

Survey paper: Text To Speech conversion for visually impaired peoples

Nisha P¹, Dr.J.Vijayakumar²

¹student of Mphil, Department of Electronics & instrumentation, Bharathiar University, Coimbatore.

²Associate Professor & Head, Department of Electronics & instrumentation, Bharathiar University, Coimbatore.

Abstract - There is an image everywhere around us and we see the image and read the text in our day-to-day life. Like bus names, bus numbers, hotel names, newspapers, etc. But the question is how Visually Impaired or blind people can recognize this text. Surely they need some assistance to read the text. In this research, the images are converted into text and the text is converted into audio output. It is mainly used for low visual persons or blind peoples to recognize the text. The field of research in Character recognition, Speech recognition and computer vision. In this research, as the recognition process is done using OCR, Raspberry Pi, MAT lab and openCV library. It recognizes characters using API, the e-Speak algorithm, PYTHON, and JAVA programming. This paper explains the purpose, implementation, and test results of the device. This project consists of capturing the image, text localization, text to audio conversion.

Key Words: Raspberry Pi, OCR, TTS, MAT lab, OpenCV

1.INTRODUCTION

There have been a number of advantages in the field to read the text help visually impaired people or blind person without much difficulty. Digital image processing is the method of computer algorithms to make picture processing on digital pictures to read the text. This paper says a study of effective prototype system to help the visually impaired people to become self-governing. At present there are abundant assistive technologies devices such as moving chair, vibrating watch and talker device are common example. Images with transcript act as central communication medium for conveying information to the blind people. Reading is one of the most common one today societies, but the low visually peoples are read the text is most difficult one in today. So we are developing the some computer vision system and audio output for low visual persons. The raspberry pi cameras are used to capture the digital images and OCR used which has function of MAT lab for converting image to text form. OCR makes it possible to apply techniques such as text to speech, text mining and machine translation. Then the TTS engines are used for text to speech with help of speaker. The output device is a headphone is connected to the speaker which can speak out the text aloud for visually impaired peoples. If you want better results you can use a high definition camera for it, otherwise, the output will be not completely accurate [1].

2. LITERATURE SURVEY

Table no -1

| Ref and Year | Technology used | Result |
|---------------------------------|---|--|
| Asha G et.al (2017) | Bilateral filter, canny edge detection, OCR, Festival TTS | The text region from the complex background and to give a high-quality input to the OCR. The text, which is the outcomes of the OCR is given to the TTS engine which provides the speech output. |
| Deepa V.Jose et.al(2014) | SQLite and Microsoft, TTS and STT | It includes the correctness of grammar and meaning with end results of achieving excellence in pronunciation. |
| Kiran Rakshana R et.al(2019) | OCR, TTS, Raspberry Pi, PYTHON, Application programming interface(API) | The OCR and TTS synthesizer were actualized to extricate the content data from images and convert it into the sound. |
| Vaibhav V et.al(2018) | ASCII, Histogram, Skewing, OCR, MAT lab | This approach can efficiently distinguish the object of attention from the background or other objects in the camera view.OCR is used to make word identification on the surrounded text fields and transform it into speech output for blind users. |
| L Latha et.al(2019) | OCR, Open source computer vision, PYTHON 3.6 | This system is useful for blind peoples to access information which is in the form of documents, texts or printed forms. |
| Shraddha Hingankar et.al(2020) | Thresholding, Segmentation, Tesseract OCR, Text to speech synthesizer | We will get represented an epitome system to scan the typing text and handheld objects for helping blind individuals. |
| Ram Nivas Duraisamy et.al(2018) | MAT lab, Raspberry Pi, Tesseract, TTS, Festival speech synthesis, pseudo code | The algorithm successfully processes the image and reads, the output come out well. This is an economical as well as efficient device for low vision people or blind peoples. |

