

Smart Reader – Text To Speech Converter

Using Raspberry Pi

Prof. Aparna Dhavlikar¹, Prachi H. Mane², Pooja B. Mhaske³, Priyanka M. Ovhal⁴

aparna.dhavlikar@mescoepune.org , manep047@gmail.com , pmhaske1234@gmail.com ,
priyankaovhal65@gmail.com

Electronics and Telecommunication Department, Modern Education Society's College Of Engineering, Pune

Abstract— Smart Reader allows user to hear the text which is given as input. It involves extraction of text from the image and converting the text to speech. This is done with Raspberry Pi and a camera module by using the OCR [optical character recognition] technique. The system consists of a webcam interfaced with raspberry pi. Raspberry pi has the audio port where the output can be heard through the headphone or the speaker. The conversion time aimed is few milliseconds. This device can help visually impaired persons to hear the text in images to be read.

Keywords— *Raspberry Pi, Optical Character Recognition.*

I) INTRODUCTION

The main part of the smart reader is optical character Recognition, which is abbreviated as OCR. It can translate the scanned images of hand written, typewritten text into the machine-encoded text. OCR makes it possible to edit the text, store it more compactly, display or print the text.

In OCR technique, first the image is converted in the black and white image. Then the image is cropped and each line is separated from row. The same procedure is repeated for each row to separate the character.

II) LITERATURE SURVEY

The table no.2.1 shown below describes the comparison of various techniques which were previously implemented. In this comparison different papers which were published by various authors have been studied and compared in tabular form considering techniques used.

Table 2.1: comparison of different techniques

NAME OF AUTHOR	TITLE	YEAR	INFORMATION
Falguni Patel, Dixita Patel, Talha Pathan, Anita Bindra Gujarat. (Journal Paper(IJSRE))	Smart Reading System	March 2017	This paper gives the idea of capturing of the image and processing or recognizing the text in Optical Character Recognition (OCR) system.
Velmurugan D, Srilakshmi, Umamaheswari S, Parthasarathy S, Arun K., Coimbatore, India. (IJAREEIE)	Smart Reader for Visually Impaired People Using Raspberry PI	March 2016	This paper addresses the integration of a complete Text Read-out system. The focus is on Implementing the image capturing technique on raspberry pi.
Mrunmayee Patil, Ramesh Kagalkar, University of Pune, India. (Research Article)	A Review on Conversion of Image to Text As Well As Speech Using Edge Detection and Image Segmentation	November 2014	The techniques of Image segmentation and Edge detection. The system goes through various phases such as preprocessing, feature extraction, object recognition, edge detection, and segmentation.

	n		
--	---	--	--

III) PROPOSED SYSTEM

Proposed system converts the input text to speech. For this we use raspberry pi kit, which is interfaced with the camera module, which captures the image. When the image is captured it is being processed with various operations with the help of Open CV software. Operations like morphological operations (erosion, dilation), color conversion, filtering is being done. After the image is processed, tessarect helps to convert the processed image in text. Then this text file is given as input to the espeak algorithm. And finally through earphone or speaker the output is heard.

a. BLOCK DIAGRAM

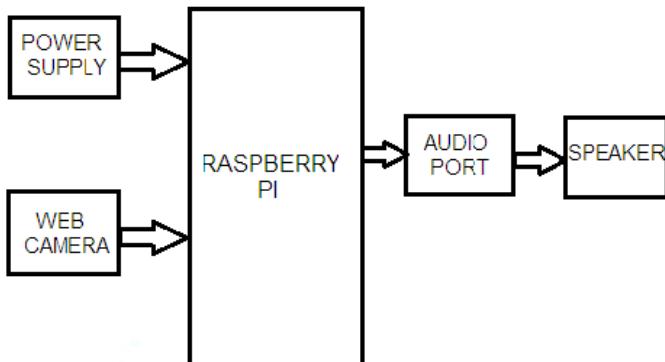


Fig 3.1: Block diagram

The web camera is connected to the USB port of raspberry pi. The raspberry pi has an OS named RASPION which process the conversions. The audio output is taken from the audio jack of the raspberry pi. The converted speech output is amplified using an audio amplifier, which is inbuilt.

