Voice Assisted Text Reading system for Visually Challenged Persons

[1]Reddi Swapna, [2]Dr. D. Lalitha Bhaskari

[1]M.Tech Scholar, Department of Computer Science and System Engineering,

[2]Assistant Professor, Department of Computer Science and System Engineering,

Andhra University College of Engineering (A), Visakhapatnam, AP, India.

Abstract: According to world health organization there are 91% peoples are visually challenged people. It is observed that they are still finding it difficult to roll their day to day life and they are difficult to find the medicinal tablet covers on their own. So we propose a technique for visually challenged people or differently abled persons to check whether they are taking a right tablet or not on medical tablet covers. This can help the visually challenged persons to read a printed text on medical strips on strips in vocal form. A web camera is used to capture the text or tablet covers image from the printed text and the captured image is analysed using Tesseract-Optical Character recognition (OCR). The detected text is then converted into speech using a compact open source software speech synthesizer, openCV. Finally, the synthesized speech is produced by the headphone by TTS method. The TTS (Text To Speech) this technology is basically used for conversion of text or image file into voice or in audio form. This TTS technology proposed to help the visually challenged peoples. It includes camera to capture image as a input text, which is then pass to TTS unit. TTS unit installed in raspberry-pi and the output of TTS is amplified by using audio amplifier and then it given to the speaker. The concept of proposed system is the idea of developing a mechanism where in machine will read the characters on medical tablet boxes which will be converted into speech thus enabling visually challenged persons to know the tablet names which they are buying or using. There are three modules in this system: feature extraction, image processing module and voice processing module. Based on these modules, the characters are converted to audio. All capture methods found in OpenCV (capture, capture continuous, capture sequence) has to be considered according their use and abilities.

Keywords: **Optical character recognition (OCR), Text to Speech (TTS), Open CV, Raspberry pi 3 model B, image Recognition, Web Camera, Speakers.**

INTRODUCTION

Visually challenged people fail to excess text using existing technology, including problems like medical tablet covers, books on their own. The proposed system uses the methodology of a camera based assistive device that can be used by visually challenged people to read Text on tablet covers. Most of the access technology tools built for people with visually challenged and limited vision is built on to basic building block of OCR software and text-to-speech (TTS) engines. This project has represented the innovative idea as well as low cost technique that is used to hear the contents of the tablet image without reading them. It combines the concept of optical character recognition (OCR) and Text to speech synthesizer (TTS) in raspberry pi. This system used to help the visually challenged people to interact with computer effectively through vocal interface. Text extraction from colour image is very difficult task for computer the text to speech conversion system is read the English alphabets and numbers that are in the image using the OCR technique and convert it into the voice format. This paper presents the design implementation and experimental result of the device. This device consist of three modules: feature extraction, image processing module and voice processing module. The Optical Character Recognition (OCR) is the process that converts the scan or printed text images into the text format for the further processing. This paper has presented the simple approach for text extraction and its conversion into speech. The testing of devices was done on raspberry pi 3 model B device. Text to speech (TTS) system produces the more natural voice that can be closely matched with human voice. The example of the speech synthesis are the voice enabled e-mail and messaging.

RASPBERRY-PI

The Raspberry-pi is a small size computer which does not have hard disk as there in PC but has an SD card of 4GB or above for storage purpose. In this project Raspberry pi 3 model B is used. The Raspberry Pi 3 Model B is the latest version of the Raspberry Pi computer. The quad-core Raspberry Pi 3 is both faster and more capable than its predecessor, the Raspberry Pi 2. the Pi 3's CPU--the board's main processor--has roughly 50%-60% better performance in 32-bit more than that of the Pi 2, and is 10x faster than the original single-core Raspberry Pi. 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.