| | | | | | |
|-------------------------------|--|---|----------------------------------|---|--|
| Aravind S et.al(2013) | MAT lab, Character recognition, Microsoft win 32 SAPI library | The method integrates both explained algorithms and novel methods for image processing text detection character identification and speech synthesis and good results have been reached. | Velmurugan et.al(2016) | Audio amplifier, LC filter, Raspbian | The simulation outcomes have been successfully validated and the device output has been tested using various samples. Our algorithm successfully processes the image and reads its outcome well. |
| Akhilesh Panchal et.al(2016) | SMF, Histogram, de-blurring, Blind de-convolution, Lab View .NET | Text verification is successfully performed using the patent matching technique. After successful verification, the text is converted into audio output. | K. Nirmala Kumari et.al(2016) | Image Text to Speech Conversion Using OCR Technique in Raspberry Pi[16] | Text is derived from the image and transformed into audio. It recognizes both capitals as well as small letters. It accepts numbers as well. |
| N. S. Lokhande et.al(2019) | Raspberry Pi, Audio amplifier, capacitor filter, OCR, Festival TTS. | The algorithms successfully process the image and its output is come out well. | D. Balaji et.al (2018) | MAT lab, TTS synthesizer, NLP, DSP, CART | In this work text into grammars and then that Grammar into speech is converted by MATLAB. The definition considers are, it does not read punctuation, Romans number. |
| R. Shantha Sleva Kumari et.al | TTS, MOS, Character To Tone, Word To Speech | The algorithm is checked for both men and women. The performance standard for the concatenative tongue is measured in terms of the Mean Opinion Score (MOS). The MOS value for the female is 4 and for the male is 3.5. The delay is decreased and mismatching of words also decreased. | Chaw Su Thu Thu et.al(2014) | Optical Character Recognition, Text To Speech, MAT lab | The mathematical numbers are including successfully transformed into text and then audio. Different type of font character is used and again it is transformed into text and then audio successfully. E-text into speech is also converted successfully. |
| A. Subbiah et.al(2016) | ADA Boost Algorithm, Raspberry Pi model B, | We should be done the Adaboost algorithm in the Raspberry PI Model for the transformation of Text to audio speech so that visually impaired peoples can easily recognize and read the data. | G.A.E. Sathish kumar et.al(2019) | Ultrasonic sensor, PIR sensor, GPS, Live stream OCR | The image capturing, transforming it into text, and object detection is analyzed many times to get accurate results. It should be giving 90-93% accurate output. |
| Itunuoluwa Isewon et.al(2014) | Text To Speech, JAVA | We have too developed a very simple and attractive GUI that enables the user to type in his/her text given in the text field in the application. Our system interfaces with a TTS engine developed for American English. | Pawan S Nadig et.al(2020) | MAT lab, segmentation, Speech API, OCR | Optical Character Recognition technique Kannada text images are completely converted into their corresponding speech. The resulting accuracy will be dependent on the training data. Also, use input language as English (India) to get a proper accent. |
| K. Lakshmi et.al(2016) | Kirsch detector, Optical character recognition, Google speech recognition Tool | The translation tools can convert the text to the coveted language and then Again by using the Google speech recognition tool can convert that changed text into voice | Ashwini V et.al(2016) | text localization, OCR, OTS | A novel text localization algorithm is proposed to localize text regions. Off-the-shelf OCR is used to perform text recognition on text localized regions and then recognized text codes are transformed to speech for a blind person. |

| | | |
|----------------------------------|---|---|
| Sagar G.k et.al(2017) | Image processing, Python | The outcomes of the simulation are tested successfully along many Samples that have been done to test the output The input image is processed efficiently by the algorithm used and cleared output is produced. |
| Muhammad Farid Zamir et.al(2020) | Tesseract OCR, TTS | The test results are evidence of this outcome that the recommended method gives excellent results in the processing of the image, extraction of text, and turn into speech. The analysis outcomes have shown that the extraction of text and its translation into audible speech is 100%. |
| Falgguni Patel et.al(2017) | Machine translation, Text-To-Speech, Text mining, OCR | It will depart the line from the image which is provided to the input and then every character will depart from the line. After that, every character will be recognized and then the text image will produce output in the form of speech. |
| Harshala V et.al(2020) | e-Speak algorithm, Raspberry Pi, OCR | Limited text is taken and transformed into speech. In our will extract more words as Possible. |

