International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

Read to Me: A Cloud Oriented Reading Assistance for Visually Impaired

Veena F Kuradagi[1], Anitha G[2]

[1]PG Student, University BDT college of Engineering, Visveswaraya Technological University, Hadadi Road, Davangere, Karnataka, India

[2]Associate Professor Dept CS &E, University BDT College of Engineering, Hadadi Road, Davangere, Karnataka, India

Abstract: ***Perusing is fundamental in day by day life for everybody. Outwardly debilitated people can read just by utilization of unique applications by them like Braille dialect. The disadvantage of this framework is that each item does not give the content in Braille. In this paper we have proposed an assistive content perusing system to help outwardly impeded people to peruse writings from different questions in their day by day lives. At first we catch the picture of the required, pre handling is performed on it. Pre handling incorporates steps like dark scale and Binarization, question of intrigue acknowledgment. In the proposed framework we are making the utilization of OTSU calculation to change over the dim scale picture into binarized one. The content districts from the caught picture are then separated and perceived by utilizing optical character acknowledgment programming (OCR). The principle calculation in OCR to be specific MODI is utilized here. This extricated content of different textual styles and sizes then can be perceived independently and afterward consolidated in a word giving its yield as sound utilizing Text-to-discourse utilizing the SAPI libraries.***

Keywords: Assistive Text reading, Binarization, OTSU algorithm, SAPI libraries, OCR, MODI algorithm

1. Introduction

Reading is very essential in our daily lives. Out of 314 million visually impaired people all around the world, 45 Million are blind and new cases being added each year. Recent developments in computer vision, digital cameras, and portable computers make it feasible to assist these individuals by developing camera-based products that combine computer vision technology with other existing commercial products such optical character recognition (OCR) systems. Printed text is all over in the type of reports, receipts, bank statements, hotel menus, school or college handouts, product packages, medical bottles etc.

There are many assistive systems available today but they have certain issues reducing the feasibility for the visually challenged persons. For example, portable bar code readers designed to help blind people identify different products, it enables the users who are blind to access

Information about these products through speech and Braille. But a big limitation is that it is very hard for blind users to find the position of the bar code and to correctly point the bar code reader at the bar code. There are systems like K Reader Mobile it runs on a cell phone and allows the user to read mail, receipts, fliers, and many other documents. However, the document to be read must be nearly flat, placed on a clear, dark surface (i.e., a no cluttered background), and contain mostly text. Furthermore, K Reader Mobile accurately reads black print on a white background but has problems recognizing colored text or text on a colored background. It cannot read text with complex backgrounds. The main aim is to develop such a system that will read the texts from complex backgrounds successfully.

2. Proposed System

The proposed system helps blind persons to read product labels. Users have to capture image and then system read out the text from image .It will be very useful for Persons those are going through optical surgery. It can be useful for road side text detection so that blind person can travel independently. The proposed system provides effective solution as Compared to most of the existing systems.

Image capture module
The image capture module will detect the image captured by the Camera attached to the goggles. This will be easy for the visually impaired person to capture the image as the camera will be situated on the goggles. The image captured will be converted into grayscale and Binarization

Image correction module
This module will correct the image by reducing the noises by mean of filtering algorithm like median filter so that the text will be effectively recognized.

Text extraction module
This module will recognize and extract the text. This will be achieved using OCR-Optical Character Recognition - is the mechanical or electronic conversion of images of typed, handwritten or printed text into machine-encoded text. We will be using here MODI algorithm of OCR.

Audio module
This module will get the extracted text as an input and it will read out the text using Text To Speech available in the mobile. If no text is recognized then default audio output will be given. This will be done using SAPI libraries.

3. Literature Survey

[1] Xiaodong Yang, Yingli Tian Chucai Yi Aries Arditi "Context-based Indoor Object Detection as an Aid to Blind Persons Accessing Unfamiliar Environments" 2010 Computer vision based indoor way finding system is

implemented for blind people to independently access unfamiliar buildings. A blind person can find different rooms and building exit or an elevator. This system includes text recognition. It detects doors based on general geometric shape, by combining edges and corners. To differentiate between an office doors from a bathroom door, it extract and recognize the text information. [2] The system use OCR and location can be delivered through speech for blind travelers.

