



## REPLICA TO SPEECH RECLAMATION FOR VISUALLY IMPAIRED

**Himangi A. Badwaik**

Electronics and Telecommunication Engineering  
Prof.Ram meghe institute of technology and  
research,Badnera,amravati  
Amravati,Maharashtra,India  
himangi.badwaik27@gmail.com@gmail.com

**Gunjan R. Balpande**

Electronics and Telecommunication Engineering  
Prof.Ram meghe institute of technology and  
research,Badnera,amravati  
Amravati,Maharashtra,India  
gunjanbalpande1805@gmail.com

**Prof. A. I. Rokade**

Electronics and Telecommunication Engineering  
Prof.Ram meghe institute of technology and  
research,Badnera,amravati  
Amravati,Maharashtra,India  
rokadeashay@gmail.com

**Shraddha N. Nagrale**

Electronics and Telecommunication Engineering  
Prof.Ram meghe institute of technology and  
research,Badnera,amravati  
Amravati,Maharashtra,India  
shraddhanagrale93@gmail.com

---

### Abstract-

*The present paper has introduced an innovative, efficient and real time cost beneficial technique that enables user to hear the contents of text images instead of reading through them. Visual impairment is one of the biggest limitation for humanity, especially in this day and age when information is communicated a lot by text messages (electronic and paper based) rather than voice. The device we have proposed aims to help people with visual impairment. In this project, we developed a device that converts an image's text to speech. The basic framework is an embedded system that captures an image, extracts only the region of interest (i.e. region of the image that contains text) and converts that text to speech. It is implemented using a Raspberry Pi and a Raspberry Pi camera. The captured image undergoes a series of image pre-processing steps to locate only that part of the image that contains the text and removes the background. Two tools are used convert the new image (which contains only the text) to speech. They are OCR (Optical Character Recognition) software and TTS (Text-to-Speech) engines. The audio output is heard through the raspberry pi's audio jack using speakers or earphones.*

**Keywords:** OCR, image pre-processing, embedded system, raspberry Pi, TTS, text extraction, voice processing.

### INTRODUCTION

The present paper has introduced an innovative, efficient and real time cost beneficial technique that enables user to hear the contents of text images instead of reading through them. Visual impairment is one of the biggest limitation for humanity, especially in this day and age when information is communicated a lot by text messages (electronic and paper based) rather than voice. Reading is very important and one aspect in our day today lives. Almost 314 million visually impaired people are there all around the universe [1], 45 Million are blind and new cases being added each year as per researches done. The device we have proposed aims to help people with visual impairment. In this project, we developed a device that converts replica i.e. an image's text to speech. Emerging technologies and recent developments in computerized vision, digit cameras, PDA and portable computers make it flexible to assist these individuals by developing image-based products that combine computer vision technology with other existing commercial technology such as optical character recognition (OCR) platforms. Printed text is one of the forms of reports, receipts, bank statements, restaurant menus, classroom hand-outs, product packages, medicine bottles, banners on road etc. There are many favouring systems available currently but they have little issues in reducing the flexibility for the visually impaired persons. The basic framework is an embedded system that captures an image, extracts only the region of interest (i.e. region of the image that contains text) and reclamation means converts that text to speech. It is implemented using a Raspberry Pi and a Raspberry Pi camera. The captured image undergoes a series of image pre-processing steps to locate only that part of the image that contains the text and removes the background. Two tools are used convert the new image (which contains only the text) to speech. They are OCR (Optical Character Recognition) software and TTS (Text-to-Speech) engines. The audio output is heard through the raspberry pi's audio jack using speakers or earphones.

## LITERATURE REVIEW

[2] In our planet of 7.4 billion humans, 285 million people are visually impaired and out of whom 39 million people are completely blind, i.e. have no vision at all. And 246 million have served visual impairment (according to the WHO). Visual impairment or vision loss is defined as the decreased ability to see clearly and cannot be fixed using glasses. Blindness is the term used for complete vision loss. The common causes of vision loss are uncorrected refractive errors, cataracts and glaucoma. People with visual impairment face a number of difficulties in normal daily activities like walking, driving and reading.

[3] Reading is important in today's life. Printed text is everywhere in the form of receipts, bank documents, reports, books. There are many systems done text to speech conversion, but they can't handle product labelling. But big limitation is use of this is difficult for blind people. This paper proposed camera based text reading to help blind peoples to read the product label. Main prototype of this system is 1) Novel based algorithm to solve aiming problem. 2) Novel algorithm of automatic text localization to remove the background of the text image. 3) Camera based framework used for binned people to read text.

