ISSN:2348-2079

Volume-7 Issue-1

International Journal of Intellectual Advancements and Research in Engineering Computations

Implementation of service robot using OCR technique in raspberry PI for visually impaired person

Prabhakaran R [1], C Thamilarasi[1], Jimin C Jiji[2]

[1]Assistant professor, Department of Electronics and Communication Engineering,
[2]UG students, Department of Electronics and Communication Engineering,
Shree Venkateshwara Hi-Tech Engineering College, Gobi, Tamilnadu, India.

ABSTRACT

Optical character recognition (OCR) is the identification of printed characters using photoelectric devices and computer software. It coverts images of typed, handwritten or printed text into machine encoded text from scanned document or from subtitle text superimposed on an image. In this project text images are converted into audio output. OCR is used in machine process such as cognitive computing, machine translation, text to speech, key data and text mining. It is mainly used in the field of research in Character recognition , Artificial intelligence and computer vision .In this project, as the recognition process is done using OCR the character code in text files are processed using Raspberry Pi device on which it recognizes character using tesseract algorithm and python programming and audio output is listened. To use OCR for pattern recognition to perform Document image analysis (DIA) we use information in grid format in virtual digital library's design and construction. This work mainly focuses on the OCR based automatic book reader for the visually impaired using Raspberry PI.

Keywords: Optical character recognition (OCR) Document image analysis (DIA),The technology of speech synthesis (TTS)

INTRODUCTION

The Blind peoples are unable to perform visual tasks. For instance, text reading requires the use of a braille reading system or a digital speech synthesizer (if the text is available in digital format). The majority of published printed works does not include braille or audio versions, and digital versions are still a minority. Thus, the development of a mobile application that can perform the image to speech conversion, whether its'a text written on a wall, a sheet of writing paper or in another support, has a great potential and utility. The technology of optical character recognition (OCR) enables the recognition of texts from image data. This technology has been widely used in scanned or photographed documents, converting them into electronic copies, which one can edit, search, play its content and easily carry. The technology of speech synthesis (TTS) enables a text in digital format to be synthesized into human voice and played through an audio system. The objective of the TTS is the automatic conversion of sentences, without restrictions, into spoken discourse in a natural language, resembling the spoken form of the same text, by a native speaker of the language. This technology has had significant progress over the last decade, with many systems being able to generate a synthetic speech very close to the natural voice. Research in the area of speech synthesis has grown as a result of its increasing importance in many new applications [1].

Author for correspondence:
Department of Electronics and Communication Engineering, Shree Venkateshwara Hi-Tech Engineering College, Gobi, Tamilnadu, India.

LITERATURE SURVEY

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. [9]

BRAILLE

Braille is writing and reading system used people who have visual impairment. Braille language is written on embossed paper. The braille characters are small rectangular blocks called cells that contain bumps called raised dots. The visually impaired person feels the arrangement of the raised dots which conveys the information. [5, 7] Braille literacy statistics of India: One out of every three blind people in the world is an Indian. It is estimated that nearly 15 million Indians are blind and out of that 2 million are children. Only 5% of the children receive education. Although braille readers, keyboards and monitors exist, they are not accessible to the rural communities and braille material is not easily and abundantly available [6].

RASPBERRY PI

The raspberry Pi is a small, low cost CPU which can be used with a monitor, keyboard and mouse to become an efficient, full-fledged computer [8]. The reason we chose Raspberry Pi micro-computer for our project is that, firstly, it is an easily available, low-cost device. RPi uses software which are either free or open source, which also makes it cost-effective. The Raspberry Pi uses an SD card for storage and its small size also gives us the advantages of portability. [3] As a part of the software development, the Open CV (Open source Computer Vision) libraries are utilized for image processing. Each function and data structure was designed with the Image Processing coder in mind [10].

EXISTING SYSTEMS AND LIMITATIONS

One of the biggest advantages of barcode readers is portability. Hence, they can be used by the visually impaired in identifying different products. An extensive database is created which contains all the information about the product. The user simply scans the bar code and the product details are listed through e-braille readers.