image processing and Optical character recognition [6][7]. We proposed, the MATLAB using for the image is transformed to text and then the text is converted to speech. We can translate text from a document and generate synthesized through a computers speaker [8]. This paper using image acquisition, image preprocessing, TTS conversion for automatic text detection and recognition based on image processing techniques. The preprocessing is used for noise removal. The approved median filter using for noise removal, the nonlinear filter used due to its excellent denoising power and computational performance [9]. The image of text print is taken from camera. The raspberry pi is connected to camera and converts into text. The preprocessing, segmentation are using for noise removal. The audio amplifier using for amplified the speech output [10]. TTS synthesis a simple character to voice translation. The alphabets (a-z, A-Z) and digits (0-9) are recorded in the order of wave data (.wav) for the database. Each character has a unique pronunciation. We play the wave data corresponding to each character read, in character to voice translation, we can also play the wave data for each word read. Once the text is read, for every word the corresponding wave file are concatenated and played. The MOS value for a female is 4 and a male is 3.5. the delay is reduced and mismatching of words also reduce [11]. The raspberry pi model for scanning the images from stereotypical forms such as bus numbers, store signs, product signs, and door numbers. The AdaBoost algorithm is used for processing the visual information converting into audio speech [12]. A TTS synthesizer system using some modules for conversion. NLP, DSP, text analysis, application of pronunciation rules, and prosody generation [17]. The JAVA programming language for using text to speech conversion in our projects. Our software is called text to speech robot [13]. Our aim of this project is text is converting to digital form. The kirsch detector is used to detect the eight-point area of every pixel in the image. Here OCR is printed text converted into machine-encoded text. The tesseract gives results with 100% accuracy [23]. Google text to speech API tool using for text into voice conversion with help of headset [14]. The web cameras are using for take an image. The images are processed by OCR software in Raspberry Pi. The TTS engine for use the text into speech and the output is given to audio amplifier. Amplifier connected to the speaker, if you want speaker can replace by the headphone [23]. The systems are simulated in MATLAB and also using by hardware. PYTHON are used in hardware to run the program [15]. TTS device consists of image processing module and voice processing module. The image processing module using for an image in text format and the voice processing module using for a text in a sound format. OCR automatically recognizes the character. The festival software is used to convert text to speech, the festival is an open-source TTS system [16]. MATLAB is using for converting the image to text and text to speech. E-text into speech also converted successfully. The preprocessing image is transformed into text by MATLAB. Microsoft Win 32 SAPI library has been applied to speech enable applications in this work. In this work only recognize the one character converted into text at once [18]. We

Nowadays the English language is most common in the reading system. The speech to text conversion using it makes easier and correct meaning. The SQLite and Microsoft speech recognition engine for using a speech-to-text conversion. The related study of the present TTS and STT algorithms are used and the quality of the outcome is improved [2]. In this paper, the ultrasonic sensor gives vibration sensing for visually impaired people to easily recognize the bus name and bus number in the bus stop with help of an audio processing algorithm [3]. The proposed project framework is the implementation of the image catching technique in MATLAB software. Most low visual people using Braille for reading a book and text is difficult to make and limited easily available. The OCR can be simulated in MATLAB for conversion. The object of attention from the background or other objects in the camera view is effectively distinguishing [4]. The openCV libraries are used for capturing the image of text and to perform the character recognition. The OCR technology is using for written text to machine encoded text. The openCV is very handle and convenient to use compared to the PC platform [5]. In this article, we do the Raspberry Pi 3 camera to capture the image, and tesseract OCR is the engine that extracts the recognized text. The good accuracy gives Tesseract and Festival with compared classical techniques of

proposed, the camera can capture the image or video according to the user command, PIR sensor detect the motion around the user, the ultrasonic sensor gives the distance between object and user. Users have headphones to intimate instructions. The Neo-6M GPS module is location updating system. User press the button the raspberry pi will update the location to the headphone, objects clearly up to 90%. PYTHON languages are used in this work [19]. In this work, OCR has converted the printed text into machine-encoded text. The captured images are using different preprocessing techniques like image acquisition, segmentation and template matching to remove the noise and suitable format. Speech API using for character to speech conversion [20] In this framework, the camera will capture the product object, ROI from the capture images are an extract from the motion-based method. The text recognition only performs the ROI. Optical character recognition is used to implement text recognition in the surrounded text regions. Recognized text is transfer into speech for blind users [21]. In this work, raspberry pi is using to read the alphabets and numbers, and the input is stored in .jpg format. The obtained image is converted to text format. The text file is processed by image processing techniques and the speaker gives an audio signal. TTS engine converts the text into an audio (.flac extension file) format [22]. The proposed system, OCR is converted the typewritten toward machine-encoded text. It's used for books convert into electronic files. The coding was done by MATLAB. TTS engine converts text into PCM digital audio [24]. This paper says, the camera is taken by text recognition and face recognition, identifies the text using OCR. E-speak algorithm using for conversion of recognizes text file to audio output. An analog signal produced by the e-speak is given to a headset to get the audio output signal [25].