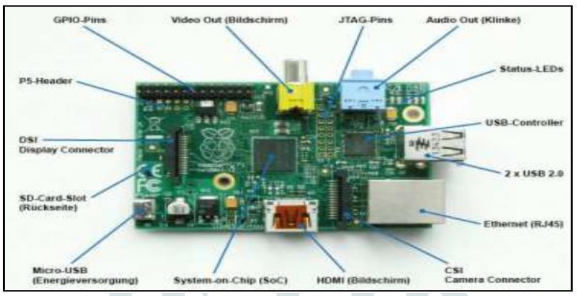

Fig.1 Architecture of Raspberry PI 3 Model B

PROPOSED SYSTEM

The main aim of this project is to convert text to speech. To do this we use optical character recognition. Optical character recognition is conversion of images of typed, handwritten or printed text into machine-encoded text from image. Here we come up with system where pattern of the image is extracted and image text is recognized using Optical character recognition.

In this project, we are going to convert image to speech on tablet covers. 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. The concept of proposed system is the idea of developing a mechanism where in machine will read the characters on medical tablet covers which will be converted into speech. There are two modules in this system: image processing module and voice processing module. This explains the text reading system for visually challenged people for their self-independent. 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.

Block diagram

This paper has presented the simple approach for text extraction and its conversion into speech using raspberry pi connected with web camera, speaker, HDMI to VGI.

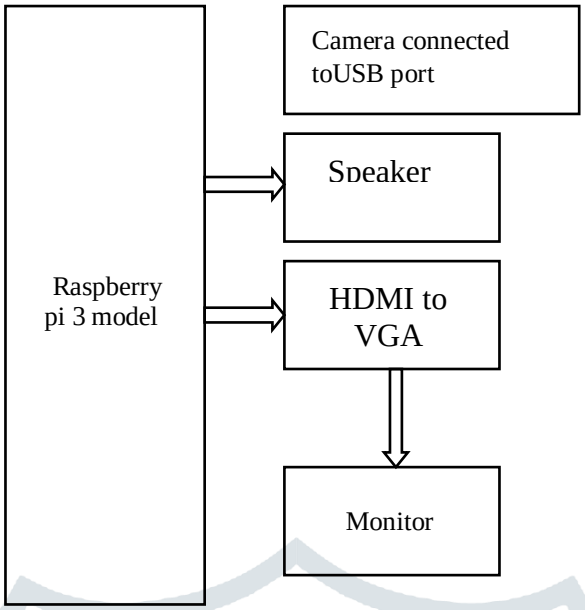

Fig.2. Block Diagram of Text to Speech

METHODOLOGY

Image Capturing: The first step is the one in which the tablet is placed under the camera and the camera captures an image of the placed tablet. The quality of the image captured will be high so as to have fast and clear recognition due to the high-resolution camera.

Pre-Processing

The pre-processing stage consists of three steps: Skew Correction, Linearization, and Noise Removal. The captured image is checked for skewing. There are possibilities of the image getting skewed with either left or right orientation. Here the image is first brightened and binarized.

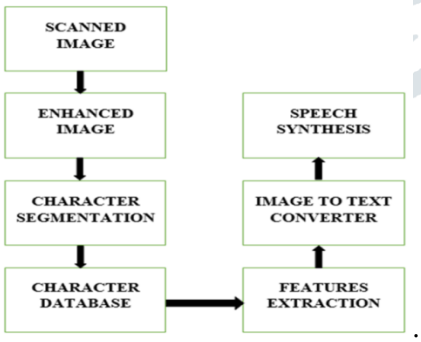

Fig.3: Flow of Process

The function for skew detection checks for an angle of orientation between ±15 degrees and if detected then a simple image rotation is carried out till the lines match with the true horizontal axis, which produces a skew corrected image. The noise introduced during capturing or due to the poor quality of the page has to be cleared before further processing.

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 the audio output.

Text to Speech

The scope of this module is initiated with the conclusion of the receding module of Character Recognition. The module performs the task of conversion of the transformed text to audible form. The Raspberry Pi has an on-board audio jack, the on-board audio is generated by a PWM output and is minimally filtered. A USB audio card can greatly improve the sound quality and volume.

