Tamil Image Text to Speech Using Rasperry PI

V.Suresh Babu[1], D.Deviga[2], A.Gayathri[3], V.Kiruthika[4] and B.Gayathri[5]

[1]Associate Professor, [2,3,4,5]UG Scholar, Department of ECE, Hindusthan Institute of Technology, Coimbatore, Tamilnadu, India.

ABSTRACT

Abstract: India has a large blind population of 37million people were visually impaired across global among which 15million people are from India. Blind students can read only audio book, Braille or with help of personal assistant. By using this project, the visually impaired people to access various text resources and enhance their knowledge. In this device, it has an camera captures the text and it convert into a speech signal that read out through the earphones of the person and it will be in an wearable devices. Then the text can be separated from an captured image by using an OPENCV (Open Source Computer Vision) library and then audio control by raspberry PI model and the program is simulated using python software which gives an audio signal as output. That enables visually impaired people can read the text easily.

Keywords: Web camera, OPENCV library, Raspberry pi, TTS converter, SVM and Speaker.

1. INTRODUCTION

This project deals with a device that assists blind people with reading printed text in an real time. The camera module is used to capture the real time image of the product. It is given to the main module. The main module is Raspberry pi which is one of its own a mini computer, which processes the image captured by a camera. Then optical character recognition technique is used to process the image. Then that captured image will be separates the text from the image by using an OPENCV (Open Source Computer Vision) library. Then desired letters from the text can be identified by using a Tesseract OCR. Then it will be converted into voice using raspberry pi. Initially an image will be captured. Then second step is pre-processing step. By using a threshold operations color image will be converted into gray scale image and then converted into binary image. Then character image will be mapped to a higher level by using a special characters and patterns in that image. Then the extracted text will be gets from an SAPI (speech application programming interface). Then it compares the input string with SAPI. Then it extracts voice which is available in library. Then it chooses the pace of voice. Hence TTS converts the extracted text into speech. Finally it gives a speech from a given image.

2. LITERATURE REVIEW

[1] Hidden Markov Models had been described the western cursive handwriting by getting an speech recognition. It also recognized speech recognition for many languages such as Chinese, Japanese, etc. In this Hidden Markov Model is mainly recognized in isolated characters especially in an Indian scripts. The of range of an Indian scripts up to maximum 98% to minimum 92%.In that composition of having many lexicon sizes they are minimum 1k to maximum 20k. This Hidden Markov Model will be an initial character and speech recognition of many methods.

[2] Holistic Word Case recognition is using a multi-layer perceptron neural network. This paper is described how a standard multilayer perceptron neural network can used to correctly classify handwritten words according to whether they contain whole upper case or whole lower case characters. This without actually having to recognize any individual characters used an optimized (6-2-1) architecture MPL neural network trained with the conventional

back-propagation algorithm it shown that it is possible to successfully classify 84% of 1061 word data set. The data set being randomly selected from a 3183word data set obtained from 12 writers each submitting approximately 150 words of both cases.

[3] Automatic Recognition of Printed Oriya Script, this paper deals with an optical character recognition system for printed Oriya, a popular Indian script. 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.

[4] Western cursive words has been recognized by an HMM using modeling and recognition methods. In this method, the cursive words handwritten words is captured by using local feature of sequences of thin fixed-width vertical frame. The input image word of each vector in quantizing can be represented as Markov chain discrete symbols. In this method handwritten words can be mostly represented as optional ligatures and sequence of characters. This model can be also used in many recognized methods as reference model. By using an ligatures in this model the words will become having interconnection network of characters in construction of model words of infinite length.

3. EXISTING METHODOLOGY

In the previous models, user can identify conversion of text messages into speech. This project is mainly constructed for the visually impaired peoples and vocally impaired peoples. By using this method visually impaired people can more advantages in their education and also in enhancing their knowledge. In this method real time text localization and recognition method was used. Initially possibility of each External Reagent being a character was estimated using these features. Then that External Reagents with locally maximal possibility were selected for next stage. By using this classification accuracy was development. An efficient clustering algorithm was used. That group of External Reagents can be converted into text and OCR classifiers have trained on synthetic fonts is employed to label character recognition. After the character recognition it will goes to an voice analyzing technology and the information of input image will be displayed in LED display. After the filter process an audio output will be extracted through the speaker and converts it into audio format by reading information in image in a synthesized voice. This can be used in other languages books, documents and others and also away from home or office. But it has main disadvantage as time consumption for this process is very high, then accuracy for text recognition very less. Hence we are proposing a new methodology for overcoming these drawbacks.

4. PROPOSED METHODOLOGY

To overcome above difficulties in previous methodology, we propose a wearable device with camera which captures the text and the captured image is first detected and extracted using OPENCV library. Then the OCR

method is employed, which converts the images of typed or printed into a machine encoded text. After capturing image OCR will check out the error using a post-processing algorithm. Then captured will converted into speech signals then it will be read out through earphones or speaker. The software employed in this system is Python. The entire system is implemented using Raspberry pi3 model.
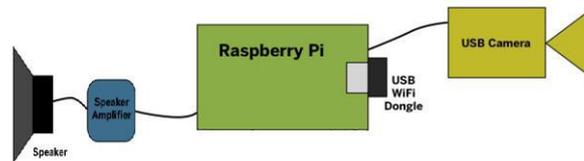

Fig 1. System Hardware Block Diagram

CAMERA: Camera is an digital device which can be used for image capturing, video recording, audio recording. Here we are using an web camera which will capture an image of input image and stores them. By using an diaphragm and adjusting the focal length in lens it will captures the image. And then transmitted to the raspberry pi for the next process.

