

청각장애인을 위한
실시간 화자 구분 음성 인식 타이핑 프로그램

Talk To Me

종합설계프로젝트2

HearUR | 2012105028 김한별
2012105037 박소은
2012105040 박지인
2012105057 송은영

Talk To Me
INDEX

1. 팀원 역할 소개
2. 주제 정리
3. 사용 도구
4. 수행 완료 항목
5. 일정표
6. 수행 예정 항목

01

팀원 역할 소개

01



박소은 [팀장]

- : 프로젝트 진행 관리
- : 아날로그 음성 변환 및 입력

02

03

04

05



김한별 [구현관리]

- : 프로젝트 구현 지휘
- : 인공지능 알고리즘 분석 및 테스트



박지인 [문서관리]

- : 문서 수집 및 제작, 정리
- : 화자 분류와 타이핑 속도 조율



송은영 [예산관리]

- : 예산 수립 및 신청, 관리
- : 화자 분류 결과 출력 및 전송

02

주제 정리

01

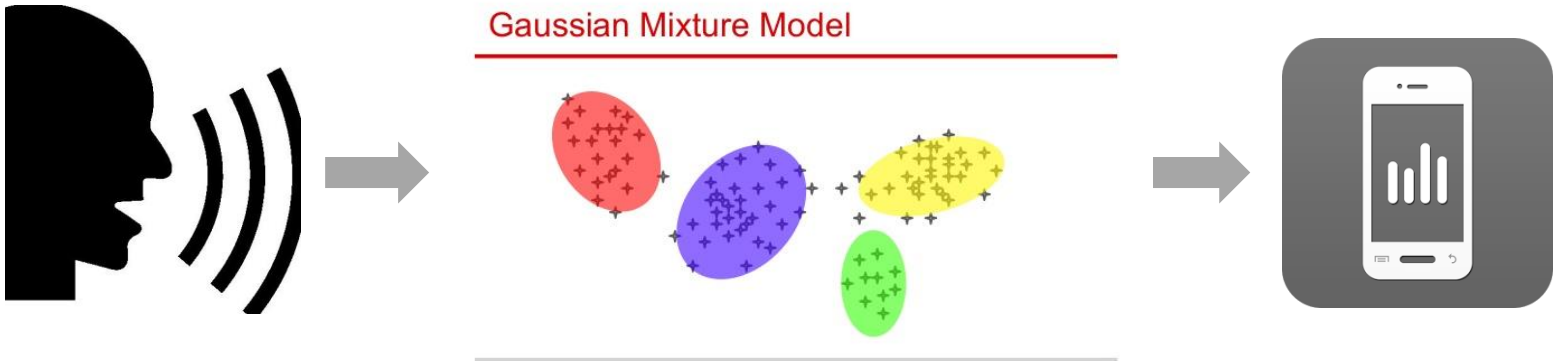
02

03

04

05

>> 주제



Unsupervised learning
가우시안 혼합 모델 (GMM, Gaussian Mixture Model)

03

사용 도구

01

02

03

04

05

>> 아마존 웹서비스



- 아마존 클라우드 서비스

01

02

03

04

05

>> LIUM & Voice ID

LIUM Speaker Diarization



- Speaking Diarization Open Source 'LIUM'
- Lium을 이용하여 화자인식에 특화시킨 Open Source 'Voice ID'

