

Image Text To Speech Conversion by Translating With Raspberry Pi

¹Shubham Machale, ²Aadityaa Yedshikar, ³Pratik Sahastrabuddhe, ⁴Shashikant Suranje, ⁵Prof.Mrunali Umale
Department of Computer Engineering, Zeal College Of Engineering and Research, Narhe, Pune.

Abstract: - The principle issue in correspondence is dialect predisposition between the communicators. This gadget fundamentally can be utilized by individuals who don't know English and need it to be meant their local dialect. The curiosity part of this exploration work is the discourse yield which is accessible in 53 distinct dialects interpreted from English. This paper depends on a model which causes client to hear the substance of the pictures in the coveted dialect. It includes extraction of from the picture and changing over the two deciphered discourse in the client wanted dialect. This is finished with Raspberry Pi and a camera module by utilizing the ideas of OCR [optical character recognition] motor, Google Speech API [application program interface] which is the Text to discourse motor and the Microsoft interpreter. This calms the explorers as they can utilize this gadget to hear the English in their own particular wanted dialect. It can likewise be utilized by the outwardly disabled. This gadget causes clients to hear the pictures being perused in their coveted dialect.

Keywords: - Raspberry Pi; Tesseract OCR engine; Google Speech API; Microsoft translator; Raspberry Pi camera board.

I. INTRODUCTION

There are as of now numerous frameworks which read pictures and give voice yield [2] [3] [4]. Be that as it may, this framework gives voice yield in any dialect wanted by the client. This is finished by catching the picture which is to be perused utilizing a raspberry pi camera module [7]. Raspberry pi [7] is a Mastercard measured single board PC. The working framework utilized is Raspbian. A 15 cm lace link is utilized to join the camera module to the raspberry pi. The coding is finished utilizing python dialect. The Optical character acknowledgment motor changes over the pictures of content into machine encoded content and spare it in a content record. Tesseract is the OCR motor which is utilized for separating the English content from the picture and putting away it in a content record [4]. The content to discourse motor proselytes content to discourse yield [4]. eSpeak is a discourse synthesizer which can without much of a stretch be utilized as a part of raspberry pi for discourse yield in English. For making an interpretation of it to different dialects Google content to discourse motor and Microsoft interpreter is utilized [5]. Google content to discourse is a screen peruser which talks the content on the screen. Microsoft interpreter is a multilingual measurable machine interpretation cloud benefit gave by Microsoft. It bolsters 53 distinctive dialect frameworks [6]. This interpreted discourse yield could be heard through speakers or head set. The stage being utilized for reproduction of this model is putty in SSH (Secure Socket Shell). Putty is a free and open-source terminal emulator

which is utilized to offer charges to the raspberry pi. By and large, it is finished utilizing MATLAB [2] yet it is distinctive here in light of the fact that interpretation module is an additional element which is impossible utilizing MATLAB..

II. RELATED WORK

V. Ajantha Devi, Dr. S Santhosh Baboo[1] In this study, we propose an optical character recognition technique using Intro Sort. Main feature of this proposed technique is that we segment images using intro sort. It reduces the comparison time for matching the pixels of an image. It reflects reduction in OCR time. Intro sort algorithm begins with quick sort and when recursion depth exceeds a level it switches to heap sort, based on the number of pixels being sorted. This approach also has advantage of recognizing number plates and text documents in very nominal time. Our approach is able to extract characters of different font sizes. Our technique is performed well in noisy images too.

Yusuf Khan , Kapil Kumar Gupta , Namrata Dhanda[5]As per theWorld Health organization (WHO), 285 million people are estimated to be visually impaired worldwide among which 90 percent live in developing countries and 45 million blind individuals world-wide Even though there are many existing solutions to the problem of assisting individuals who are blind to read, none of them provide a reading experience that in any way equivalent to that of the sighted population. In specific, there is a need for a portable text reader that is reasonable and readily existing to the blind community. Addition to the specially enabled in the IT revolution is both a social responsibility as well as a computational test in the rapidly advancing digital world today. This work suggests a smart reader for blind people using raspberry pi. This paper lectures the assimilation of a complete Text Read-out system designed for the blind. The system consists of a webcam interfaced with raspberry pi which accepts a page of printed text. The OCR (Optical Character Recognition) package.

