면서 공동체 사회 만들다 끝	36 많다 사람 들 한 공간 에 살다 면서 공동체 사회 만들다 끝

<그림 7. 형태소 분석 및 형태소 재구성>


총 단어 수	형태소 분석 오류	형태소 분석 성공	정확도
107	2	105	약 98.13%

<표 2. 형태소 분석 결과>

<그림 7>의 오른쪽 그림은 현재 Konlpy API를 사용한 뒤 형태소 분석 오류 사례를 보여주며, 왼쪽 그림은 성공된 사례로써 비교분석하여 <표2>에서와 같이 정확도를 확인할 수 있다.

2.2.4 수화 번역 영상 생성

자막 파일의 각 문장의 시작시간과 끝시간의 차를 계산하고, 문장 번역 결과에 따라 수화들을 연결한다. 그리고 연결된 파일을 문장의 지속시간에 맞추어 재생 속도를 조절한다. 이와 같은 작업을 반복하여 각각의 문장에 해당하는 수화영상들을 다시 연결하여 최종적으로 하나의 수화번역 영상을 만드는 기능을 구현하였다. 동영상의 연결은 openCV 패키지를 이용하였고, 동영상의 재생속도를 조절은 moviepy 패키지를 이용하였다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	수화 인 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 1.3	2019-APR-18

3 수정된 연구내용 및 추진 방향

3.1 수정사항

3.1.1 OpenPose 제외

초기에는 수화데이터 수집 시, 자연스러운 수화 동작의 연결을 위해 OpenPose로 탐지하여 수화 동작의 스켈레톤 데이터를 따오고 이를 애니메이션으로 가공하여 사용자에게 제공할 계획이었다. 하지만 수화는 손 동작 뿐만 아니라 표정도 의미에 포함되는 복합적인 언어이기 때문에 애니메이션으로 가공할 경우 오히려 실제 사람이 하는 수화보다 품질이 떨어질 수 있다는 점과, OpenPose를 이용해 수화를 만드는 것에 대한 기술적인 당위성이 부족하다는 계획서 발표 피드백을 고려하여 OpenPose는 프로젝트 요소에서 제외시키기로 결정하였다. 대신 수집한 원본 수화영상을 연결하여 보여주는 방식으로 진행하기로 하였다.

3.1.2 Backend 개발 환경 변동


Backend 개발환경에 있어서 웹서버는 Apache에서 Nginx로, 개발 프레임워크는 Php 기반의 CodeIgniter에서 Python기반의 Django로 변경하였다. Django는 성능과 속도가 빠르면서 가볍고, 현재 프로젝트 개발이 Python으로 대부분 이루어지고 있기 때문에 호환성 역시 고려하여 채택하게 되었다.

서버 개발환경이 바뀐에 따라 DataBase에도 변동사항이 있다. Django에는 자체 DataBase인 SQLite를 지원하는데, 가볍게 사용할 수 있고 따로 미들웨어를 쓰지 않고 라이브러리만 포함시켜 사용할 수 있는 장점이 있다. 하지만 한 번에 한 명의 사용자만이 데이터베이스를 사용할 수 있어 멀티 유저를 지원하는 우리의 프로젝트와 맞지 않다는 판단하에 SQLite는 테스트용으로 사용하며, 웹 서버의 DataBase로는 MongoDB를 채택하였다. MongoDB의 사용이유는 2.1.1.2에서 다루었다.

3.1.3 Android Studio 앱에서 web 서비스로 변동

Web 서비스는 반응형 웹 디자인을 사용하여 모바일 web으로도 호환이 가능하다. 이는 누구나 사용하는 웹 브라우저로 접근할 수 있고 유저가 사용하기에 편리하다는 장점이 있다. 앱은 기본적으로 '설치'가 되어야 하는 반면 웹은 모든 기기에서 빠르게 접근이 가능하다. 또한 앱은 OS별로 각각 수정해야 하며, 수정 이후에도 마켓의 정책 기준에 부합하지 못하면 등록이 거부될 수 있는 단점이 있다. 반응형 앱은 하나의 화면에 많은 콘텐츠를 표현할 때 적합하지 않다. 우리는 쇼핑몰같이 많은 콘텐츠를 표현하지 않으므로 반응형 웹의 단점에 해당하지 않는다. 이러한 이유로 웹 서비스가 더 유용하다고 생각했다.

3.1.4 형태소 분석 언어를 R언어에서 Python으로 대체

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	수화 인 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 1.3	2019-APR-18

R과 Python은 모두 통계를 위한 유명한 프로그래밍 언어이다. R은 통계학자를 염두에 두고 개발된 언어이며 강력한 시각화 기능을 제공하고 있고 Python은 초보자에게 이해하기 쉬운 언어로 평가된다. Python은 R보다 학습 조건이 까다롭지 않기 때문에 코드 작성 시간을 줄여주고 코드를 재사용 할 수 있다. 특히 시중에 공개된 꼬꼬마, 코모란, 트위터, 한나눔, 온전한눈 다섯개 형태소 분석기를 한꺼번에 묶어 편리하게 사용할 수 있도록 만들어진 Konlpy 패키지는 파이썬을 지원한다.

4 향후 추진계획

4.1 향후 계획의 세부 내용

4.1.1 선택적 서비스

사용자는 자막 서비스와 수화 서비스를 필요에 따라 선택하여 원하는 서비스를 제공받을 수 있다. 첫번째 계획으로 동영상 등록할 때 선택하는 방식과 두번째 계획으로 자막과 수화 서비스를 전제로 두고 영상이 실행되는 동안 선택적으로 서비스 사용 유무를 설정할수 있는 방식을 고려하고 있다.

4.1.2 URL을 이용한 업로드 기능 추가


현재 프로젝트에서 제공하는 업로드 방법은 사용자의 컴퓨터에 ‘보고자 하는 영상’이 저장되어있어 파일을 웹사이트에 업로드하는 방식으로 상당히 제한적이다. 최근 트렌드는 동영상을 다운 받아 시청하기보다는 스트리밍 또는 비디오 플랫폼을 이용하여 시청하기 때문에 URL을 이용한 업로드를 추가하려고 한다.

4.1.3 단어 유사도 검사


2.2.4에서 형태소에 맞는 수화 동영상을 데이터베이스에서 검색할 때 만약 동음이의어가 나온다면 단어 유사도 검사를 통해 동음이의어 중 하나를 선택한다고 기술하였다. 이를 구현하기 위해 Python의 Word2Vec과 Gensim 패키지를 이용하여 단어를 벡터화하고 단어간 유사도를 구한다.

4.1.4 전문가 자문

본 프로젝트의 사용자는 청각장애인으로 프로젝트의 완성도는 청각장애인이 직접 사용했을 때 동영상에 대한 내용 이해도, 수화와 자막의 일치율에 따라 결정된다. 따라서 수화전문가, 청각장애인을 만나 프로젝트 목적에 대한 견해 및 자문을 구할 계획이다. 직접 사용하였을 때, 불편한 사항이나 추가로 필요한 사항을 자문받아 완성도를 높이려고 한다. 현재 국민대학교 중앙동아리 ‘손말사랑회’ 수화 동아리에

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	수화 인 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 1.3	2019-APR-18

문의하여 전문가와의 연락을 시도 중이다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	수화 인 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 1.3	2019-APR-18

5 고충 및 건의사항

1. 수화데이터가 부족하여 데이터를 가공하여 사용하기에 어려움을 겪고 있습니다.
2. 국립국어원 한국수어사전에 있는 수화영상을 크롤링해서 사용했으나 저작권 문제로 데이터 사용에 어려움을 겪고 있습니다.