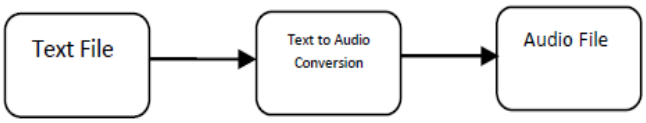

Fig.4. Text to Speech

As the recognition process is completed, the character codes in the text file are processed using Raspberry Pi device on which recognize a character using Tesseract OCR algorithm and python programming, the audio output listens. Text-to-speech (TTS) is a type of speech synthesis application that is used to create a spoken sound version of the text in a computer document, such as a help file or a Web page. TTS can enable the reading of computer display information for the visually challenged person, or may simply be used to augment the reading of a text message. Speech output requires a few audio software packages to be installed on your RPi. They may be already there but it does no harm to try to install these listed below anyway. The installer will let you know if the package is already present on your RPi. Text obtained from the image is then localized and converted to text format. The recognized text codes are output to users in speech.

WORKING

This system captures the document image placed in front of the camera which is connected to IoTRasberriy pi through USB .After selecting the process button the 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 TESSERACT library. Using Text-to-speech library the 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 USB.

OCR (Optical Character Recognition)

OCR is main element. It can convert the scanning image into editable text. OCR is implemented in this project to recognize characters which are then read out by the systems through a speaker.

TTS (Text to Speech)

It is method that scans and reads English alphabet and numbers. Webcam focused on text and it taking a picture. There are some delay is required. After the delay, it taking the picture and the processed by raspberry- pi

So that the steps for character recognition.

- Web camera capture the image then image can read.
- Pre-processing is done in 2nd step. Colour image is converted into gray scale and gray scale is converted into the binary image.
- Character is extracted then resizes the image.
- Load templates that can be matched.
- Remove the background
- Edge detection is done in last step of character recognition, in that open the text file and write the file.so that the output is stored in text format.

Above process is done in the image processing remaining process are done in voice processing. The image is converting into the text and text is converted into the speech.

HARDWARE IMPLEMENTATION

- Raspberry Pi 3
- Adaptor 5V/2A
- USB webcam
- VGA to HDMI cable
- 16GB microSD card
- Speakers these all hardware devices connected to Raspberry pi 3 model B via USB cables.

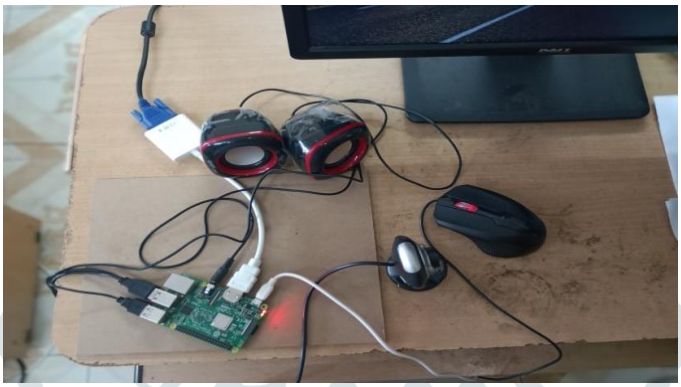

Fig.5. Hardware connection of Raspberry PI 3 Model B

SOFTWARE IMPLEMENTATION

Operating system: Raspbian (Debian)

Language: Python2.7

Platform: Tesseract, OpenCV (Linux-library)

Library: OCR engine, TTS engine

The operating system under which the proposed project is executed is Raspbian OS which is derived from the Debian operating system. The algorithms are written using the python language which is a script 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.

XI. RESULT

The connections are made and the setup is placed for our convenience. All the components like monitor, web camera, speakers, etc. are connected to the raspberry pi device. To run the code snippet, First, right click on the code and open with Python2 IDLE console. Then, run the code snippet. The above screen is opened with a welcome voice note, saying “please keeps the USB camera facing to image”.