RASPBERRY PI: The raspberry pi is a single computer board with credit card size, that can be used for many task that your computer. The main purpose of designing the raspberry pi board is, to encourage learning experiment and innovation for school level students.

Features of compared with raspberry pi 2, pi3 has:

 - A 1.2GHz 64bit quad-core ARMv8 CPU.
 - 802.11 IN Wireless LAN
 - Bluetooth 4.1
 - Bluetooth Low Energy(BLE)
 - 512 MB RAM
 - 2 USB Integrated USB Hub
 - 10/100 Wired Ethernet

USB PORT: USB Port is connectable port that can be mainly used for transmitting the audio and video data in one device to another device. In mainly High Definition Multimedia Interface can be also used for the transmitting the data in High Definition Multimedia Interface compliant sources. This can be also used in conversion of analog and digital standards. By using this HDMI-USB port we can control the operation of microcontroller in the system.

SOFTWARE TOOLS

 - Text extraction algorithm.
 - Text to speech synthesis.

> ➢ Integration layer with tesseract OCR.
> ➢ Python software.

What is python?

Python is an interpreter, object-oriented, high–level programming language with dynamic semantics. It is high-level built in data structures combined with dynamic typing and dynamic binding and make it very attractive for Rapid Application Development. It is very easy to learn and syntax emphasizes readability.

Features of python:

> ➢ Batteries Included.
> ➢ Everything is Object.
> ➢ Interactive Shell.
> ➢ Strong Introspection.
> ➢ Cross Platform.

Python Operators

> ➢ Python language supports following type of operators.
> ➢ Arithmetic Operators.
> ➢ Comparison Operators.
> ➢ Logical (or Relational) Operators.
> ➢ Assignment Operators.
> ➢ Conditional (or ternary) Operators.

Python Execution Model

During execution model source mode can be translated to byte code than it run the python code in virtual machine.

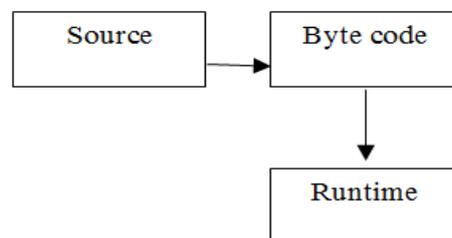If it is interpreted then code is automatically compiled.

Fig 2 Python Execution Models

DATABASE OPERATIONS

> ➢ Integration components.
> ➢ Games, images.
> ➢ XML, robot and more.

USERS OF PYTHON

- ➤ Google.
- ➤ PBS.
- ➤ NASA.
- ➤ Library of Congress.

ADVANTAGES OF PYTHON

- ➤ Easy to write.
- ➤ Easy to maintain.
- ➤ Open source
- ➤ Great standard libraries
- ➤ Powerful and fast.
- ➤ Dynamic and run immediately

DISADVANTAGES OF PYTHON

- ➤ Native loops are slow.
- ➤ Less control than C.
- ➤ Global interpreter lock.

5. RESULT

The result obtained from the procedure described above are indicated in the figures

Fig 3: Hardware Assembly

Fig 4: Output Image

6. CONCLUSION

This project provides a book reading concept for the blind utilizing a local-sequential scan. The system includes a text tracking algorithm that extracts words from a close-up camera view. Text to speech synthesis is a rapidly growing aspect of computer technology and is increasingly playing a more important role in the way we interact with the system and interfaces across a variety of platforms. The planned system gives a very simple method for text to speech conversion. Text inputs like the alphabets, sentences, words, and numbers are given to the system. Text to speech conversion is achieved and receives a better result which is audible and perfect. This system is very much used in the web applications, email, readings, mobile application and so on for making an intelligent speaking. To extract the text regions from a complex backgrounds book text localization algorithm is proposed. Using tesseract library the image will be converted into data and the data detected from the image will be shown on status bar. The obtained data will be pronounced through the ear phones or speaker.

REFERENCE

[1] Ajantha.V, Santhosh Baboo,"Embedded Optical Character Recognition On Tamil Text Image Using Raspberry PI, International journal of computer science trends and technology(IJCST)-Volume 2 Issue 4, Jul-Aug 2014.

[2] Sagar G.K, "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, vol.7, issue-7

[3] Velmurugan.V, "A Smart Reader for Visually Impaired People using Raspberry PI", ISSN 2321 3361, Vol.6.

[4] Nagaraja.L, "Vision Based Text Recognition using Raspberry PI", International Journal of Computer Applications (0975-8887).

[5] Tiji M Jose, "Recognition of Tamil Handwritten Character using Daubechies wavelet Transforms and Feed-Forward Back propagation Network" , International Journal of Computer Application, Vol.64, no.8.

[6] R.Jagadeesh Kannan, R.Prabhakaran, "A Comparative Study of Optical Character Recognition for Tamil Script" European Journal of Science Research, ISSN 1450-216X, VOL.3, NO4, (4009), PP.570-582.

[7] Aparna.K.G, A.G.Ramakrishnan, "A Complete Tamil Optical Recognition System".

[8] Naveen Sankaran, and C.V.Jawahar, "Recognition of Printed Devagiri Text Using BLSTM Neural Network", IEEE 2012.

[9] Nikil Mishra "Image Text to Speech Conversion Using Raspberry Pi & OCR Techniques", IJSRD-International Journals for Scientific Research and Development, Vol.5, Issue 08, 2017, ISSN (online): 2321-0613.

[10] Paneer Selvi, V.Suresh Babu "Optical Character Recognition For Malayalam Text Using Multi Class Support Vector Machine, International Journal of Application Research, ISSN 0973-4562, Volume 10, Number 16 (2015).