

DOCUMENT ANALYSIS AND RECOGNITION USING ANDROID PLATFORM

Prof. Jagruti Wagh, Priyanka Pawar, Ajay Kasabe, Ganesh Lavhare
Department of Computer Engineering
Marathwada Mitra Mandal's College Of Engineering, Karvenagar, Pune
Savitribai Phule Pune University, India

ajayk21111990@gmail.com pawar.priyanka.171@gmail.com, gany0312@gmail.com

Abstract: Document Reader using Android is the concept to use smart phone in a more smart way. We were imagine one thing which slightly magical and impressive is that whatever document we want to actually read that the same document can we listen with doing some other work?

Answer is yes, we can design a system which can be listen instead of reading. Our document reader using Android project topic can fulfill this idea. To make such kind of a system we are using daily used camera with good camera quality. In our busy life style people are not spend more time for a reading books, novels, news papers and any other physical paper books which is more difficult to carry with them, because lack of time peoples are ignore such hobby's but our system will overcome on all these problems and get a excellent solution. Using this system people will actually get knowledge without holding book in hand without reading also. System will provide facility to listen book. If we want to read some document or book or newspaper but we don't have that much time to hold in hand and read then we can simply take a snap by phone camera and listen by running our system/application.

Our purpose was to create an application that allows a user to jump to the correct section of an audio book by taking a picture of the page they are reading in a reading in a paper book. We used image processing algorithms such as Image Binaralization, Contour Extraction method/ sobel algorithm and Morphological Transformation.

Keywords: Binaralization, sobel algorithms, morphological transformation

1. INTRODUCTION

Today, avid readers can access their media in a variety of ways. E-readers are portable, and audio books allow listeners to perform other tasks while following a story however, some traditionalists prefer the physical experiences associated with paper books. Within the course of a single novel, a reader might have use for all three formats. For example, a user

may need to go grocery shopping while in the midst of reading a paper book. If they were to switch to the audio book format, they could continue enjoying the novel while driving a car. On day trips, a quick skim of a digital book in a cell phone might be preferable.

Manually synchronizing these formats can be cumbersome, especially since E-readers have different page schemes depending on font size, audio book audio files aren't precisely divided into separate paragraphs, and the same novel by different publishers have different page schemes.

Now that large databases are easily accessible through the cloud, a means of automatically syncing all three media formats is possible. Companies like Amazon have implemented methods of synchronizing audio books and digital e-reader formats. We propose to bridge the gap between paper books and audio books using image processing. The main challenge is matching chapter of a physical book to a time stamp in the corresponding audio book.

2. OPERATING ENVIRONMENT

Mobile application software will support Android phones with at least 2.1 version of Android.

3. DESIGN AND IMPLEMENTATION CONSTRAINTS

At least 1 Android mobile device needed to show a working model. Also system should have sufficient RAM to get process properly.

The project will support only English written scripts to get converted into English or any language. Mobile camera quality will be good enough to create good quality image of written text data.

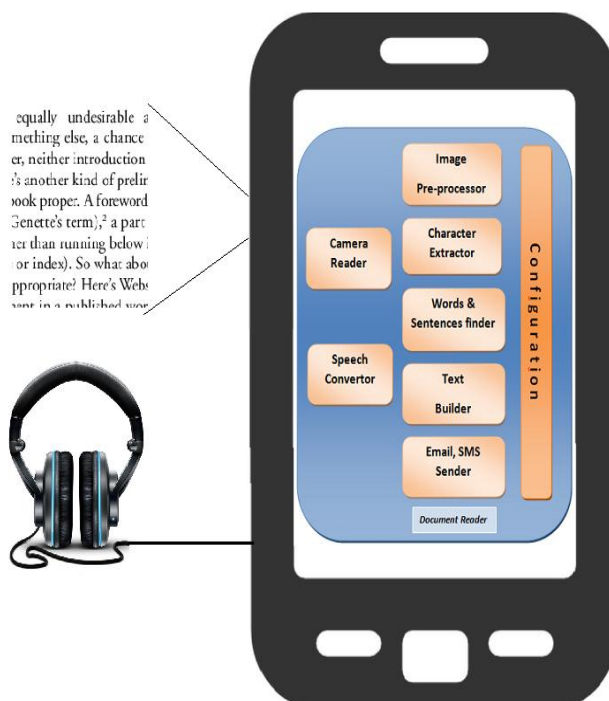


Fig.1: Architecture of Document Reader App

4. WORKING

Scanned pages of a physical book will train the classifier while pictures taken by an Android-based smart phone will be converting that image text into audio. About multiple paragraph reading we can Take small snaps and arrange it sequence wise.