Fig.4. Input Screen Setup

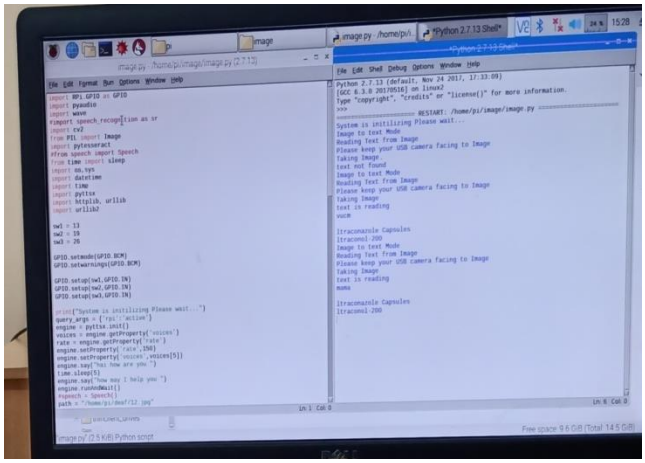

Fig.5. Running code snippet

Fig.6. Taking image from USB Camera

CONCLUSION AND FUTUREWORK

The voice assisted text reading system for visually impaired is discussed. The output is shown for the various input data set like only text inputs, text with images merged etc. Optical Character Recognition is used to predict the input text with pre-loaded database template. Both the characters are compared if it matches then using text to speech synthesizer, speech output is produced. Text to speech can convert the text on image into sound. For this conversion does not require internet connection. It is very easy to use, so the visually disabled person can independently use this device. the user feels easier to read the text in the form of speech using Optical Character Recognition and speech synthesizing. The proposed work is tested with different input sets in printed text format where noise parts are removed and text is extracted to predict the text aurally and is further enhanced by identifying the adjacent character recognition for effective reading to avoid the discontinuity. A output produced as audio output to read the corresponding input which helps the visually disabled people to read any printed text in vocal form. In the Future, some of the enhancements can be done so that it reads out the text printed from the shining surfaces like tablet foils where the light gets reflected and make it difficult to extract the text and avoid erroneous data output.

REFERENCES

[1]. Anusha Bhargava, Karthik V. Nath, PritishSachdeva and MonilSamel “Reading Assistant for the Visually Impaired” International Journal of Current Engineering and Technology, 2015.

[2]. Chaw Su Thu and TheingiZin Implementation of "TEXT TO SPEECH CONVERSION" international journal of engineering research and technology volume 3, Issue 3, March-2014.

[3] 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.

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

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

[6] 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 .

[7]. 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.

[8]Ainsworth, W., "A system for converting English text into speech," Audio and Electroacoustics, IEEE Transactions on , vol.21, no.3, pp. 288-290, Jun 1973.

[9] Fushikida, Katsunobu; Mitome, Yukio; Inoue, Yuji, "A Text to Speech Synthesizer for the Personal Computer," Consumer Electronics, IEEE Transactions on , vol.CE-28, no.3, pp.250-256, Aug. 1982.

[10] Hertz, S., "English text to speech conversion with delta," Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '86. , vol.11, no., pp.2427-2430, Apr 1986

[11] Lynch, M.R.; Rayner, P.J., "Optical character recognition using a new connectionist model," Image Processing and its Applications, 1989., Third International Conference on , vol., no., pp.63-67, 18- 20 Jul 1989

[12] S. Furui, "Speaker independent isolated word recognition using dynamic features of speech spectrum", IEEE Transactions on Acoustic, Speech, Signal Processing, Vol.34, issue 1, Feb 1986, pp. 52-59.

[13] Leija, L.; Santiago, S.; Alvarado, C., "A system of text reading and translation to voice for blind persons ," Engineering in Medicine and Biology Society, 1996. Bridging Disciplines for Biomedicine. Proceedings of the 18th Annual International Conference of the IEEE, vol.1, no., pp.405-406 vol.1, 31 Oct-3 Nov 1996.

[14] Tanprasert, C.; Koanantakool, T., "Thai OCR: a neural network application,"TENCON '96. Proceedings. 1996 IEEE TENCON. Digital Signal Processing Applications, vol.1, no., pp.90-95 vol.1,
26-29 Nov 1996

[19] Breen, A.P., "The future role of text to speech synthesis in automated services," Advances in Interactive Voice Technologies for Telecommunication Services (Digest No: 1997/147), IEE Colloquium on , vol., no., pp.6/1-6/5, 12 Jun 1997