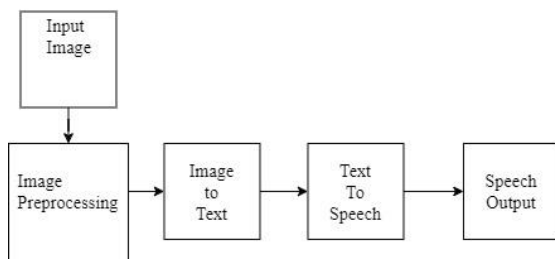


Fig -1: Basic block diagram of Text To Speech conversion

3. CONCLUSIONS

The TTS format is successfully converted by using MATLAB. The proposed algorithm is efficient system for conversion of text to speech format by using visually impaired peoples. The Raspberry Pi camera is captured the image for converting the text format. The OCR is convert image into text. TTS is using for TTS conversion. These algorithms are converting the text into audio format, and the output is clearly with set volume for low visual persons or blind peoples. The future work will be high accuracy of converted text into speech helping for blind peoples.

REFERENCES

- [1] Asha G. Hagargund, Sharsha Vanria Thota, Mitadru Bera, Eram Fatima Shaik(2017) "Image to Speech Conversion for Visually Impaired", International Journal of Latest Research in Engineering and Technology, ISSN: 2454-5031, Issue 06, Vol. 03, No. 0, pp. 09-15.
- [2] Deepa V. Jose, Alfateh Mustafa, Sharan R(2014) "A Novel Model for Speech To Text Conversion", International Refereed Journal of Engineering and Science, ISSN(online): 2319-183X,(print) 2319-1821, Issue 1, Vol. 3, No. 0, pp. 39-41.
- [3] Kiran Rakshana R, chitra C(2019) "A Smart Navguide System for visually Impaired", International Journal of Innovative Technology and Exploring Engineering, ISSN: 2278-3075, Issue 6S3, Vol. 8, No. 0, pp. 0.
- [4] Vaibhav V. Govekar, Meenakshi A(2018) "A Smart Reader for Blind People", International Journal of Science Technology & Engineering, ISSN: 2349-784X, Issue 1, Vol. 5, pp. 0.
- [5] L Latha, V Geethani, M Divyadharshini, P Thangam(2019) "A Smart Reader for Blind People", International Journal of Engineering and Advanced Technology, ISSN: 2249-8958, Issue 6S3, Vol. 8, No. 0, pp. 0.
- [6] Shraddha Hingankar, Prachi Tardekar, Santoshi Pote(2020) "A Smart Reader for Visually Impaired Individuals", International Research Journal of Engineering and Technology, P-ISSN: 2395-0072, E-ISSN: 2395-0056, Issue 07, Vol. 07, No. 0, pp. 0
- [7] Ram Nivas Duraisamy, Sathya Manoharan(2018) "A Smart Reader for Visually Impaired People", International Journal of Latest Engineering and Management Research, ISSN: 2455-4847, Issue 10, Vol. 03, No. 0. pp. 39-46.
- [8] Aravind S, Roshna E(2013) "A Text Reding System for the Visually Disabled", International Journal of Research in Computer Application & Management, ISSN: 2231-1009, Issue 12, Vol. 3, No. 0, pp. 0.
- [9] Akhilesh Panchal, Shrugal Varde, M.S Panse(2016) "Automatic Scene Text Detection and Recognition system for visually Impaired People", International Journal for Research in Emerging Science and Technology, E-ISSN: 2349-761, Issue 6, Vol. 3, No. 0, pp. 0.
- [10] N. S. Lokhande, P.B. Pawar, S.J. Shelke, A.R. Wagh(2019) "Book Reader for Visually Impaired Using Raspberry Pi", International Research Journal of Engineering and Technology, P-ISSN: 2395-0072, E-ISSN: 2395-0056, Issue 02, Vol. 06, No. 0, pp 0.