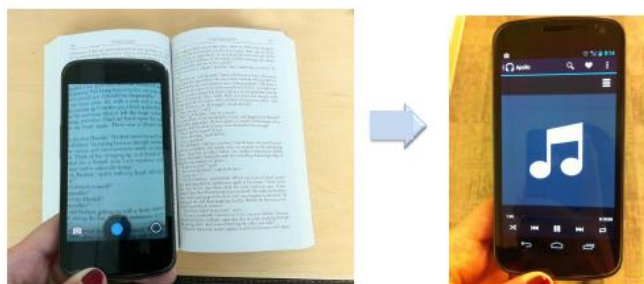


Fig.2: Working steps of Document Reader App

5. PROJECT FLOW

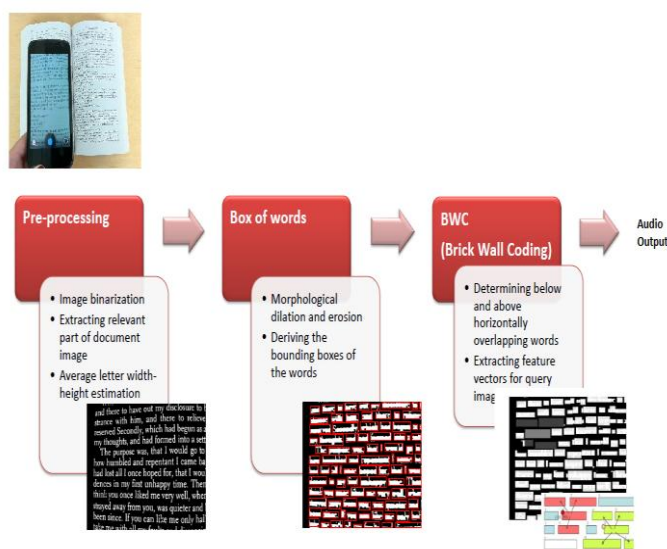


Fig.3: Steps of processing Algorithm

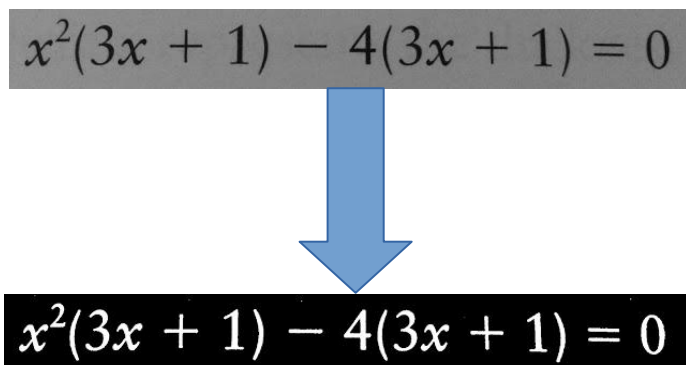
6. EXTERNAL ALGORITHM

We cannot check size..... We just match pattern (x.....X).For make under same pattern we can use image processing make a square of both character and take a mean of character and get number.Three main method of Extracting is follows:-

- 1) Image Binarization
- 2) Contour Extraction method / sobel algorithm
- 3) Morphological Transformation

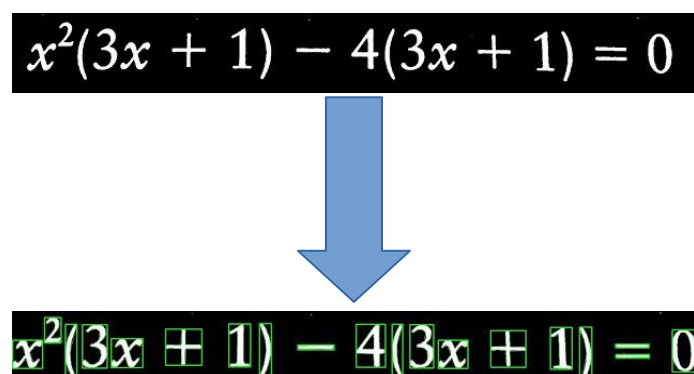
6.1 Image Binarization

This is also known as Adaptive Thresholding algorithm. The function transforms a grayscale image to a binary image