[4] In the development of OCR of this script is difficult because of large number of characters have to be recognized. This proposed system the digitalized document image is first passed through pre-processing modules like skew correction, line

segmentation, zone detection, word and character segmentation, etc. hence the modules have been developed by combining some conventional techniques with some newly proposed ones.

[5] Proposed that this project introduced smart device that useful for visually impaired peoples. This project uses the technology of camera based device that can be useful for blind people to read documents. The work is implementing capturing image technique in an embedded system. In this system has a camera as an input device to detect the printed text for reorganization and that scanned text is process by a software OCR.

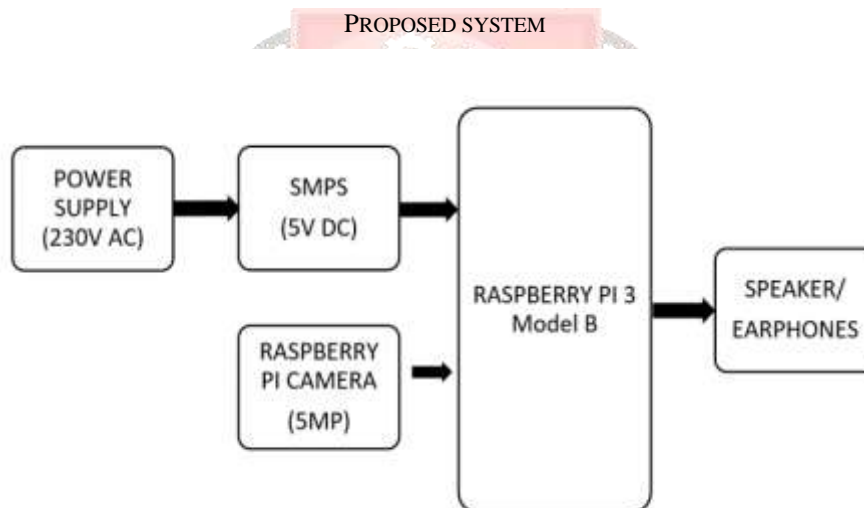


Figure No.1.1 Block Diagram of Speech Converter from Text Images For Blind

Raspberry pi use to detect the text from image using Open CV python programming OCR ENGINE The extraction of the text in the image is done using optical character recognition (OCR). OCR is a field of research in pattern recognition, artificial intelligence and computer vision. It is the conversion of the images of typed, handwritten or printed text into a digital text or computer format text. Earlier OCR versions had to be trained in each character of a text with its specific font. Today, advanced OCRs are available that have a high degree of accuracy, support a wide variety of image formats, languages and fonts. For our project, we have used Tesseract OCR. It is the most accurate open source OCR engine and is powered by Google. It can be used on the Linux, mac and windows platforms.

### 1. OCR ENGINE

The extraction of the text in the image is done using optical character recognition (OCR). OCR is a field of research in pattern recognition, artificial intelligence and computer vision. It is the conversion of the images of typed, handwritten or printed text into a digital text or computer format text. Earlier OCR versions had to be trained in each character of a text with its specific font. Today, advanced OCRs are available that have a high degree of accuracy, support a wide variety of image formats, languages and fonts. For our project, we have used Tesseract OCR. It is the most accurate open source OCR engine and is powered by Google. It can be used on the

Linux, mac and windows platform. The newest Tesseract version, 3.4 supports a hundred languages. However, images must undergo a number of pre-processing stages like noise removal, scaling etc. otherwise the output will be of low quality.

## 2. TTS SOFTWARE

The process of converting text to speech by a computer is called speech synthesis. A text to speech system(TTS) is used to perform speech synthesis. A TTS is composed of two parts: front end and back end. The front end converts the text to a symbol, for example, a number. Each symbol generated is assigned a phonetic. The back end then converts the phonetic into sound. In our project, we have used Festival TTS. Festival is the most widely used open source TTS. It has a wide variety of voices and support English, Spanish and welsh language. We have used the English language

