International Journal of Science and Engineering Applications
Volume 9–Issue 04,49-52, 2020, ISSN:-2319–7560

Reading Device for Blind People using Python, OCR and GTTS

Supriya Kurlekar
SITCOE (Yadrav), India

Onkar A. Deshpande
SITCOE (Yadrav), India

Akash V. Kamble
SITCOE (Yadrav), India

Aniket A. Omanna.
SITCOE (Yadrav), India

Dinesh B. Patil.
SITCOE (Yadrav), India

Abstract: This paper presents the reader for Blind people, developed on Raspberry Pi 2. It uses the Optical character recognition technology for the identification of the printed characters using image sensing devices and computer programming [1]. It converts images of typed or printed text into machine encoded text. In this research these images are converted into the audio output (Speech) through the use of OCR and Text-to-speech synthesis. The conversion of printed document into text files is done using Raspberry Pi which again uses PyTesseract library and Python programming. The text files are processed & convert into the audio output (Speech) using GOOGLE Text-to-speech (gTTS) & python programming language and audio output is achieved.

Keywords: Character recognition, Pi Camera, Raspberry Pi 2, Python Programming, Text To Speech (TTS), Speech Output.

1. INTRODUCTION

This kind of system helps visually impaired people to interact with computers effectively through vocal interface. Text Extraction from color images is a challenging task in computer vision. Text-to-Speech is a device that scans and reads English alphabets and numbers that are in the image using OCR technique and changing it to voices. Now a day's SMS is one of the most popular way of communication using mobile phone but visually impaired people cannot use this.

This project has been built around Raspberry Pi processor board. It is controlling the peripherals like Camera and speaker which act as an interface between the system and the user. Optical Character Recognition or OCR is implemented in this project to reco gnize characters which are then read out by the system through a speaker. The camera is mounted on a stand in such a position that if a paper is placed in front of camera, it captures a full view of the paper into the system. Also, when the camera takes the snapshot of the paper, it is ensured that there are good lighting conditions. The content on the paper should be written in English and be of good font size.

When all these conditions are met the system takes the photo, processes it and if it recognizes the content written on the paper. After this it speaks out the content that was converted in to text format in the system from processing the image of the paper. In this way Reading Device for Blind People helps a blind person to read a paper without the help of any human reader.

2. WORKING PRINCIPLE

When we run the Python Program, this system captures the image placed in front of the picamera which is connected to Raspberry Pi .After captured document image undergoes Optical Character Recognition(OCR) Technology.
OCR technology allows the conversion of scanned images of printed text or symbols into text or information that can be understood or edited using a computer program. In our system for OCR technology we are using Pytesseract library.
After that Convert image into text, text convert into speech using Text-to-speech library we use GOOGLE Text-to-speech library using this data will be converted to audio. Camera acts as main vision in detecting the image of the placed document, then image is processed internally and separates label from image by using open CV library and finally identifies the text which is pronounced through voice. Now the converted text into audio output is listened either by connecting headsets via 3.5mm audio jack or by connecting speakers via Bluetooth.

3. BLOCK DIAGRAM

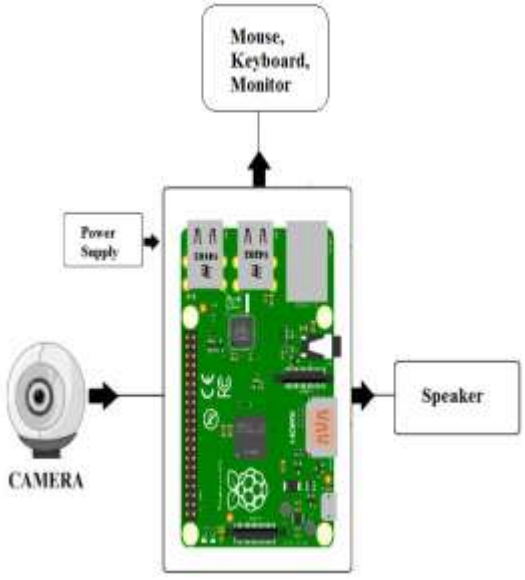

Figure.1 Block diagram of Reading Device for Blind People

4. HARDWARE IMPLEMENTATION

Figure.2 Reading Device for Blind People

Raspberry Pi is a low cost, credit card sized computer that connects to monitor and uses standard keyboard and mouse. The hardware components of the Raspberry Pi include power supply, storage, input, monitor and network.