[3]Xiaodong Yang, Shuai Yuan, and YingLi Tian

"Recognizing Clothes Patterns for Blind People by Confidence Margin based Feature Combination" clothes pattern can be recognized using this system. There are many clothes patterns. This system is classifying clothes patterns into 4 categories: stripe, lattice, special, and pattern less. In this system texture analysis methods only focused on textures varying with distinctive pattern changes. Due to large intra class variations in each clothes pattern category. It cannot achieve level of accuracy for clothes pattern recognition. Extracting statistical and structural feature from image wavelet sub bands can be a solution of this problem.

Detecting Boris, Epshtein Eyal, Ofek Yonatan Wexler "Text in Natural Scenes with Stroke Width Transform"2010. A novel image operator used to find the value of stroke width for each image pixel. [4] It is used in text detection in natural images. The suggested operator is data dependent and local, which makes it fast and it is strong enough to reduce the need for scanning windows or multi-scale computation. Extensive testing shows that the suggested scheme outperforms the latest published algorithms. Its simplicity allows the algorithm to detect texts in many fonts and languages.

Asif Shahab, Faisal Shafait, Andreas Dengel "ICDAR 2011 Robust Reading Competition Challenge 2: Reading Text in Scene Images" 2011 International Conference on Document Analysis and Recognition of Text in natural scene images is becoming a prominent research area because imaging devices like mobile phones are available[5]. The ICDAR 2011 Robust Reading Competition was organized to evaluate the performance of recent algorithms in recognizing and detecting text from complex images.

In [6] Sneha Sharma, Dr. Roxanne Canosa, advisor ``Extraction of Text Regions in Natural Images" 2007. The detection and extraction of text regions in an image is a well known problem in the computer vision research area. The goal of this project is to compare two basic approaches to text extraction in natural (non-document) images: edge-based and connected-component based. The algorithms are implemented and evaluated using a set of images of natural scenes that vary along the dimensions of lighting, scale and orientation. Accuracy, precision and recall rates for each approach are analyzed to determine the success and limitations of each approach. Recommendations for improvements are given based on the results.

In [7] Dimitrios Dakopoulos and Nikolaos G. Bourbakis, Fellow, ``Wearable Obstacle Avoidance Electronic Travel Aids for Blind: A Survey" IEEE 2010

The variety of portable or wearable navigation systems have been developed to help blind people during navigation in outindoor or indoor environments. There are three main categories of these systems: Electronic travel aids, and position locator devices, electronic orientation aids. This paper is a comparative survey of wearable obstacle detection systems to users and informs the research community about the capabilities of these systems and about the progress in assistive technology for visually impaired people. The survey is based on various performance parameters and features of the systems that classify them in categories.

In [8] Bharat Bhargava, Pelin Angin, Lian Duan " A Mobile-Cloud Pedestrian Crossing Guide for the Blind" This system help blind and visually-impaired persons to detect the status of pedestrian signals at street for safe outdoor navigation.This system propose a mobile-cloud collaborative approach for Context aware outdoor navigation, where it use the computational power of resources made available by cloud computing providers for real-time image processing. The proposed system architecture has the advantages of being minimal infrastructural reliance and extensible, thus allowing for wide usability.The proposed approach is for real-time crossing guidance for blind pedestrians.

In [9] Yingli Tian, Chucai Yi ``Assistive Text Reading from Complex Background for Blind Persons". This paper presents a system for blind persons to read text from object and signage that are held in the hand. The system read text from complex backgrounds and then communicates this information aurally. They design a novel text localization algorithm to localize text regions in images with complex backgrounds, by learning gradient features of distributions of edge pixels in an Adaboost model and stroke orientation. Optical character recognition (OCR) software is used to recognize Text characters in the localized regions and transformed into speech outputs. The performance of the proposed system is evaluated on ICDAR 2003 Robust Reading Dataset.