6.2 Contour Extraction method

Contour extraction is nothing but the word identification. For that we can use algorithm which is known as sobel algorithm. Contour function to detect and find all the "blobs" or shapes. (Each letter) and make a square of both character and take a mean of character



6.3 Identifying Text Boxes

For identified the text boxes we can Use 'Morphological Transformation' to find letters those that are part of a common word. I.e. separating word in sentences



7. TTS ENGINE

- TTS Engine is the most important thing in Document reader system.
- TTS engine is already present in android
- Whenever we want to use it then Initialize first
- With the proper suitable language. Setting the language.
- Input should be in text format.
- Set peach (speed / audio voice male or female)
- **Then final actual string to convert into audio**

8. EXISTING SYSTEM

According to Document Analysis and Recognition (ICDAR), 2011, they described a system for full text search in which they augment three probabilistic text retrieval methods with knowledge about expected OCR errors. The approach used confusion information for specific characters, along with bi-gram probabilities of character occurrences to create multiple possible search terms for each initial search term. After performing the search with each new term, the validity of returned documents is based on the confusion and bi-gram occurrence probabilities. The results claim increases from 2-3% in recall with decreases or 4-5% in precision. Author used a similar approach in which they use confusion statistics to generate an enhanced finite state machine for query terms in Japanese text.

9. PROPOSED SYSTEM

Here in our application we are going to focus on the act to make our application portable to be used by the people who are in a hurry or cannot read. Old age persons cannot read a lot of the times so they can use our application to convert the text into audio and listen to it whenever they want. The user is in no need to have a technical knowledge to use this application. To achieve this we are going to make available different languages in which the user wants to listen and the text will be provided by our application if he wants to send it to another person. This is an off line (does not require Internet connection to work) Optical Text Recognition application which can be used to convert text (English ONLY!!!! for now) on a paper to editable digital text on your device. Extracted/converted text can be listening in audio format.

10. CONCLUSION & FUTURE SCOPE

The implementation stage involves careful planning, investigation of the existing system and its constraints on implementation, designing of methods to achieve changeover and evaluation of changeover methods. The application will allow the user to listen to any text's image captured by user its peak thus satisfying the customer's need of reading and understanding document in audible format. Porting to other Smartphone platforms like iOS, BlackBerry OS, Windows Mobile 7 etc.

We will focus on extracting features from images of a physical book's pages and convert them in to audio. Scanned pages of a physical book will train the classifier while pictures taken by an Android-based Smartphone will be converting that image text into audio.

ACKNOWLEDGEMENT

The success and final outcome of this project required a lot of guidance and assistance from many people and we are extremely fortunate to have got this all along the completion of my project work. Whatever I have done is only due to such guidance and assistance and I would not forget to thank them.

We are thankful to and fortunate enough to get constant encouragement, support and guidance from all teaching staffs of our department which helped us in successful completion of our project.

REFERENCES

- [1]. *Real-time document image retrieval for a 10 million pages database with a memory efficient client* Kazutaka Takeda, Koichi Kise, and Masakazu Iwamura. In *Document Analysis and Recognition (ICDAR)*, 2011 International Conference on, pages 1054_1058, 2011.
- [2]. K. Takeda, K. Kise, and M. Iwamura, "Real-time document image retrieval for a 10 million pages database with a memory efficient and stability improved LLAH," *Proceedings of the 11th International Conference on Document Analysis and Recognition (ICDAR2011)*, 2011.
- [3]. T. Nakai, K. Kise, and M. Iwamura, "Real-time retrieval for images of documents in various languages using a web camera," *Proceedings of the 10th International Conference on Document Analysis and Recognition (ICDAR2009)*, pp. 146–150, July 2009.