The disadvantage with this product is that the user might not be able to point the bar code reader in the correct direction [2]. Another approach is optical enhancement solutions such as an optical zooming device that expands the braille character. However, not all visually impaired people need to know braille language. Some methods aim at converting text to speech. This is accomplished using a scanner, speakers and a computer. This method is efficient only with simple scanned documents. It cannot extract text from an image with a complex background. [4]

PROJECT OVERVIEW

Types of Graphics

Optical Character Recognition is the mechanical or electronic conversion of images of typed, handwritten or printed test into machine coded text, whether from a scanned document, a photo of a document, of from subtitle text superimposed on an image. The basic junction of OCR Technology is to automatically capture printed text present scanned images and convert it into a text searchable document. A text to speech (TTS) synthesizer is a computer based system that read text aloud automatically, regardless of whether the test is introduced the system that can read text aloud automatically, regardless of whether the is introduced by a computer input stream or a scanned input submitted to implemented by both hardware and software. Speech is often based concatenation of natural speech ie. Units that are taken from natural speech altogether to form a word or sentence.

Software Specifications

- The operating system is raspbian.
- Language: Python.

- Platform: OpenCV
- Library: OCR engine, TTS engine.

This paper consists of a prototype system of assistive text reading. The concept of proposed system is the idea of developing specs reader based text reading system for visually impaired persons. There are three different modules in this system: Camera module, Optical Character Recognition

Module and Text-To-Speech Module This explains the text reading system for visually impaired users for their self-independent.

HARDWARE SPECIFICATIONS.

- Raspberry Pi 3(Model B)
- Broadcom BCM2837 processor
- LCD Display
- Capacitors.
- Resistors.
- Battery.
- Speaker.

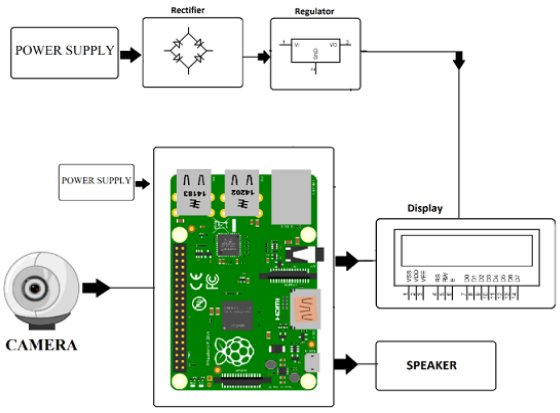
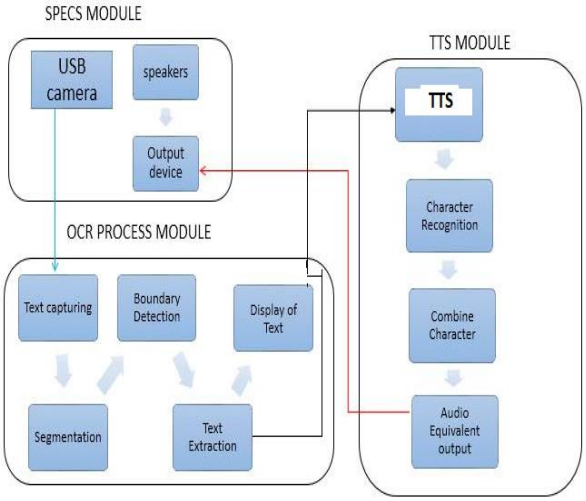


Fig. 1.Hardware features of the proposed system.

PROPOSED SYSTEM

When capture button is clicked this system captures the product image placed in front of the web camera which is connected to Raspberry pi through USB .After selecting the process button the captured label 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 Flite library the data will be converted to audio. Camera acts as main vision in detecting the label image of the product or board then image is processed internally and separates label from image by using open CV library and finally identifies the product and identified product name is pronounced through voice. Now it identifies received label image is converted to text by using tesseract library. Once the identified label name is converted to text and converted text is displayed on display unit connected to controller. Now converted text should be converted to voice to hear label name as voice through ear phones connected to audio jack port using flite library.

In this research, the camera acts as input for the paper. As the Raspberry Pi board is powered the camera starts streaming. The streaming data will be displayed on the screen using GUI application. When the object for label reading is placed in front of the camera then the capture button is clicked to provide image to the board. Using Tesseract library the image will be converted into data and the data detected from the image will be shown on the status bar. The obtained data will be pronounced through the ear phones using Flite library.



