Review of Text to Voice Conversion for Visually Impaired Person by using Camera.

[1]Mr. Sumit Chafale., [2]Ms. Priyanka Dighore [3] Ms.Dipika Panditpawar., [4]Mr. Khushal Bhagawatkar,[5]Mr. Pravin Donade.
[1]Asst. Professor, [2]Student, [3]Student [4]Student [5]Student
[1,2,3,4,5] Department of Electronics &Telecommunication Engineering
[1,2,3,4,5] Suryodaya college of Engineering & Technology, Nagpur, India

Abstract: We propose a text to voice conversion for visually impaired person by using camera. Helps the visually impaired person to read the text label and product packing from handheld object in their daily lives. The system framework consist of three components: First, image capture component capture the image with help of camera. Then that image is send to the Data Processing. Second, Data Processing component is used for to filter the text from surrounding and will be recognized by OCR. Third, Text to Speech module is used to get audio output and we also employ the Microsoft's speech synthesizer for text to speech conversion. Since the project should be portable to assist the visually impaired persons, the entire application is based on MATLAB.

Index Terms: **Camera, Optical character recognition (OCR), Stroke with filtering, Binarization Hand- held objects, MATLAB.**

I. INTRODUCTION

The reading is very important in daily lives. Text is present all over in real world in the form of printing reports, product packages, medicines, bank statements, etc. In India there are 21.9 million disabled people are present. Out of which 48.5% people are visually impaired. One of the major problems they face is that understanding text information which is not in Braille language. Now a day's there is a need of basic education like to read text label from hand-held objects for children and adults. There is a one solution, it is given by teacher but it requires more effort, and takes more time also different techniques are necessary

These paper present a system of text to voce conversion for visually impaired person by using camera. We can capture the image of report documents with help of image capturing device such as digital camera, and mobile phones. Here we are using Text to speech module. This module is basics of phonics of English. Recognized texts are passed serially to Text to Speech module with the help of Serial to USB convertor. In Text to speech module text information is converted into speech or audio and we also employ the Microsoft's speech synthesizer for text to speech conversion. Input to speech synthesizer is text and it produces output as an audio stream. Visually impaired person can hear the the audio output with the help of earphone or speaker.

By using MATLAB we can acquire data from hardware device using serial port as sound card and major data directly into MATLAB for visualization and analysis. Graphical User Interface (GUI) is main part of the software application interaction between human and computer.

II. LITERATURE SURVEY

The state of lacking the visual perception due to physiological or neurological factors is called blindness. The lack of integration in the growth of the optic nerve or visual center of the eye represents partial blindness. The total blindness is the full absence of the visual light perception. In this proposed work, a virtual eye which is simple, cheap, friendly user is designed and implemented, in which the mobility of both blind and visually impaired people is improved in specific areas. This proposed work is a wearable one and it consists of a head hat, mini hand stick and foot shoes are used, by which the blind person can navigate alone safely and easily and avoid any obstacles that may be encountered, whether fixed or mobile, prevent any possible accident. The ultrasonic sensor is the main component which is used for the blind people to scan predetermined area by emitting-reflecting waves. The input to Arduino microcontroller is the Signals received from the barrier objects. The commands which are issued are carried out by the microcontroller and the status of the given appliance is communicated or it is derived to the earphone using raspberry pi speech synthesizer.

The proposed system is fast, cheap and it is easy to use, it gives an innovative and affordable solution to the visually impaired and blind people in third world countries. To help the blind persons to read text labels and product packaging from hand-held objects in their daily lives a camera-based assistive text reading framework is proposed and developed. An effective and efficient motion-based method is proposed first in which Region of Interest (ROI) is defined, a novel text localization algorithm is proposed by learning gradient features of stroke orientations and an Ada boost model is used for the distributions of edge pixels. Text characters in the localized text regions are then binarized and it is recognized by off-the-shelf optical character recognition (OCR) software. The output to the blind user in speech is the recognized text codes. The braille system of reading is the method which is used by the majority of the blind people for reading books and documents, but this method found to be difficult to make and less readily available. This difficulty has to be minimized by developing a device that could bring relief to the agonizing tasks that the

