1. 과제명 : STT, SVM을 이용한 청각장애인용 통화 SW

2. 팀원 : 25-1기 최환종 (PL), 25-1기 안중환, 25-1기 정다비치

3. 아이디어 제안

청각장애인 또는 듣기가 불편한 고령자들은 전화통화에 어려움을 겪는다. 이런 문제에 도움을 주기 위해 첫째로 상대방의 음성을 Text로 변환하여 전화중인 청각장애인에게 보여준다. 두 번째로 듣지 못하는 청각장애인은 텍스트만 보고는 상대방이 어떤 감정으로 말을 하였는지를 알기 어렵다. 우리는 상대방의 음성을 이용해 감정을 분석하여 청각장애인에게 감정까지 보여준다.

4. 개발 방법

1. STT를 이용한 텍스트 변환  
   통화음성을 Android에서 묵음단위로 녹음하여 Core Server로 보낸 후 서버에서 Google STT Server에 Http 요청으로 JSON Object를 얻어오도록 하였습니다. 또한 이 음성을 감정결과를 얻기 위한 원형 Signals로 사용합니다. 또한 실시간 처리와 Connection 관리를 위해서는 GCM방식은 적합하지 않아, Connection Server에 메시지 Stream을 두는 방식으로 끊김에 상관없고 데이터의 손실이 없게 처리하였다.
2. 에너지의 통계적 분포(PLP)를 이용한 지도학습:

* 음성 전처리 및 SVM 지도학습:

EMO DB를 사용하여 0~2700Hz구간에서 6개의 감정자질들을 추출하기 위해 Wav Signal을 FFT이후 피처추출에 적합하도록 정규화하고, PLP 일부 과정을 처리한 후 신호의 세기에 무관하도록 Scaling하였다. 주파수영역으로 바뀐 신호를 6등분하고 각각의 파워를 구하여 특징으로 사용하였다. 분류정확도를 향상시키기 위해 한국드라마를 추가로 Training Set으로 사용하였다. SVM은 RBF커널을 사용하였다.

* Android byte버퍼 전처리:

FFT시간을 줄이기 위한 작업으로 묵음구간과 잡음을 최대한 제거하여 Signal을 줄이고 16Khz에서 5333hz로 샘플링주파수를 재조정하여(1~2700hz까지 획득) 처리시간을 향상시킨 후 전처리 작업을 하도록 하여 3초이내에 감정을 얻어낼 수 있게 하였다. 또한 저주파Index에서 같은 패턴을 보여 저주파의 에너지 값을 조정하여 고주파를 강조하는 방법으로 패턴을 구별하였다.

5. 프로젝트 기여도 평가

|  |  |  |  |
| --- | --- | --- | --- |
| **팀원** | **기여도(%)** | **역할** | **근거** |
| 최환종 |  |  |  |
|  |  |  |
| 안중환 |  |  |  |
|  |  |  |
| 정다비치 |  |  |  |
|  |  |  |

6. 목표관리 평가서

|  |  |  |
| --- | --- | --- |
| **구현목표** | 해당 대회 수상 등급 | **가중치** |
| **상세내용** | A+ : 대통령상, 국무총리상, 장관상  A- : 기관장상 [기관장상, 4팀]  B+ : 장려상 [단체장상, 7팀]  C+ : 본선진출 [32팀] | 100 |
| **평가** | **근거** | |
| **B+** | 글로벌 SW 공모대전 장려상 수상 | |