01

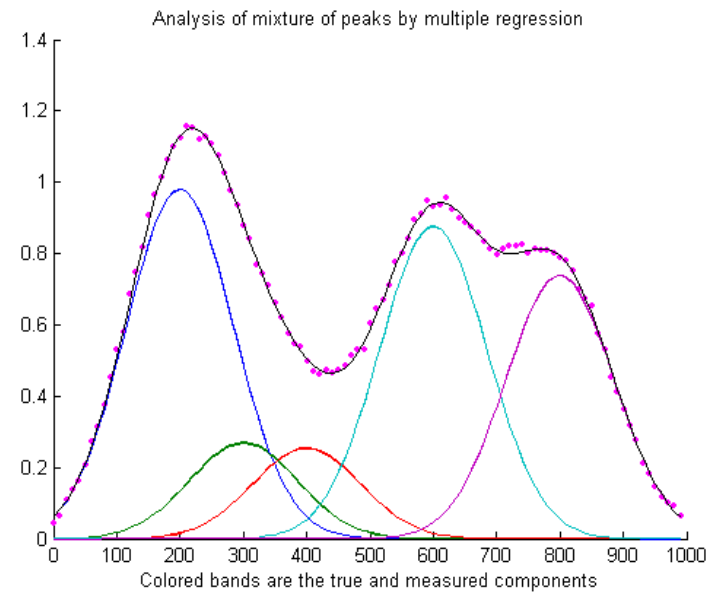
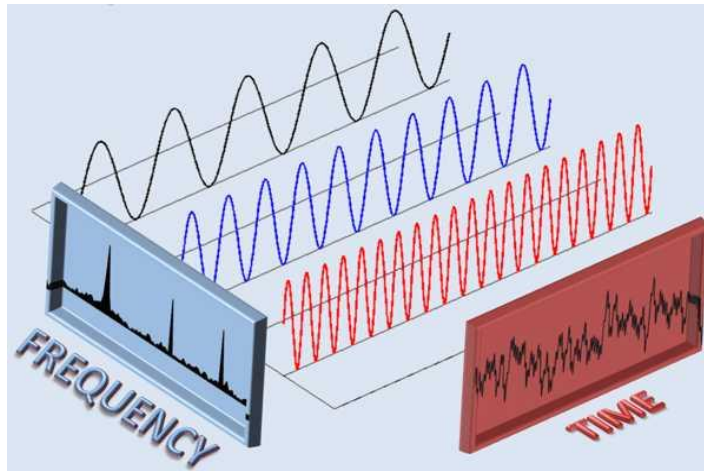
02

03

04

05

>> MFCC & GMM



- 입력된 음성에 FFT를 적용해 추출한 MFCC사용
- EM 알고리즘으로 파라미터를 구한 GMM사용

04

수행 완료 항목

01

02

03

04

05

>> 디비 생성

```
[>>> from voiceid.db import GMMVoiceDB  
[>>> db = GMMVoiceDB('./db')  
[>>> db.add_model('./output/out', 'BYEOL')  
True  
_
```

- 저장해놓은 음성으로 GMM모델을 적용한 디비 생성

- 안드로이드로 목소리를 입력 (WAV)
- 입력받은 목소리를 서버로 전달 (BYTE)

01

02

03

04

05

>> 화자식별

```
converting_file
diarization
trimming
voice matching
```

```
*****
```

```
speaker  S0
```

```
00:00:00,000 to 00:00:04,180
```

```
The system identified this speaker as 'BYEOL'!
```