- CPU: Broadcom BCM2836 900MHz quad-core ARM Cortex-A7 processor
- RAM: 1 GB SDRAM
- USB Ports: 4 USB 2.0 ports
- Network: 10/100 Mbit/s Ethernet
- Power Ratings: 600 mA (3.0 W)
- Power Source: 5V Micro USB
- Size: 85.60 mm × 56.5 mm
- Weight: 45 g (same as Raspberry Pi B+)
- 802.11n Wireless LAN
- 40 GPIO pins
- Full HDMI port
- Combined 3.5mm audio jack and composite video
- Camera interface (CSI)
- Display Interface (DSI)
- Micro SD card slot

Piamera
The Raspberry Pi camera module can be used to take high-definition video, as well as stills photographs. The camera module is very popular in home security applications, and in wildlife camera traps.

- 5MP sensor
- Wider image, capable of 2592x1944 stills, 1080p30 video
- 1080p video supported
- CSI
- Size: 25 x 20 x 9 mm

HDMI to VGA Converter
It is used to connect the Raspberry Pi board to the Projectors, Monitors and TV.

5. SOFTWARE IMPLEMENTATION

5.1 Programming Explanation

5.1.1 Python-Tesseract

Python-Tesseract is an optical character recognition (OCR) tool for python. That is, it will recognize and "read" the text embedded in images.

Python-Tesseract is a wrapper for Google's Tesseract-OCR Engine. It is also useful as a stand-alone invocation script to Tesseract, as it can read all image types supported by the Pillow and Leptonica imaging libraries, including jpeg, png, gif, bmp, tiff, and others. Additionally, if used as a script, Python-Tesseract will print the recognized text instead of writing it to a file.

Functions

- **get_tesseract_version** Returns the Tesseract version installed in the system.
- **image_to_string** Returns the result of a Tesseract OCR run on the image to string
- **image_to_boxes** Returns result containing recognized characters and their box boundaries
- **image_to_data** Returns result containing box boundaries, confidences, and other information. Requires Tesseract 3.05+. For more information, please check the Tesseract TSV documentation
- **image_to_osd** Returns result containing information about orientation and script detection.
- **run_and_get_output** Returns the raw output from Tesseract OCR. Gives a bit more control over the parameters that are sent to Tesseract.

Installation

- pip install pytesseract

5.1.2 GTTS (Google Text-to-Speech)

GTTS (Google Text-to-Speech), a Python library and CLI tool to interface with Google Translates text-to-speech API. Write spoken mp3 data to a file, a file-like object (byte string) for further audio manipulation, or stdout. Or simply pre-generate Google Translate TTS request URLs to feed to an external program.

Features

- Customizable speech-specific sentence tokenizer that allows for unlimited lengths of text to be read, all while keeping proper intonation, abbreviations, decimals and more;
- Customizable text pre-processors which can, for example, provide pronunciation corrections;
- Automatic retrieval of supported languages.

Installation

- pip install gTTS

Module

- from gtts import gTTS
- tts = gTTS('hello')
- tts.save('hello.mp3')

Operating system: Raspbian (Debian)
Language: Python2.7
Platform: Pytesseract, OpenCV (Linux-library)
Library: OCR engine, Google TTS engine

The operating system under which the proposed project is executed is Raspbian which is derived from the Debian operating system. The program is written using the python language. The functions in algorithm are called from the

OpenCV Library. OpenCV is an open source computer vision library, which is written under C and C++ and runs under Linux, Windows and Mac OS X. OpenCV was designed for computational efficiency and with a strong focus on real-time applications. OpenCV is written in optimized C and can take advantage of multi-core processors.

6. FLOW OF PROCESS

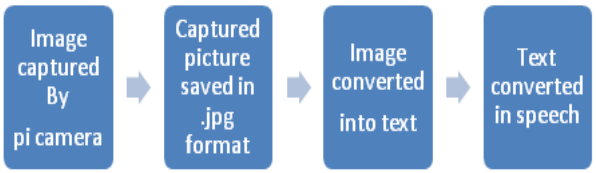

Figure.2 Flow of Process

6.1 IMAGE CAPTURING

The first step is the one in which the document is placed in front of the Picamera and the Picamera captures an image of the placed document. The quality of the image captured will be high so as to have fast and clear recognition due to the high-resolution camera.

6.2 IMAGE TO TEXT CONVERTER

Python-Tesseract is an optical character recognition (OCR) tool for python. That is, it will recognize and "read" the text embedded in images.