visually impaired has to go through. Due to the 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. To build texture representations, deep convolutional neural networks are used for recent approaches. Nevertheless, the representation of texture are unclear and it is invariances to categorical variations. The recent CNN-based texture descriptors for recognizing is the work of systematic evaluation and attempts to understand the nature of invariances captured by these representations. A general purpose texture descriptor which is excellent is bilinear CNN model, which is proposed recently and to other CNN-based descriptors, it is favorably comparable on various texture and scene recognition benchmarks. This model acquires translationally invariant and without requiring spatial jittering of data it obtains better accuracy on the image Net dataset of data compared to corresponding models trained with spatial jittering. The technique which we proposed is based on recent work in which the pre-images are visualized and by providing a means to understand categorical properties which are captured by these representations. Finally, we show preliminary results on how a unified parametric model of texture analysis and synthesis can be used for attribute-based image manipulation.

III.PROPOSED SYSTEM AND METHODOLOGY

Figure.1 gives the block diagram of the steps to be followed for image processing, text reading and to give audio output.

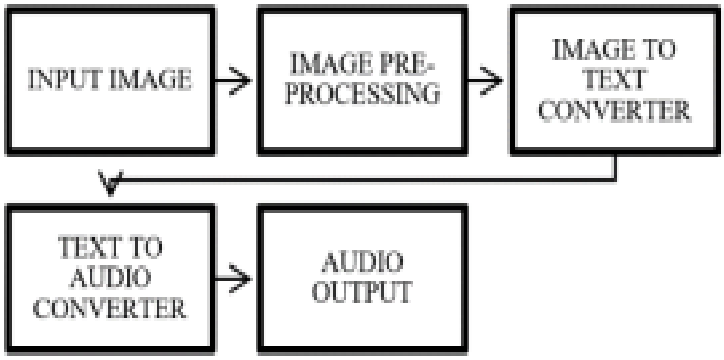

Figure.1 Block diagram of project using OCR

Image capture: This step consists of color to gray scale conversion, edge detection, noise removal, warping and cropping and thresholding. The image is converted to gray scale as many Open CV functions require the input parameter as a gray scale image. Noise removal is done using bilateral filter. Canny edge detection is performed on the gray scale image for better detection of the contours. The warping and cropping of the image are performed according to the contours. This enables us to detect and extract only that region which contains text and removes the unwanted background. In the end, Thresholding is done so that the image looks like a scanned document. This is done to allow the OCR to efficiently convert the image to text.

Data Processing: The above diagram (fig.1) shows the flow of Text-To-Speech. The first block is the image pre-processing modules and the OCR. It converts the pre-processed image, which is in .png form, to a .txt file. We are using the Tesseract OCR.

Audio Output: It is the artificial production of human speech. A computer system used for this purpose is called a speech synthesizer, and can be implemented in software or hardware. A text-to- speech (TTS) system converts normal language text into speech.

IV. APPLICATION

a. For visually impaired people in public places like mall, bank statement, hospital, hotel menus, any shops etc.
b. For normal and illiterate people.
c. It is use for foreign language. Such type of technology will help people travelling abroad for understanding sign, destination of buses and trains, names of streets etc.

V. FUTURE SCOPE

We have proposed a system to read printed text on various handheld object, natural seen images, book covers, hotel menus and so on.Our future work will extend our project to handle non-horizontal text string. It can effectively extract the text regions from images with complex background. In future we will enhance this work with higher accuracy and large number of input samples.
In future we can improve this project for handicap people also by using speech to text module for writing purpose

REFERENCES

[1] National Census of India, [Online]. Available:http://censusindia.gov.in/Census_And_ You /disabled_population.aspx

[2] Scan Talker, Bar code scanning application to help Blind Identify over one million products(2006).[Online].Available:http://www.freedomscientific.com/fs_news/PressRoom/en/ 2006/ScanTalker2-Announcement_3–30–2006.asp

[3] K Reader Mobile User Guide, knfb Reading Technology Inc. (2008). [Online]. Available: http://www.knfbReading.com

[4] International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 17 (2018) pp. 12995-12999 @Research India Publications.

[5] International Research Journal of Engineering and Technology (IRJET) Volume:04 issue:03|Mar-2017.

[6] International Journal of Advance Research in Computer and Communication Engineering Vol. 5, Issues6, June 2016.

[7] International Journal of Innovative Research in Science, Engineering and Technology. Vol.7, Issue12, December 2018.