R.Aravind , R.P.Jagadesh , M.Sankari ,N.Praveen, Mrs. S.Arokia Magdaline[6] Text Detection from Natural Scene Images: Towards a System for Visually Impaired PersonsWe propose a system that reads the text encountered in natural scenes with the aim to provide assistance to the visually impaired persons. This paper describes the system design and evaluates several character extraction methods. Automatic text recognition from natural images receives a growing attention because of potential applications in image retrieval, robotics and intelligent transport system. Camera-based document analysis becomes a real possibility with the increasing resolution and availability of digital cameras

Raja Venkatesan.T, M.Karthigaa[2]In this paper an innovative, efficient and real-time cost beneficial technique that enables user to hear the contents of text images instead of reading through them as been introduced. It combines the

concept of Optical Character Recognition (OCR) and Text to Speech Synthesiser (TTS) in Raspberry pi. This kind of system helps visually impaired people to interact with computers effectively through vocal interface. Text Extraction from colour 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. This paper describes the design, implementation and experimental results of the device. This device consists of two modules, image processing module and voice processing module. The device was developed based on Raspberry Pi v2 with 900 MHz processor speed.

Raja Venkatesan.T, M.Karthigaa, P.Ranjith, C.Arunkumar,[3]A Majority of the visually impaired use Braille for reading documents and books which are difficult to make and less readily available. This gives rise to the need for the development of devices that could bring relief to the agonizing tasks that the visually impaired has to go through. Due to digitization of books there are many excellent attempts at building a robust document analysis system in industries, academia and research labs, but this is only for those who are able to see. This project aims to study the image recognition technology with speech synthesis and to develop a cost

effective, user friendly image to speech conversion system with help of Raspberry Pi. The project has a small inbuilt camera that scans the text printed on a paper, converts it to audio format using a synthesized voice for reading out the scanned text quickly translating books, documents and other materials for daily living, especially away from home or office. Not only does this save time and energy, but also makes life better for the visually impaired as it increases their independency. Performance of KDDC is compared with the most prevailing kernels.

III. PROPOSED SYSTEM

We will centre on separating highlights from pictures of a content and change over them in to sound. Filtered pages of content will prepare the classifier while pictures taken by camera will change over that picture content into sound. We plan to accomplish more writing research on archive picture recovery, join and enhance them by keeping memory require little in the meantime. We will probably find the ideal exchange off between preparing speed and coordinating exactness, and to make a showing of this idea on raspberry pi. The camera would take a photo of content and naturally play back the applicable time stamp of its relating book recording.

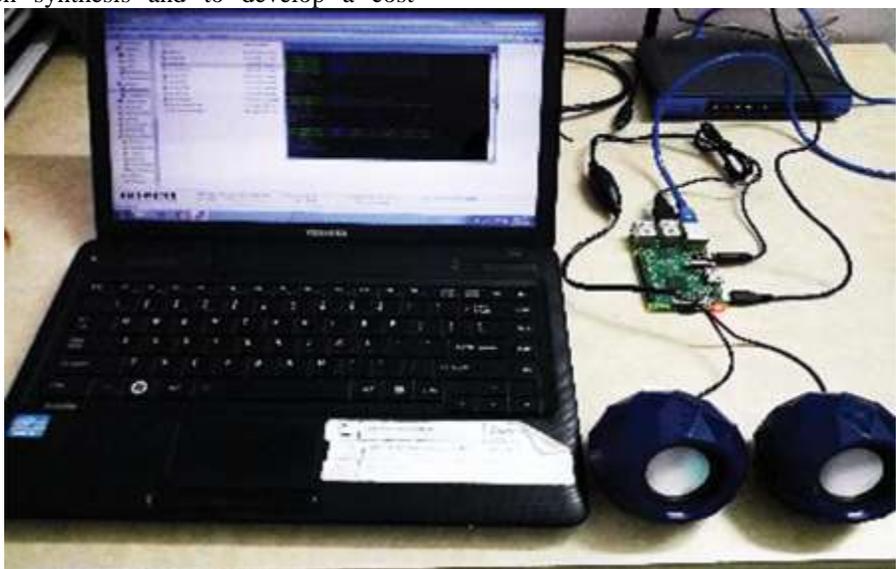


Figure 1: Sample UI

System Hardware Design

The hardware consists of the following parts: Raspberry pi camera module, Raspberry pi 3 [model B] mounted with SD card, speakers, Internet connection via Ethernet or Wi-Fi, laptop. The Fig 2 gives the block diagram of system hardware design.