RASPBERRY PI MODULE WITH OCR & TTS

MODULES DESCRIPTION

Image processing: In the first step the device is moved over the printed page and the inbuilt camera captures the images of the text. The quality of the image captured will be so high so as to have fast and clear recognition due to the high resolution camera. Letters will be extracted and converted into digital form.

Preprcoessing: It consists of three steps: Skew Correction, Linearization and Noise Removal. The captured image is checked for skewing. There are possibilities of image getting skewed with either left or right orientation. Here the image is first brightened and binarized. The function for skew detection checks for an angle of orientation and if detected then a simple image rotation is carried out till the lines match with the horizontal axis, which produced a skew corrected image. The noise introduced during capturing or due to poor quality of the page has to be cleared for further processing.

Segmentation: This operation seeks to decompose an image of sequence of characters into sub image of individual symbol (characters). The binarized image is checked for interline spaces. If inter line spaces are detected then the image is segmented into sets of paragraphs across the interline gap. The lines in the paragraphs are scanned for horizontal space intersection with respect to the background. Histogram of the image isused to detect the width of the horizontal lines. Then the lines are scanned vertically for vertical space intersection. Here histograms are used to detect the width of the words. Then the words are decomposed into character width computation.

Feature Extraction: In this stage we gather the essential features of the image called feature maps. One such method is to detect the edge in the image, as they will contain the required text.

Tesseract: This software is used to convert the image file to text file by extracting the texts from the image and storing it in the file with.txt extension. Google Text to Speech API: This is used to convert the output content into voice which can be heard on the speakers.

FUTURE SCOPE

The future work will focus on optimize the algorithms to improve the computing efficiency, design user interface to study with blind and visually impaired subjects. Furthermore, we should test more types of objects to make our algorithm widely used. It is also essential to make more experiments to further approve our method.

CONCLUSION

We plan to implement text to speech conversion technique using raspberry pi. The simulation results have been designed and the hardware output has to be test different samples. Our algorithm will successfully process the image

and reads it out clearly. This is an economical as well as efficient device for the visually impaired people. We are going to apply our algorithm on many images and will find whether it successfully complete the conversion. The device is compact and helpful to the society.

Acknowledgment

We are expressing our thanks to all Faculty members and Skilled Assistants of Electronics and Communication Engineering department and my Friends who helped me in every possible way. Last but not least I thank my Parents for their moral support.

REFERENCES

[1]. AlexandreTrilla and FrancescAlías, "Sentence-Based Sentiment Analysis for Expressive Text-to-Speech", IEEE Transactions on Audio, Speech, and Language Processing, 21(2), 2013, 223-233.

[2]. Alías F. Sevillano X. Socoró J. C Gonzalvo X. (2008),„Towards high-quality next-generation text-to-speech synthesi", IEEE Trans. Audio, Speech, Language Process,16(7), 2008, 1340-1354.

[3]. Balakrishnan G. Sainarayanan G. Nagarajan R. andYaacob S,Wearable real-time stereo vision for thevisually impaired", 14(2), 2007, 6–14.

[4]. Chucai Yi. YingLiTian.AriesArditi.,Portable Camera-based Assistive Text and Product Label Reading from Hand-held Objects for Blind Persons", IEEE/ASME Transactions on Mechatronics, 3(2), 2014, 9-10.

[5]. Deepa Jose V. and Sharan R„A Novel Model for Speech to Text Conversion", International Refereed Journal of Engineering and Science (IRJES), 3(1), 2014, 39-41.

[6]. Goldreich D. and Kanics I. M. „Tactile Acuity is Enhanced in Blindness", International Journal of Research And Science, 23(8), 2003, 3439–3445.

[7]. Shaalini, R., et al. "Human motion detection and tracking for real-time security system." International Journal of Advanced Research in Computer Science and Software Engineering, 3(12), 2013.

[8]. Roopa, V. G. "Real Time Visual Tracking of the People using Video Camera with Reduced Time Complexity." Unique Journal of Engineering and Advanced Sciences 1(2), 2014.

[9]. Joao Guerreiro and Daniel Gonçalves „Text-to-Speech: Evaluating the Perception of Concurrent Speech by Blind People", International journal of computer technology, 6(8), 2014, 1-8.

[10]. J. Liang D. and Doermann H,Camera-based analysis of text and documents: a survey," International Journal on Document Analysis and Recognition, 7(6), 2005, 83-200.