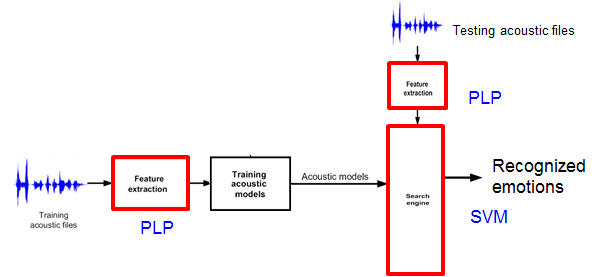
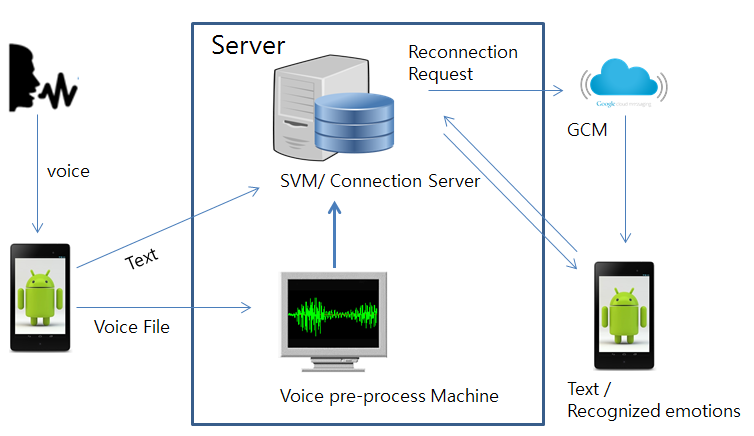
1. 과제명 : STT, SVM을 이용한 청각장애인용 통화 SW개발

2. 팀원 : 최환종, 안중환, 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



(시스템 구성도)

1. STT를 이용한 텍스트 변환: Google Speech To Text Engine을 이용하여 발화자의 음성을 그대로 Text로 변환하여 서버를 통해 상대방에게 보내준다. Google STT엔진은 지속적인 기계학습이 이루어지기 때문에 그 정확도가 매우 높으므로 결과를 그대로 전송한다. 이때 양쪽 socket연결의 불안정성을 해결하기 위해 GCM을 보조도구로 사용한다.
2. 에너지의 통계적 분포(PLP)를 이용한 지도학습:

2-1) 기쁨(Happiness), 슬픔(Sadness), 화남(Anger), 보통(Neutrality)의 4가지 감정으로 분류한다.

2-2) EMO-DB의 감정 별 Sample음성 500개를 학습데이터로 사용하여 전 처리한다.

① 잡음제거→ ②진폭 평균화→ ③ 고정분할→ ④ PLP분석→ ⑤구간별 에너지 값 추출

2-3) 에너지량 Code Book 생성(Bag of word기법)

2-4) Code Book을 이용하여 Training Set 생성 (Label 지도)

Training Set → [ ① Label { ②감정자질 후보군 ③감정자질 TF } ]

2-5) 완성된 Set을 기계학습의 Training Set으로 사용

2-6) 감정분석: 감정 분석을 위한 음성을 전 처리과정을 거쳐 Set을 만든 후 SVM에 Label요청을 한다. 추가적인 학습을 위해 Label이 포함된 Training Set을 다시 학습시킨다.

5. 전체 개발 일정

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **팀원** | **역할** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** |
| 최환종 | Connection서버 구현 |  |  |  |  |  |  |  |  |  |  |  |  |
| GCM연동/Connection안정화작업 |  |  |  |  |  |  |  |  |  |  |  |  |
| Training Set생성 알고리즘 구현 |  |  |  |  |  |  |  |  |  |  |  |  |
| SVM 분류기 구현 |  |  |  |  |  |  |  |  |  |  |  |  |
| 안중환 | 안드로이드 설계 및 모듈개발 |  |  |  |  |  |  |  |  |  |  |  |  |
| 음성 전처리 모듈 구현 |  |  |  |  |  |  |  |  |  |  |  |  |
| PLP변환 알고리즘 구현 |  |  |  |  |  |  |  |  |  |  |  |  |
| 정다비치 | 결과 시각화를 위한 UI구현 |  |  |  |  |  |  |  |  |  |  |  |  |
| 샘플데이터 지도학습 |  |  |  |  |  |  |  |  |  |  |  |  |
| 샘플데이터 추가 수집/표본화 |  |  |  |  |  |  |  |  |  |  |  |  |
| 공통 | Training Set Label분류 |  |  |  |  |  |  |  |  |  |  |  |  |
| 통합 및 디버깅 |  |  |  |  |  |  |  |  |  |  |  |  |

6. 목표관리 평가서 (구, 평가 받고 싶은 카테고리)

|  |  |  |
| --- | --- | --- |
| **구현목표** | **음성이 실시간으로 처리되어 전송되는가?** | **가중치** |
| **상세내용** | 음성이 청각장애인에게 TEXT + 감정으로 실시간 전송이 되는가?  - 전화와 실시간으로 청각장애인 기기에서 TEXT로 보여줘야 한다.  (TEXT는 실시간, 감정은 시간차를 허용).  - 대화가 가능할 정도의 처리 속도  - 유동IP에 관여 받지 않는 안정적인 데이터 통신 | 25 |
| **구현목표** | **음성을 이용한 감정분석이 가능한가?** | **가중치** |
| **상세내용** | 음성을 이용한 감정 파악이 가능한가?  - 감정이 실린 음성을 기쁨, 슬픔, 화남, 보통으로 표시하는가?  - 다양한 목소리(음색, 크기)에도 정확한 감정분석이 가능한가?  - 성별에 관계없이 올바른 감정결과를 얻을 수 있는가?  - 잡음환경에서의 감정분석이 가능한가?  - 화남/기쁨의 감정구분이 잘되는가? | 40 |
| **구현목표** | **시각화** | **가중치** |
| **상세내용** | 음성에 따른 감정을 충분히 시각적으로 표현하였는가?  - TEXT를 위젯 형식으로 통화App과 함께 볼 수 있는가?  - 감정결과값을 사용자가 쉽게 확인할 수 있는가? | 10 |
| **구현목표** | **완성도** | **가중치** |
| **상세내용** | 1. 모바일 UI 및 사용편의성  2. 감정분류의 정확도  3. 서버의 학습 능력 및 피드백  4. 끊김이나 지연 없는 안정성 | 25 |