A Smart Reader for Blind People

Mr. Vaibhav V. Govekar  
ME Student  
Department of Electronics & Telecommunication Engineering  
BSIOTR, Pune

Prof. Meenakshi A  
Assistant Professor  
Department of Electronics & Telecommunication Engineering  
BSIOTR, Pune

Abstract

This work introduces smart reader by using raspberry pi. The system consists of a camera interfaced with computer which receives the printed or handwritten text. This work proposes the novel implementation of smart book reader with raspberry pi controller. The system consists of a webcam interfaced with raspberry pi which accepts a page of printed text. The experimental results are done with the help of raspberry pi controller. Controller coding for the Raspberry pi is done through PYTHON language. The audio output is obtained after the captured image is converted to text. The image captured by the camera is converted to text and displayed in the form window and then the text is obtained as audio output. Raspberry pi has the audio port where the output can be heard through the headphone or the speaker. Once the image is converted to text raspberry pi takes few mille seconds to convert it as a voice output. This system is validated with both simulation and experimental verification it achieves the text document is converted with speech for the use of visually impaired people.

Keywords: Image processing, OCR, TTS, Raspberry pi, Python

I. INTRODUCTION

Visually impaired people fail to excess text using existing technology, including problems with alignment, focus, accuracy, mobility and efficiency. The proposed project uses the methodology of a camera based assistive device that can be used by people to read Text document. Most of the access technology tools built for people with blindness and limited vision are built on to basic building block of OCR software and text-to-speech (TTS) engines. Optical character recognition (OCR). Total Blindness = Visual acuity less than 3/60 digitize and reproduce texts that have been produced with non-computerized system. The storage space can be minimized.

II. PROBLEM STATEMENT

In this project, we are going to convert image to speech. The image in the form of scanned image or real time captured image is converted in text using OCR Optical Character Recognition. Then the text is converted into speech using TTS Text To Speech conversion. Optical character recognition (OCR) is the one of the simplest method used to converts (Electronic or mechanical) printed images or text, handwritten documents into machine-encoded text, whether it is scanned Document or a photo, a scene photo. The OCR API returns the extracted text, along with information about the location of the detected text in the original image back to the device app for further processing (such as text-to-speech) or display. Text-to-speech is the functions that convert the text into digital audio signal.

III. PROJECT BACKGROUND

The proposed project uses the methodology of a camera based device that can be used by people to read Text document. The framework is on implementing image capturing technique in a software system based on MATLAB. A Maximum number of the blind people use Braille for reading documents and books which are complicated to make and less readily available. This gives rise to the need for the improvement of devices that could bring relief to the excruciating tasks that the blind people have to go beginning to end.

IV. SYSTEM DEVELOPMENT

A. Image Acquisition:

A. In the First stage the device is in the moving on the printed page and the inbuilt camera which is attached to the device is captures the images of the text. The quality of the image captured will be high so as to have fast and clear recognition due to the high resolution camera.
B. Pre-processing:
Pre-processing stage consists of three steps: Twist Correction, Linearization and Noise removal. The captured image is checked for skewing. There are different possibilities of image getting twisted with either left or right orientation. Here the image is first enhanced and binaries. The function for twist detection checks for an angle of orientation and if the twist is detected then a simple image rotation is carried out till the lines match with the true horizontal axis, which produces a twist corrected image. The noise introduced during capturing or due to poor quality of the page has to be cleared before further processing.

C. Segmentation
After pre-processing, the filtered image is send to the next phase i.e. segmentation. In this phase the images are modified into sub-image of separate symbol (characters). Inter line spaces are tested by the binaries images. If inter line spaces are observed then the image is segregated into number of paragraphs across the interline gap. As per background the horizontal gap of lining is scanned properly. Histogram of the image is examining the actual width of the horizontal lines. Next step is to scan the line of vertical space intersection. Hence histogram helps to detect the width of the words. Then the words are segregated into characters using character width computation.

D. Feature Extraction
The input data objects are group into a set of features through feature extraction. The features extracted will help in extracting the related information from the input data in order to analyze the feature matching. This method is parallel to edge orientation histograms, scale-invariant feature convert descriptors, and shape contexts, but it is differ from above as compared to dense grid of uniformly spaced cells and it overlaps local contrast normalization for enhancing accuracy.

E. Image to Text Converter
The ASCII values of the recognized characters are processed by Raspberry Pi board. Here each of the characters is matched with its corresponding template and saved as normalized text transcription. This transcription is further delivered to audio output.

F. Text to Speech Converter
The character rectification is done in this stage. This is called as Character Recognition. The module performs the task of conversion of the transformed text to audible form. In this phase the raw text is converted to equivalent words. The raw text contains symbols like numbers and abbreviations. This process is called as text normalization. In the back end the synthesizer converts the symbolic representation into sound.

G. Simulation Environment
The image to text and text to speech conversion is done by the OCR software installed in PC. The conversion which is done in OCR can be simulated in MATLAB. The conversion process in MATLAB includes the following processes.
1) Binary image conversion.
2) Complementation.
3) Segmentation and labelling.
4) Isolating the skeleton of character.

H. Sample Image
The following image which is captured by the camera contains the following word. This image is in the jpeg format which has to be converted into text.

---

**Fig. 1: Character Detection and Recognition (OCR)**

Paragraphs across the interline gap. As per background the horizontal gap of lining is scanned properly. Histogram of the image is examining the actual width of the horizontal lines. Next step is to scan the line of vertical space intersection. Hence histogram helps to detect the width of the words. Then the words are segregated into characters using character width computation.

---

All rights reserved by www.ijste.org
I. Binary Conversion:

In this section sample image is converted into binary format. The image which was a 3D image initially is converted to 2D image.