In [10] Yasar Guneri Sahin*, Baris Aslan, Sinan Talebi, Ayberk Zeray ``A SMART TACTILE FOR VISUALLY IMPAIRED PEOPLE"2015. The impaired persons have many difficulties in society. One of the most significant difficulties is traveling because of inappropriate city designs. Recent developments in technology have enabled several facilities, such as tactile paving surfaces, to improve their lives, but so far there is no comprehensive solution to the problems they face. This study proposes a new, cost efficient and simple system, which consists: tagged paths and batons to make traveling alone possible. The proposed system is available for Android mobile devices and IOS and consists of two software applications, "Out Guide", and "In Guide" for indoor and outdoor environments respectively.

4. System Design

The proposed system helps blind persons to read product labels. Users have to capture image and then system read out the text from image .It will be very useful for Persons those are going through optical surgery. It can be useful for road

side text detection so that blind person can travel independently. The proposed system provides effective solution as Compared to most of the existing systems.

Read with Android

Mobile phones are one of the most commonly used electronic gadgets today. Here, we intend to develop a modular and friendly application using cloud based OCR platform and the built in Android TTS for producing an audible result of the text file.

The architecture is one of the representation of the theoretical system that defines the structure, behavior along with more specifications about the system. It indicates the fundamental organization of the system that describes various components in it and their relationship with each other and the corresponding environment.

Figure 1: System architecture

The Above figure .1 shows overall system architecture. When user want to read any object then check for internet connection is there or not. If internet connection is there in the mobile user should click on the camera button and capture the object photo which he wants to read. Then it immediately uploaded to OCR cloud for recognition and data extraction. Here image will be converted to text file. Then recognized data (.txt file) export Read with Cloud application. Read with Cloud application sends data for speech synthesis. Now the text file is send to default android text to speech (TTS) engine. TTS engine outputs speech for the recognized text file. After all these process audio plays to the user.

5. Proposed Technique

The most advanced optical character recognition systems, such as ABBYY Fine Reader OCR, are focused on replicating natural or "animal like" recognition. In the heart of these systems lie three fundamental principles: Integrity, Purposefulness and Adaptability. The principle of integrity says that the observed object must always be considered as a "whole" consisting of many interrelated parts. The principle of purposefulness supposes that any interpretation of data must always serve some purpose. And the principle of adaptability means that the program must be capable of self-learning.

OCR processing is CPU intensive. Therefore, in an effort to keep the power consumption mobile application to minimum, and to increase the speed of the OCR, we selected to have the processing be done using the ABBYY Cloud OCR SDK. Its service is platform in dependent due to the fact it is accessible through Web API and is not running on the device itself. So a Web API provided by ABBYY can be developed to run under any OS platform.

6. Result and Analysis

The application only requires an Android phone as Its hardware. In order to achieve a faster performance in running the application, we recommend a minimal specification of the device to be 400MHz and 128M RAM. The device must have 5MP as the minimal resolution of the camera in order to have a resolution of the captured image between 150dpi to 600dpi. The application is compatible with all the API levels 15 (Ice-cream sandwich) to 23 (Lollipop). The packages used are: Android Studio (IDE for Android Programming), Android local T S and ABBYY OCR SDK Cloud. Android's notable features such as open source platform, multiple screen for multitasking, custom ROM, and open source libraries for Text-To Speech Synthesis superseded Android over other OS versions for the work presented here.