The text to be heard is captured by focusing the camera on it. The captured image is processed by the OCR software installed in raspberry pi. The captured image is converted to text. The text is converted into speech by the TTS engine. The final output is given to the audio amplifier from which it is connected to the speaker. The speaker can also be replaced by a headphone for convenience.

IV) IMPLEMENTATION

a. Raspberry pi kit

We are using raspberry pi kit 3 because of its advanced features. The Raspberry Pi 3 Model B kit is of the third generation Raspberry Pi. It can be used for many applications. The Raspberry Pi 3 Model B has a more powerful processor, which is 10 times faster. Additionally it contains wireless LAN and Bluetooth.

Table 4.1: Specifications of raspberry pi kit.

SR NO .	PARAMETER	VALUE
1	Processor	1.2 GHz Quad core ARM Cortex-A53 802.11 b/g/n
2	Dimensions	85x56x17mm
3	Power	Micro USB socket 5v1,2.5A
4	Ethernet	10/100 BaseT Ethernet Socket
5	Audio Output	Audio Output 3.5mm jack
		USB 4 x USB 2.0 Connector
6	Camera	15-pin MIPI Camera Serial Interface (CSI-2)

The given Figure shows the raspberry pi module and its parts. Fig 3.1 contains kit module.



Fig 4.1: Raspberry Pi Kit

Fig 4.2 contains parts of raspberry pi kit.

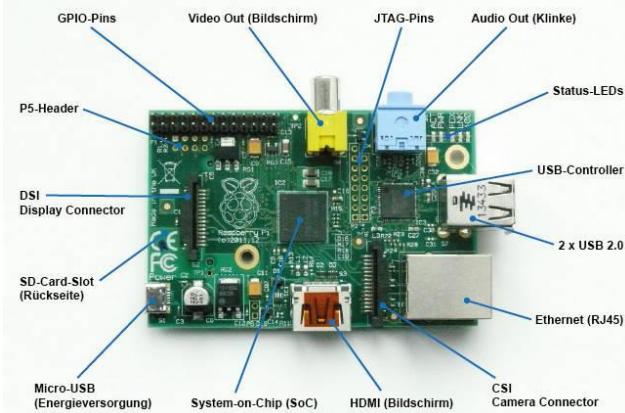


Fig 4.2 parts of kit

b. Web camera:

The camera module is connected with the camera serial interface of the raspberry pi using the USB. This helps in capturing a 5 MP resolution image.

The Figure shows a camera module can be connected via a USB cable to the kit.

The specifications of raspberry pi camera module are: 5MP of image resolution supports video and still images.



Fig 4.3: Camera Module.

c. Speaker:

Speaker is used to make the text audible to us after conversion.



Fig 4.4: Speaker module

d. Methodology

At input, coloured image is captured; the coloured image is converted into the gray scale image through RGB to Gray converter. Various morphological operations such as erosion, dilation is performed on the image. Here we get noiseless image. The image is converted into black and white image through thresholding. These all steps are done with the help of Open CV. This image is converted into text file through tessarect algorithm. It is given as input to the espeak engine. And the speech can be heard by the headphones or speaker.