- 가장 높은 유사성을 보이는 화자가 선택됨

05

일정표

01

02

03

04

05

01

02

03

04

05

06

07

08

09

10

11

12

주제 선정

수행 계획 수립

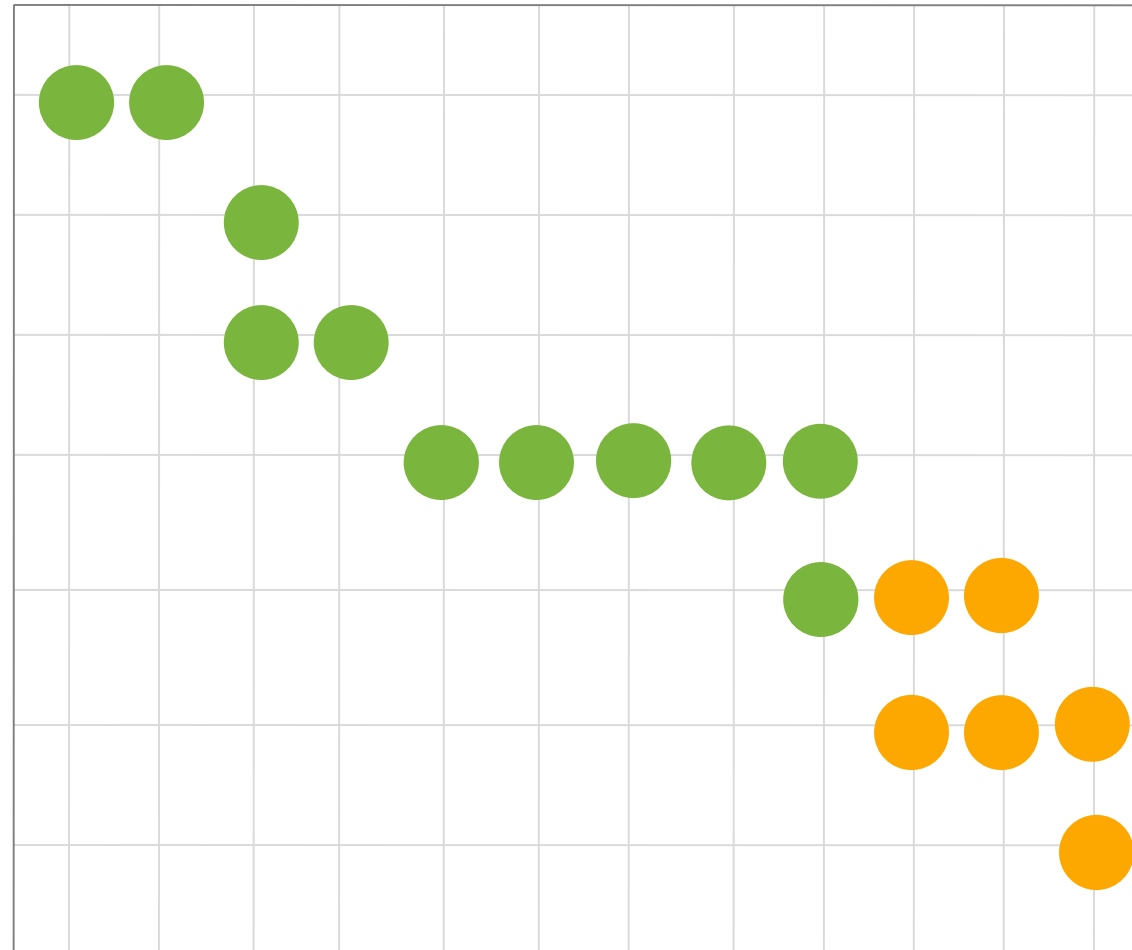
시스템 설계

기본 기능 구현

정확도 개선

테스트

결과 보고 및 발표



06

수행 예정 항목

01

02

03

04

05

>> 타이핑과 인물 구분 기능의 결합



- 타이핑 속도와 인물구별 속도차에 주의
- 안드로이드 출력화면 제작

01

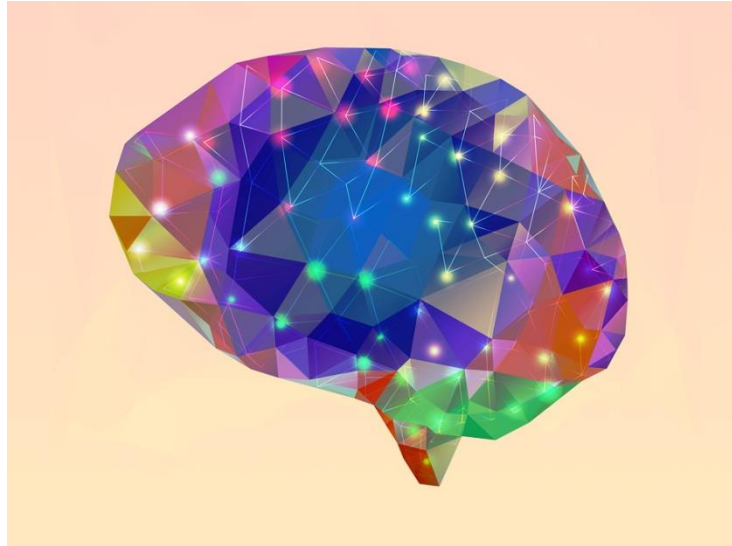
02

03

04

05

>> 테스트 & 정확도 개선



- 효과적이고 정확한 출력을 위해 수많은 테스트 실행
- 많은 수행으로 인공지능 알고리즘의 정확도 개선

청각장애인을 위한
실시간 화자 구분 음성 인식 타이핑 프로그램
Talk To Me

Thank you