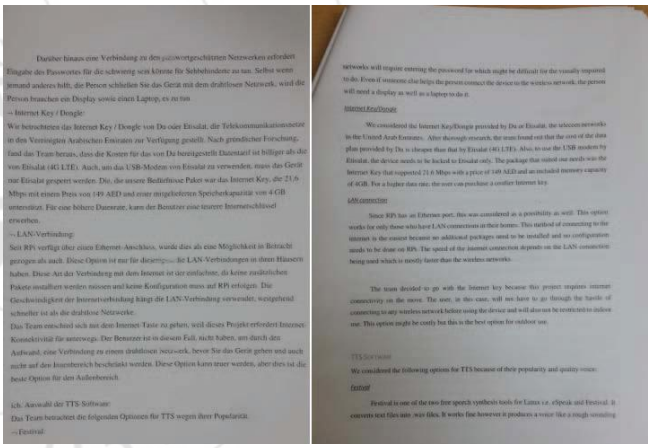

Figure 2: Text image captured by Android smart phone

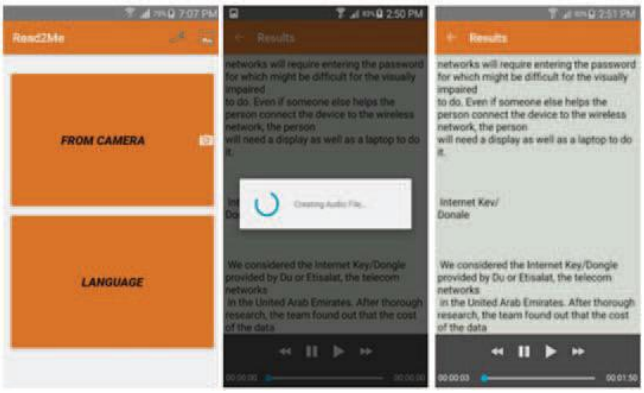

Figure 3: (a) Main screen view of the application (b) Result of OCR (c) Text to be read out loud

However, the software is commercial and requires an internet connection. Since ABBYY incorporates pre-processing and post processing stages for the OCR-ed text, therefore it

stands out as the most optimum platform for characters recognition. It eliminates the overhead cost of improving image quality before extracting text from the image.

This software is not only limited to the recognition of the text documents but as well as barcode recognition, hand-printed text recognition, business card recognition and supports up to 198 recognition languages including French, English and German in varied font styles such as Normal (Ariel, Times New Roman or Courier), Gothic, typewriter text, magnetic ink characters and matrix . Moreover, after running tests for the comparison of recognition accuracy, it was found out that ABBYY has an accuracy of 95.96% compared to 89.78% of Tesseract an alternate OCR engine.

7. Testing Read to Me on Android

a) Accuracy
The pictures (German and English) captured and shown in Figure 2, passed the OCR and TTS conversion process with an accuracy of 99.9%, with only 2 letters being misread because they were underlined (for example, y is misread as v) or faded (and sometimes E is misread as F), as shown in Figure 2.

b) Latency
The application returned the text (in English) containing 334 words in 12 seconds. The text was 12pts Times New Roman. For German, the application took 13 seconds containing 327 words with 12 pts font. Creating the audio file takes about 5 seconds for both languages. Therefore, it could be deduced that the total time from capturing the image to hearing the speech output is approximatey17 seconds.

c) Usability
The application uses voice directives, as well as voice confirmations, to inform the visually impaired user of what the application is currently doing, however playing/pausing the audio can be an issue here, since there are no voice labels for that and the blind person will possibly face some difficulty finding those buttons on the touch screen. Moreover, the blind person can also touch one of these buttons on the screen by mistake, but that will be easily detected by the user since on touching any of the buttons, the corresponding action will be committed.

d) Power Consumption
The application takes 10.39 MB of memory when installed. To check how much battery, the application uses up, the Smart Manager pre-installed on the Samsung S4 is used. Figure 4 contains screenshots from the Smart Manager which shows that the application, when used, only takes up 1% of battery and a RAM of 27.51 MB and CPU usage of 2.15%