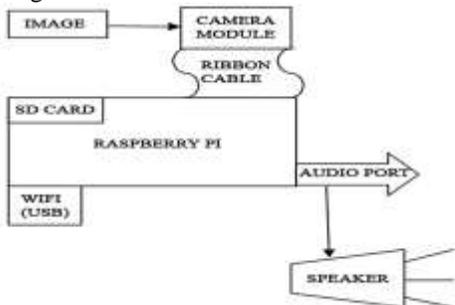


Fig2: H/W Design

System Software Design

The framework programming configuration comprises of different stages which help is delivering the final product. A machine comprehends the content of the picture and gives a voice yield with the preparing of these stages. The different stages are: Image to content transformation (OCR), Text to discourse and Microsoft interpreter. Fundamentally, the camera catches the picture and stores it as a picture record with .jpg expansion. The OCR motor [3] changes over it from picture record to content document by removing the numbers and characters of English letter set just gave the content is printed. It can't perceive manually written writings. It is then changed over to a .flac record by running an order utilizing eSpeak [3], a content to discourse motor. The flac document is given as a contribution to a python program

which gives a deciphered discourse and content yield utilizing Google content to discourse motor and Microsoft interpreter [5]. The Fig. 3 gives the piece chart of framework programming outline.

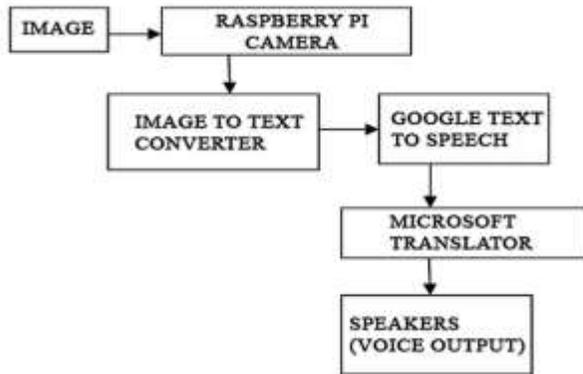


Fig 3: S/W Design

IV. CONCLUSION

In this work, Subsequent to examining the advancement pattern of current versatile human services innovation, this article shows other portable medicinal services demonstrate in view of distributed computing. This portable application can be gotten to and information can be shared crosswise over gadgets utilizing cloud.

REFERENCES

- [1] V. Ajantha Devi, Dr. S Santhosh Baboo (Jul-Aug 2014), "Optical Character Recognition on Tamil Text Image Using Raspberry Pi" International Journal of Computer Science Trends and Technology (IJCT) – Vol. 2 Issue 4.
- [2] Raja Venkatesan.T, M.Karthigaa, P.Ranjith, C.Arunkumar, M.Gowtham, "Intelligent Translate System for Visually Challenged People" International Journal for Scientific Research & Development (IJSRD), ISSN (online): 2321-0613, Vol. 3, Issue 12, 2016.
- [3] Anusha Bhargava, Karthik V. Nath, Pritish Sachdeva, Monil Samel (April 2015), "Reading Assistant for the Visually Impaired" International Journal of Current Engineering and Technology (IJCET), E-ISSN 2277 – 4106, P-ISSN 2347 – 5161, Vol.5, No.2.
- [4] K 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
- [5] Yusuf Khan , Kapil Kumar Gupta , Namrata Dhanda , "Optical Character Recognition on Tamil Text Image Using Raspberry Pi" International Journal of Computer Applications (0975 – 8887) Volume 127 – No.1, October 2015
- [6] R.Aravind , R.P.Jagadesh , M.Sankari ,N.Praveen, Mrs. S.Arokia Magdaline," Neural Network Based Assistive System For Text Detection With Voice Output" International Research Journal of Engineering and Technology (IRJET) Volume: 04 Issue: 04 Apr -2017