- [11] R. Shantha Sleva Kumari, R. Sangeetha "Conversion of English Text To Speech (TTS) Using Indian Speech Signal", *Mathematical and Computational Methods in Electrical Engineering*, Vol. 0, No. 0, pp. 0.
- [12] A. Subbiah, T. Arivukarasu, M.S. Saravanan, V. Balaji(2016) "Camera Based Label Reader for Blind People", *Sadguru Publications*, ISSN: 0972-768X, Vol. 0. No. 0, pp. 0.
- [13] Itunuoluwa Isewon, Jelili Oyelade, Olufunke Oladipupo(2014) "Design and Implementation of Text To Speech Conversion for Visually Impaired People", *International Journal of Applied Information System*, ISSN: 2249-0868, Vol. 7, No. 2, pp. 0.
- [14] K. Lakshmi, T. Chandra Sekhar Rao(2016) "Designs and Implementation of Text To Speech Conversion using Raspberry Pi", *International Journal of Innovative Technology and Research*, Issue 6, Vol. 4, No. 0, pp. 4564-4567.
- [15] Velmurugan. D, Srilakshmi, Umamaheshwari. S, Parthasarathy.S,Arun(2016) "Hardware Implementation of Smart Reader for Visually Impaired People using Raspberry Pi", *International Journal of Advanced Research in Electrical,Electronics and Instrumentation Engineering*, ISSN(print): 2320-3765, ISSN(online): 2278-8875, Issue 3, Vol. 5, No. 0, pp. 0.
- [16] K. Nirmala Kumari, Meghana Reddy J(2016) "Image Text To Speech Conversion using OCR Technique in Raspberry Pi", *International Journal of Advanced Research in Electrical,Electronics and Instrumentation Engineering*, ISSN(print): 2320-3765, ISSN(online): 2278-8875, Issue 5, Vol. 5, No. 0, pp. 0.
- [17] D. Balaji, M. Rajukumar, P. Sethu, V. Hariharan, S. Sekar(2018) "Implementation of Text To Speech with Natural Voices for Blind People ", *Paideuma*, ISSN: 0090-5674, Issue 2, Vol. 11, No. 0, pp. 49-53.
- [18] Chaw Su Thu Thu, Theingi Zin(2014) "Implementation of Text To Speech Conversion", *International Journal of Engineering Research & Technology*, ISSN: 2278-0181, Issue 3, Vol. 3, No. 0, pp. 0.
- [19] G.A.E. Sathish kumar, M. Munindhar(2019) "Intelligent Assistive System for Visually Disabled Persons", *International Journal of Recent Technology and Engineering*, ISSN: 2277-3878, Issue 4, Vol. 8, No. 0, pp. 0.
- [20] Pawan S Nadig, Pooja G, Kavya D, R Chaithra, Radhika A D(2020) "Kannada Text To Speech Conversion System", *International Journal of Innovative Technology and Exploring Engineering*, ISSN: 2278-3075, Issue 10, Vol. 9, No. 0, pp. 0.
- [21] Ashwini V, Mhaske, Mahesh S. Sadavarte, R Chaithra, Radhika A D(2016) "Portable Camera Based Assistive Text Reading from Hand Held Objects for Blind Person", *International Journal of Advanced Research in Computer and Communication Engineering*, ISSN(print): 2319-5940, ISSN(online): 2278-1021, Issue 6, Vol. 5, No. 0, pp. 0.
- [22] Sagar G.k, Shreekanth T(2017) "Real Time Implementation of Optical Character Recognition Based TTS System using Raspberry pi", *International Journal of Advanced Research in Computer Science and Software Engineering*, ISSN: 2277-128X, Issue 7, Vol. 7, No. 0, pp. 0.
- [23] Muhammad Farid Zamir, Khan Bahadar Khan, Shafquat Ahmmad Khan, Eid Rehman(2020) "Smart Reader for Visually Impaired People Based on Optical Character Recognition", *ResearchGate*, Vol. 0, No. 0, pp. 79-89.
- [24] Falgguni Patel, Dixita Patel, Talha Pathan, Anita Bindra(2017) "Smart Reading System", *International Journal of Scientific Research in Engineering*, Vol. 1(3), No. 0, pp. 33-37.
- [25] Harshala V. Gode, Prashant R. Indurkar(2020) "Text Reading for Visually Impaired Person using Raspberry Pi", *International Research Journal of Engineering and Technology*, P-ISSN:2395-0072, E-ISSN:2395-0056, Issue 02, Vol. 07, No. 0, pp. 0