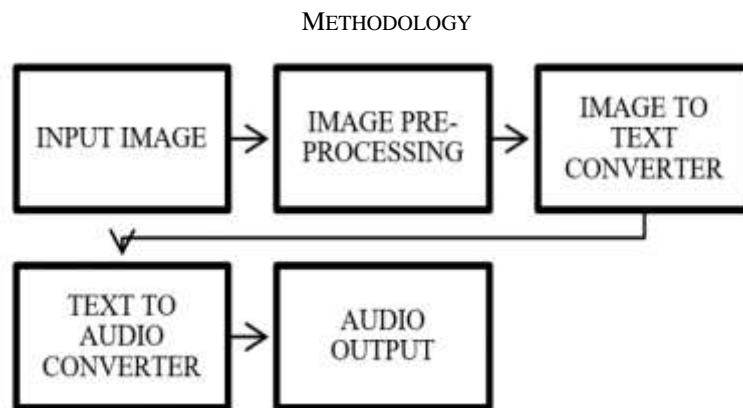


Figure No.1.2 Methodology Block Diagram

**Image Acquisition:-** In this step, the inbuilt camera captures the images of the text. The quality of the image captured depends on the camera used. We are using the Raspberry Pi's camera which 5MP camera with a resolution of 2592x1944.

**Image Pre-Processing:-** This step consists of colour to gray scale conversion, edge detection, noise removal, warping and cropping and thresholding. The image is converted to gray scale as many Open CV functions require the input parameter as a gray scale image. Noise removal is done using bilateral filter. Canny edge detection is performed on the gray scale image for better detection of the contours. The warping and cropping of the image are performed according to the contours. This enables us to detect and extract only that region which contains text and removes the unwanted background. In the end, Thresholding is done so that the image looks like a scanned document. This is done to allow the OCR to efficiently convert the image to text. **Image to text conversion:** The above diagram (fig.3) shows the flow of Text-To-Speech. The first block is the image pre-processing modules and the OCR. It converts the pre-processed image, which is in .png form, to a .txt file. We are using the

Tesseract OCR. **Text to speech conversion:** The second block is the voice processing module. It converts the .txt file to an audio output. Here, the text is converted to speech using a speech synthesizer called Festival TTS. The Raspberry Pi has an on-board audio jack, the on-board audio is generated by a PWM output.

## ACKNOWLEDGMENT

The Inspiration and guidance are valuable in aspects of life, especially what is academics. "Experience is the best teacher", is an old age. the satisfaction and pleasure that accompany the gain of experience would be incompletes without mentioning the people the people who made it possible. It gives us immense pleasure in submitting the report paper on "IMPROVE RESCUE SYSTEM FOR BOREWELL ACCIDENTS OF INNOCENTS" we take this opportunity to express my sincere gratitude to our supervisor Prof A.I. Rokade Department of Electronics and telecommunication engineering, PRMIT&R, Badnera. For the continuous support of our study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped us in all the time of research and writing of this paper. We would also like to thank to our Head of Department Dr. S.M. Deshmukh for his constant support and encouragement. We are also very thankful to all staff members of Electronics and Telecommunication Engineering, PRMIT&R, Badnera. Whose encouragement & suggestions help us to complete my project work. At last, we are thankful to our friends whose encouragements & constant inspirations helped us to complete this project work verbally & theoretically.

## CONCLUSION

The system enables the visually impaired to not feel at a disadvantage when it comes to reading text not written in braille. The image pre-processing part allows for the extraction of the required text region from the complex background and to give a good quality input to the OCR. The text, which is the output of the OCR is sent to the TTS engine which produces the speech output. To allow for portability of the device, a battery may be used to power up the system. The future work can be developing devices that perform object detection and extracting text from videos instead of static images.

## REFERENCES

- [1] World Health Organization. (2009). 10 facts about blindness and visual impairment [Online]. [www.who.int/features/factfiles/blindness/blindness\\_facts/en/index.html](http://www.who.int/features/factfiles/blindness/blindness_facts/en/index.html).
- [2] World Health Organization. 10 facts about blindness and visual impairment. 2015. Available from: [http://www.who.int/features/factfiles/blindness/blindness\\_facts/en/](http://www.who.int/features/factfiles/blindness/blindness_facts/en/)
- [3] T.Rubesh Kumar, C.Purnima "Assistive System for Product Label Detection with Voice Output For Blind Users" International Journal of Research in Engineering & Advanced Technology 2014.
- [4] Velmurugan.V, "A Smart Reader for Visually Impaired People using Raspberry PI", ISSN 2321 3361, Vol.6.
- [5] Anusha Bhargava, Karthik V. Nath, Prithish Sachdeva and Monil Samel "Reading Assistant for the Visually Impaired" International Journal of Current Engineering and Technology, 2015.
- [6] Raspberry pi cook book by simon monk.
- [7] R. Smith. "An Overview of the Tesseract OCR Engine", USA: Google Inc.
- [8] [http://elinux.org/RPi\\_Text\\_to\\_Speech\\_\(Speech\\_Synthesis\)](http://elinux.org/RPi_Text_to_Speech_(Speech_Synthesis))