Binary 0 represents black colour of the characters. Binary 1 represents white colour of the characters.

1) Boundary Marking:
The area of the text is bordered and the boundary for each character is isolated. The boundary for each character is programmed and varies from 0 to 255 bits of characters occupying memory in the database.

2) Segmentation and Labeling:
The separated blocks of characters are segmented and are automatically labeled for identity. Image segmentation converts the digital images into multiple segments such as set of pixels, which is well known as super pixels.

3) Forming Character Skeleton
For large preservation the connectivity of original reason the processes used is known as Skeletonization. This largely preserves connectivity of pervious section. While throwing away most of the original foreground pixels. For its working consider the foreground regions in the input binary image. This is made of uniform low-burning material. It starts fire sat all points along the boundary of this section and sees fire move into the interior. Fire extinguish where two differ boundaries meet and the points at which this happens form the so called "quench line"

4) Audio Output:
The programming codes are run in Python and corresponding output is generated. The output is in the form of audio. The audio is heard using headphone or speaker connected to the system. Each character of the word is spelled out first and then the entire word is read out.

V. HARDWARE REQUIREMENT

A. Raspberry Pi

The Raspberry Pi 3 Model B is the latest version of the $35 Raspberry Pi computer. The Pi isn't like your typical machine, in its cheapest form it doesn't have a case, and is simply a credit-card sized electronic board -- of the type you might find inside a PC or laptop but much smaller. The quad-core Raspberry Pi 3 is both faster and more capable than its predecessor, the Raspberry Pi 2. For those interested in benchmarks, the Pi 3's CPU--the board's main processor--has roughly 50-60 percent better performance in 32-bit more than that of the Pi 2, and is 10x faster than the original single-core Raspberry Pi (based on a multi-threaded CPU benchmark in SysBench). Compared to the original Pi, real-world applications will see a performance increase of between 2.5 x- for single-threaded applications--and more than 20x--when video playback is accelerated by the chip's NEON engine. Unlike its predecessor, the new board is capable of playing 1080p MP4 video at 60 frames per second (with a bitrate of about 5400Kbps), boosting the Pi's media center credentials. That's not to say, however, that all video will playback this smoothly, with performance dependent on the source video, the player used and bitrate. The Pi 3 also supports wireless internet out of the box, with built-in Wi-Fi and Bluetooth. The latest board can also boot directly from a USB-attached hard drive or pen drive, as well as supporting booting from a network-attached file system, using PXE, which is useful for remotely updating a Pi and for sharing an operating system image between multiple machines.
B. Python

Python is an easy-to-learn programming language that has some really useful features for a beginning programmer. The code is quite easy to read when compared to other programming languages. In addition to its simple language structure and an interactive shell with which to experiment, Python has some features that greatly augment the learning process. Python doesn’t have a lot of complicated symbols, like braces ({}), hashes (#), and dollar signs ($), which make other programming languages a lot more difficult to read. Python provides interfaces to all major commercial databases. Python supports GUI applications that can be created and ported to many system calls, libraries and windows systems, such as Windows MFC, Macintosh, and the X Window system of Unix. It provides a better structure and support for large programs than shell scripting.

C. Camera Module

The v2 Camera Module has a Sony IMX219 8-megapixel sensor (compared to the 5-megapixel OmniVision OV5647 sensor of the original camera). The Camera Module can be used to take high-definition video, as well as stills photographs. It’s easy to use for beginners, but has plenty to offer advanced users if you’re looking to expand your knowledge. There are lots of examples online of people using it for time-lapse, slow-motion, and other video cleverness. We can also use the libraries we bundle with the camera to create effects. We can read all the gory details about IMX219 and the Exmor R back-illuminated sensor architecture on Sony’s website, but suffice to say this is more than just a resolution upgrade: it’s a leap forward in image quality, color fidelity, and low-light performance. It supports 1080p30, 720p60 and VGA90 video modes, as well as still capture. It attaches via a 15cm ribbon cable to the CSI port on the Raspberry Pi. The camera works with all models of Raspberry Pi 1, 2, and 3. It can be accessed through the MMAL and V4L APIs, and there are numerous third-party libraries built for it, including the PiCamera Python library. See the Getting Started with PiCamera resource to learn how to use it.

VI. CONCLUSION

In the Smart Reader of visually impaired, we have proposed a simple and efficient algorithm for conversion of image to text. This method can effectively distinguish the object of interest from background or other objects in the camera view. OCR is used to perform word recognition on the localized text regions and transform into audio output for blind users.

ACKNOWLEDGEMENT

I am thankful to Prof. Meenakshi A for his valuable support and guidance, also thankful for making this project successful.

REFERENCES

[1] “Text to speech Device for Patients with Low Vision” AnnisaIstiqomahArrahmah, AuliaRahmatika, Samantha Harisa, HasballahZakariaRichard Mengko NOV 2015.
[2] “Portable Camera-Based Assistive Text and Product Label Reading from Hand-Held Objects for Blind Persons”, IEEE/ASME TRANSACTIONS ON MECHATRONICS, 2013
[3] “A Smart Reader for Visually Impaired People Using Raspberry Pi” -D.Velmurugan1, M.S.Sonam2S.Umamaheswari3, S.Parthasarathy4, K.R.Arun5 Assistant professor1, UG Student2, 3, 4, 5 MARCH 2016.
[4] “Camera Reading For Blind People”- Roberto Netou, NunoFonsecaa 2014.
[5] “IoT Based Data Processing for Automated Industrial Meter Reader using Raspberry Pi” Pratik D. Kute1, V. N. More2 Electronics and Telecommunication Department College of Engineering, Pune.