Python-Tesseract is a wrapper for Google's Tesseract-OCR Engine. It is also useful as a stand-alone invocation script to Tesseract, as it can read all image types supported by the Pillow and Leptonica imaging libraries, including jpeg, png, gif, bmp, tiff, and others. Additionally, if used as a script, Python-Tesseract will print the recognized text instead of writing it to a file.

6.3 TEXT TO SPEECH

gTTS (Google Text-to-Speech), a Python library and CLI tool to interface with Google Translates text-to-speech API. Write spoken mp3 data to a file, a file-like object (byte string) for further audio manipulation, or stdout. Or simply pre-generate Google Translate TTS request URLs to feed to an external program.

- Customizable speech-specific sentence tokenizer that allows for unlimited lengths of text to be read, all while keeping proper intonation, abbreviations, decimals and more;
- Customizable text pre-processors which can, for example, provide pronunciation corrections;
- Automatic retrieval of supported languages.

7. CONCLUSION

Text-to-Speech device can change the text image input into sound with a performance that is high enough and a readability tolerance of less than 2%, with the average time processing less than three minutes for A4 paper size. This portable device, does not require internet connection, and can be used independently by people. Through this method, we can make editing process of books or web pages easier. To extract text regions from advanced backgrounds, we've got projected a completely unique text localization formula supported models of stroke orientation and edge distributions. The corresponding feature maps estimate the worldwide structural feature of text at each component. Block patterns project the projected feature maps of a picture patch into a feature vector.

Adjacent character grouping is performed to calculate candidates of text patches ready for text classification. Associate degree Adaboost learning model is utilized to localize text in camera-based pictures. OCR is employed to perform word recognition on the localized text regions and rework into audio output for blind users. During this analysis, the camera acts as input for the paper. Because the Raspberry Pi board is high-powered the camera starts streaming. The streaming knowledge are going to be displayed on the screen victimization GUI application. Once the item for text reading is placed ahead of the camera then the capture button is clicked to produce image to the board.

Using Tesseract library the image are going to be born-again into knowledge and also the knowledge detected from the image are going to be shown on the standing bar. The obtained knowledge are going to be pronounced through the ear phones using Text-to-speech synthesis.

8. REFERENCES

[1] International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 05 Issue: 06 | June-2018 www.irjet.net p-ISSN: 2395-0072 © 2018, IRJET | Impact Factor value: 7.211 | ISO 9001:2008 Certified Journal | Page 1639 Raspberry Pi Based Reader for Blind People Anush Goel1, Akash Sehrawat2, Ankush Patil3, Prashant Chougule4, Supriya Khatavkar5 1Student, Department of Electronics Engineering, BVDU COE, Dhankawadi, Pune 2Student, Department of Electronics Engineering, BVDU COE, Dhankawadi, Pune 3,4,5Professor, Dept. of Electronics Engineering, BVDU COE, Dhankawadi, Pune, Maharashtra, India

[2] Ms.AthiraPanicker Smart Shopping assistant label reading system with voice output for blind using raspberry pi, Ms.Anupama Pandey, Ms.Vrunal Patil YTIET, University of Mumbai ISSN: 2278 – 1323

[3] International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Vol. 5, Issue 10, Oct 2016 2553 www.ijarcet.org ,Volume 7, Issue 4. April 2018. GSM based Message Reception for Visually Impaired Person. Supriya Kurlekar. (SITCOE,Yadrav). Prachi Herle.

[4] Dimitrios Dakopoulos and Nikolaos G.Bourbakis Wearable Obstacle Avoidance Electronic Travel Aids for Blind IEEE Transactions on systems, man and cybernetics, Part C (Applications and Reviews). Vol. 40, issue 1, Jan 2010.

[5] William A. Ainsworth A system for converting English text into speech IEEE Transactions on Audio and Electroacoustics, Vol. 21, Issue 3, Jun 1973

[6] Michael McEnancy Finger Reader Is audio reading gadget for Index Finger IJECCE Vol. 5, Issue 4 July-2014.

[7] N Giudice, G Legge, Blind navigation and the role of technology, in The Engineering Handbook of Smart Technology for Aging, Disability and Independence, AA Helal, M Mokhtari, B Abdulrazak, Eds. Hoboken, NJ, USA: Wiley, 2008

[8] Chen J Y, J Zhang, et al. Automatic detection and recognition of signs from natural scenes, IEEE Trans. Image Process., January 2004 ;13: 87–99.

[9] D Dakopoulos, NG Bourbakis, Wearable obstacle avoidance electronic travel aids for blind: A survey, IEEE Trans. Syst., Man, Cybern, January 2010; 40: 25–35.