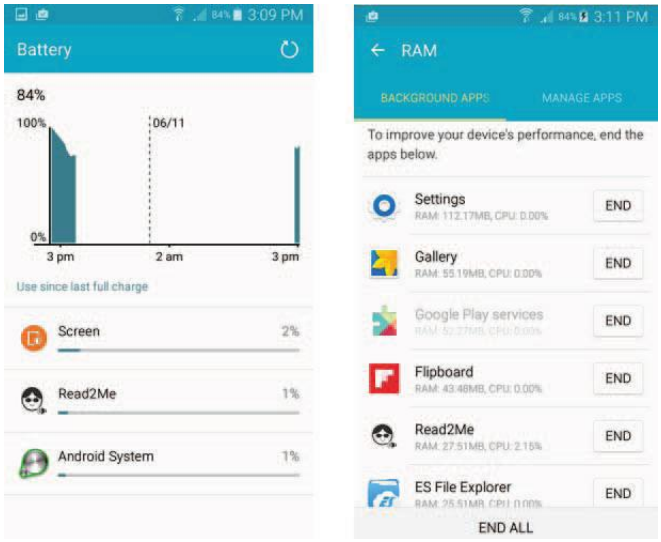

Figure 4: Power and RAM Consumption for Read2Me Application

e) Portability
The application requires only a smart phone with an internet connection, today's smart mobiles are light in weight with the trend towards smaller and lighter models. The phone used for testing was Samsung S4 which only weighs 130 g.

f) Cost
The cost of Samsung Galaxy S4 varies depending on where it is purchased from but it ranges from 700-900 AED ($189 to $240). However, any reasonably equipped mobile for a cheaper price should be able to run the application. As with the Raspberry Pi device, a free trial of the OCR service is used. The university's wireless network was also used for this application. Therefore, their costs are not included in the estimate.

8. Conclusion

We propose an assistive system to read printed text on objects for assisting blind persons. To solve the common problems of blind people we have proposed a method in which the blind people will click the image. This method can effectively separate the objects of interest from complex background and other objects in the camera view OCR is used to perform word recognition of the localized text regions and transform into audio output for blind people.

References

[1] World health organization official website. August 2014. [Online]. Available at http://www.who.int/mediacentre/factsheets/fs282/en. Accessed on March 4, 2015.
[2] R. Keefer., & N. Bourbakis. 'Interaction with a Mobile Reader for the Visually Impaired'. 21st IEEE International Conference with Artificial Intelligence Tools. 18.03 (2009): 229-236. Web.
[3] R.Shilkrot.,&P.Maes.(2014,May.1).FingerReader: A wearable device to support text reading on the go.[Online]. Available: http://fluid.media.mit.edu/sites/default/files/paper317.pd

f .
[4] R.Neto. &N.Fonseca. 'Camera Reading for Blind People'. Volume 11, 11.11 (2014) 1200-1209.[Online].Available at http://www.sciencedirect.com/science/article/pii/S22120 1731400362 4
[5] Nagaraja L, et al, "Vision based text recognition using raspberry PI", *National Conference on Power Systems & Industrial Automation (NCPSIA 2015).*
[6] Nagarathna, Sowjanya V, M. "Product label readin system for visually challenged people", International journal of Computer Science and Information Technology Resreach, Vol 3, Issue 2, 2015.
[7] M. Jeon, A. Ayala-Acevado, N. Nazneen, B. Walker, O. Akanser, '"Listen2dRoom": Helping Blind Individuals Understand Room Layouts' in CHI '12 Extended Abstracts on Human Factors in Computing Systems, Austin, TX, U.S.A., 2012, pp. 1577 – 1582.
[8] "Tesseract-OCR," 2015. [Online]. Available: https://code.google.com/p/tesseract-ocr/. [Accessed 28 Oct 2015].
[9] Abbyy.technology, 'Supported OCR Text/Print Types [Technology Portal]', 2015. [Online]. Available: https://abbyy.technology/en:features:ocr:supported_ocr_ text_types. [Accessed: 23- Oct- 2015].
[10] 'Smart Implementation of Text Recognition (OCR) for Smart Mobile Devices', The First International Conference on Intelligent Systems and Applications, pp. 19-24, 2012.
[11] "espeak text to speech," 2015. [Online]. Available: http://espeak.sourceforge.net/. [Accessed 25 Oct 2015].
[12] "ArchLinux: Festival," 13 Oct 2015. [Online]. Available: https://wiki.archlinux.org/index.php/Festival. [Accessed 25 Oct 2015].
[13] "Raspberry Pi," 05 Feb 2014. [Online]. Available: https://www.raspberrypi.org/forums/viewtopic.php?f=38 &t=68693. [Accessed 28 Oct 2015].
[14] "Android Developers Blog," 23 September 2009. [Online]. Available: http://android-developers.blogspot.ae/2009/09/introduction-to-text-to-speech-in.html. [Accessed 28 Oct 2015].