**소프트웨어 프로젝트 II AD 프로젝트**

**개발 계획서 및 보고서**

****

|  |  |
| --- | --- |
|  |  |
| 과 목 | **소프트웨어 프로젝트 II** |
| 담당 교수 | **윤성혜 교수님** |
| 제출 일자 | **2020.11.30** |
| 전공 | **소프트웨어학부** |
| 학번 | **20203086** |
| 이름 | **송수인** |

1. **요구사항 명세서**
2. **기능적 요구사항**

* 사용자에게 직접 이미지 파일 이름을 입력 받을 수 있어야 한다.
* 파일 이름이 잘못됐다면 사용자에게 이를 알리는 메시지가 출력되어야 한다.
* 저시력자를 위해 디스플레이 속 글자의 크기를 조절할 수 있어야 한다.
* 디스플레이에 출력되는 텍스트는 최대한 원본 이미지의 형식과 비슷해야 한다.
* 입력받아 OCR 기능이 수행된 결과가 자동으로 음성파일로 변환되어야 한다.
* 음성파일의 이름은 이미지 파일 이름과 같아야 한다.
* 텍스트가 출력되는 디스플레이는 사용자에 의해 수정되지 말아야 한다.
* 사용자가 직접 음성 파일을 업로드할 수 있어야 한다. (TTS 기능만 이용할 경우 대비)

1. **사용자 인터페이스 요구사항**

**2-1. Looking4U의 사용자 인터페이스를 구성하는 요소는 다음과 같다**.

**2-1-1. Text 관련 요소**

* **Display** (QTextEdit)

OCR 결과로 생성된 텍스트를 사용자에게 표시하는 창. 기본 텍스트1. 크기는 12pt.

* **Control text size slider** (QSlider)

Display에 표기된 텍스트의 크기를 조절할 수 있는 슬라이드 바. (12 ~ 40으로 범위 지정)

* **Text control buttons** (QPushButton)

1. **Clear Text**

디스플레이에 표기된 모든 텍스트를 삭제.

1. **Add New Text**

OCR을 수행할 새로운 이미지 파일 업로드를 위한 버튼 ( 이미지 파일이 업로드 되지 않았을 때 예외 고려)

**2-1-2. 음성 관련 요소**

* **Sound List** (QTableWidget)

OCR 결과로 전달된 텍스트를 음성 파일로 변환한 결과( .wav)와 사용자가 임의로 추가한 오디오 파일을 표시하는 테이블 ( 사용자가 오디오 파일을 업로드하지 않은 경우 예외 고려)

* **Control Sound file buttons** (QPushButton)

1. **Add List**

QFileDialog를 이용하여 사용자가 직접 파일 탐색기를 통해 파일을 선택할 수 있도록 설계

1. **Del List**

사용자가 클릭한 요소를 삭제할 수 있도록 하는 버튼

* **Play Control buttons** (QPushButton, QSlider)

재생, 일시정지, 정지 기능은 필수. 다음 파일 재생이나 배속 기능은 수준 내에서 구현이 가능할 시에 시도.

**3. 비기능적 요구사항**

3-1. 이 소프트웨어의 구현에는 Python과 PyQt5를 이용한다.

3-2. OCR 기능 구현을 위한 외부 라이브러리로 Pytesseract를 이용한다.

3-3. TTS 기능 구현을 위한 외부 API로 gTTs( Google Text To Speech ) API를 이용한다

**2. 소프트웨어 구조 설계**

|  |  |  |
| --- | --- | --- |
| ***모듈*** | **클래스** | **역할** |
| guiAD | MainWindow | 사용자 인터페이스의 모든 위젯들을 포함하는 메인 윈도우 |
| imgToText | Ocr | 이미지 파일을 텍스트로 변환하는 OCR 기능과 변환된 텍스트를 음성 파일로 변환하는 TTS 기능을 포함한 핵심 기능이 수행되는 모듈 |
| Player | SoundPlayer | 음성 출력과 관련된 모든 통제를 수행하는 모듈 |

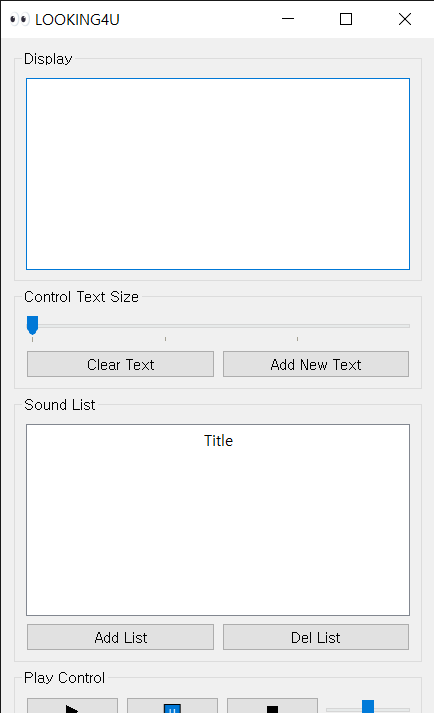
< 프로그램 개요 >

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **클래스** | **메서드** | **입력인자** | **출력인자** | **기능** |
| MainWindow | initUI | - | - | 사용자 인터페이스 구성 |
| changeFontSize | - | - | Display 내의 텍스트 크기 조정 |
| clearText | - | - | Display 내 모든 요소 삭제 |
| newText | - | - | 사용자에게 이미지 파일 이름을 입력받아 ImgToText.py로 전달 |
| tableChanged | - | - | Sound List 내에 선택된 요소들 처리 |
| addList | - | - | Sound List에 추가할 음성 파일을 사용자에게 입력받음 |
| delList | - | - | Sound List 내의 음성 파일 삭제 |
| btnClicked | - | - | 음성 파일 재생과 관련된 버튼 제어 |
| volumeChanged | - | - | 음성 파일 재생 시, 소리 조절 |
| createPlaylist | - | - | Player.py에게 전달할 재생 목록 생성 |
| Ocr | ocr\_tesseract | File name | textResult | 사용자로부터 입력받은 파일 이름인 file name 인자를 image\_to\_string 함수에 넣어 텍스트 생성 |
| textToSpeech | textResult | .wav 음성파일 | 생성된 텍스트를 넘겨받아 음성 파일로 변환 |
| SoundPlayer | Play | - | - | 음성 파일 재생 |
| Pause | - | - | 음성 파일 일시정지 |
| Stop | - | - | 음성 파일 정지 |
| createPlaylist | Playlist | Playlist(QMediaPlaylist) | 음성 파일 재생을 위한 QMediaPlaylist 생성 |
| updateVolume | Volume slider value | - | 음성 파일 Volume 조절 |

**< 상세 설계 >**

**3. 개발 결과**

**3-1. 프로그램 초기화면**

****

**3-2. 보완점 및 해결책**

1. 한글 인식률이 다소 떨어지며 곡면과 반사가 심할수록 인식률이 감소한다.

* 한글 학습 데이터 양을 늘려 인식률을 높이고 이미지 컷팅을 통해 곡면과 빛 반사를 보정한다.

1. 저시력자 분들이 사용하기에 인터페이스가 단순하지 않다.

* UI 및 창의 크기를 키우고 글자 크기 범위를 조절한다.

1. 텍스트의 양이 많을 때, Display 내에서 형식이 깨지는 현상이 발생한다.

* 정확한 원인은 알 수 없으나 이미지 내의 텍스트 외 요소가 유력한 원인으로 파악되므로 이를 잘라내거나 흐릿하게 처리하여 인식률을 높이는 방법 고안