

Project Research Document – Optical Character Recognition Software

X00088318 – Stephen Plummer

Detailed Discussion

For my fourth year project I have chosen to develop an OCR style application to archive and convert documents to text files. I feel this will benefit many people including publishing companies who wish to archive old publications that currently only remain on paper.

While these pages and books could easily just be scanned I feel text editions that can be edited would be very helpful. To be able to open up the document and ‘modernise’ the text or to copy sections to improve Google translate or to translate yourself.

Image scans can be inefficient. For instance if an old piece of text is slightly illegible, text in the crevice of the spine or sloppy, handwritten text will be just as illegible or even more so on a scan, depending on the quality of scanner used. To scan a document and convert the words to text on your screen would allow you to check for errors and correct them, if you know what a certain word is that may come across as a different one to some people.

My application will be developed on the Android system, which I believe currently to be the most reliable and popular system utilizing popular programming languages for example Java and to an extent JSON, although having some plugins that make the development process a different enough language. Apple devices however use mainly their own ‘Objective-C’ language, which is quite different to most common languages available.

I hope for my app to have a simple GUI, allowing possibly a log-in feature, which will utilize Google accounts, saving your scans to text documents within Google Drive.

The bulk of the application will be camera-focused, allowing a user to simply open the camera and take an image of the text they would like to convert and then the process will take place – however, I would like to further the use of Drive and allow previously scanned documents to be converted in the same way, instead of opening the camera.

Before jumping headfirst into this application. I’ve decided to research the Android Studio app and some Android programming examples namely found here - <http://www.codota.com/> which will be helpful including snippets on opening the camera from within an app.

Existing Applications in this domain

<i>Differences</i>	<i>Similarities</i>	<i>App</i>
Limits the portion of the image taken from the camera.	Allows an image to be taken and a file to be selected to convert	SmartText
Allows image enhancement to clarify the text.	Allows image to be taken to convert	OCR Instantly Free
Saves OCR'd images to Notes for easy access	Allows image to be taken to convert	OCR

Platform, Technologies and Libraries

I plan to develop this app for the Android platform. To do this I will need use of Android Studio and/or eclipse for writing and compiling the code.

There is a library in use by Google called Tesseract I would like to try to implement. However there are others to try and see which will work more accurately such as ABBYY's OCR library.

I will need knowledge of JSON, Java and possibly other languages, which may provide a learning curve, however this would be ideal to gain more broad experience with programming.

The Risks

One of the risks involved with starting development on this app for me is that I have so far been unfamiliar with the programming process for android. I, however, have experience with Java among other languages which should aid in my learning of a new one