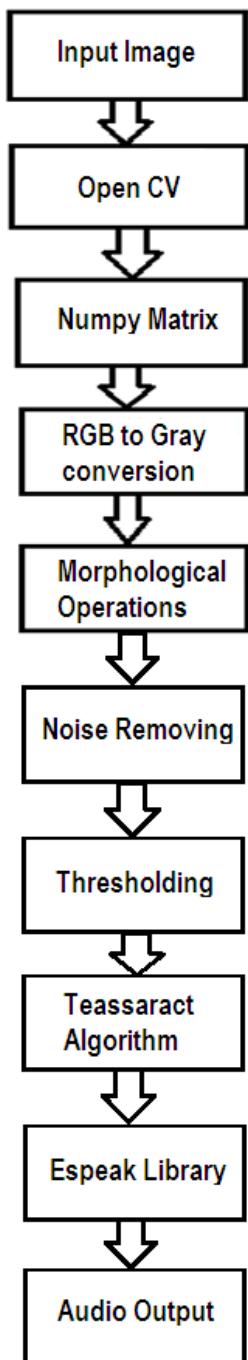


Fig 4.5. Software block diagram

e. Flow diagram

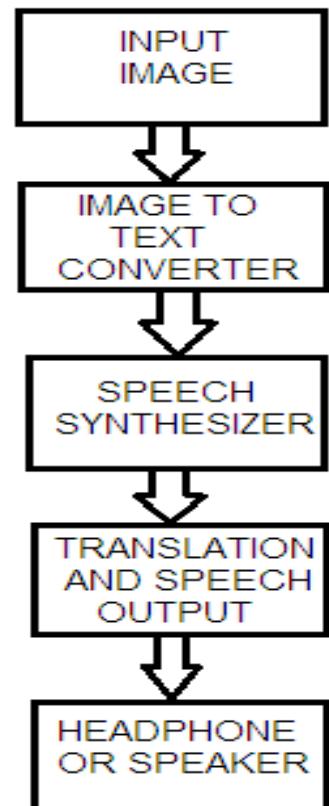


Fig 4.6 : Flow Diagram

INPUT IMAGE:

The input image is captured by camera.

IMAGE TO TEXT CONVERTER:

The image which is to be read is captured by the camera module. The tesseract algorithm is installed in raspberry pi. This algorithm is used to convert the image file to text file and storing it with .txt extension.

SPEECH SYNTHESIZER:

It is a software used to synthesize speech from text. A TTS engine convert's written text to speech .espeak is a software which can be easily used in raspberry pi by installing espeak engine.

SPEAKER OUTPUT:

The output is heard by connecting the speakers to the audio jack. The translated speech output can be heard through the speakers or headphones.

V) ANALYSIS AND RESULT



Fig.5.1 Input Image



Fig.5.2 Color Converted image (Removed Noise)



Fig 5.3 Output image(Thresholded image)

Figure 5.1 shows that the captured image is in RGB form. The image is converted in the gray scale form by using RGB to Gray conversion, morphological operations. After these image processing we get noiseless image ,which is shown in figure 5.2. Thresholding technique is used to get the final output image, which is figure 5.3. This thresholded image is given to the tessarect algorithm,which will convert it into the text file.and this text file is given as the input to the espeak engine and then the outut sound is heard.

VI) CONCLUSION

We have proposed an image to speech converter using raspberry pi kit. The device is compact and helpful to the society. This device converts input text into the sound. This system can convert the text to speech in approximately 8-10 milliseconds.

Future enhancements can be made in this system, which will help to convert the input text in different languages or in native languages.

BIBLIOGRAPHY

[1] Nirmala Kumari, Meghana Reddy J (May 2016), "Image Text to Speech Conversion Using OCR Technique in Raspberry Pi"International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (IJAREEIE), ISSN (Print): 2320 -3765, ISSN (Online): 2278 - 8875, Vol. 5, Issue 5.

[2] Gopinath, aravind , pooja et.al.."Text to speech conversion using matlab"International journal of emerging technology and advanced engineering. volume 5, issue 1, (January 2015).

[3]Rithika.H, B.Nithya Santoshi "Image to text converter using raspberry pi"2016 IEEE.

[4]Velmurugan.d,Srilakshmi,Umamaheshwari.S,Parthasarathy .S,Arun K.R."Hardware implementation of smart reader for visually imapaired people using raspberry pi",Research article.,IJESC,volume 6 issue no.3.