IOSR Journal of Engineering (IOSR JEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 46-51

Voice Assisted Text Reading System For Visually Impaired Persons Using TTS Method

A.Sankaran[1], R.Barath Priyadarshan[2], A.Mohamed Thaaric[2], S.Nataraj
Assistant Professor[1], UG Student[2]
Computer Science and Engineering
Manakula Vinayagar Institute of Technology, Pondicherry

Abstract: *According to the World Health Organization, out of 7.4 billion population around 285 million people are estimated to be visually impaired worldwide. It is observed that they are still finding it difficult to roll their day today life and it is important to take necessary measure with the emerging technologies to help them to live the current world irrespective of their impairments. In the motive of supporting them We have proposed a smart spec for the blind persons which can perform text detection thereby produce a voice output. This can help the visually impaired persons to read any printed text in vocal form. A specs inbuilt camera is used to capture the text image from the printed text and the captured image is analyzed using Tesseract-Optical Character recognition (OCR). The detected text is then converted into speech using a compact open source software speech synthesizer, eSpeak. Finally, the synthesized speech is produced by the headphone by TTS(Text-to-Speech) method. In this project Android is the main target for the implementation, as it provides an interface between camera and image processing results, while also performing functions to manipulate peripheral units.*

Keywords: *OCR, TTS, Audio module, Android*

I. Introduction

The number of visually impaired persons is increasing thanks to uncontrolled polygenic disorder, age connected causes, eye diseases, traffic accidents, and alternative causes. Cataract is leading explanation for cavity and impairment. Mobile applications that offer the support to sand-blind person became an important device in sand-blind person's life. Recent advances in mobile technology, camera, pc vision and camera based mostly} application create it potential to support sand-blind persons by developing camera based application that

Prevalence of blindness (per thousand)
As per the estimate of 2016

Types of blindness	Urban	Rural	Total
Total blindness	4.43	5.99	5.40
Economic blindness	11.14	15.44	13.83
One eye blindness	8.23	7.00	7.46

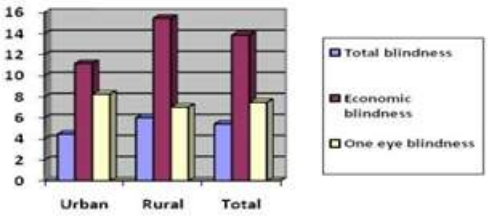

Fig 1. Prevalence Of Blindness

mix pc vision with alternative existing technology like optical character recognition (OCR) system. With the fast development of camera-based applications on good phones and handy devices, understanding the photographs taken by these devices has gained growing attention from the pc vision community in recent years which can be useful for these people. the most focus of our analysis is that the sand-blind person will get info concerning written text, text boards, scene text, hoardings, and directions on traffic sign boards in audio kind. With now of read, the system style for a camera primarily based reading system that extract text from matter board and establish the text characters from the captured image and at last, matter info are going to be regenerate into speech.

We propose a camera-based helpful text reading framework to assist blind persons scan text labels and merchandise packaging from hand-held objects in their daily lives. To isolate the object from littered

backgrounds or different close objects in the camera read, we have a tendency to 1st propose Associate in Nursing economical and effective motion based technique to outline a locality of interest (ROI) within the video by asking the user to shake the item. This technique extracts moving object region by a mixture-of-Gaussians-based background subtraction technique. within the extracted ROI, text localization and recognition square measure conducted to amass text data. To mechanically localize the text regions from the item ROI, we have a tendency to propose a completely unique text localization formula by learning gradient options of stroke orientations Associate in Nursing distributions of edge pixels in an Ad boost model. Text characters within the localized text regions square measure then binarized and recognized by ready-to-wear optical character recognition package. The recognized text codes square measure output to blind users in speech.

II. Related Work

Yi, C., Tian, Y., & Arditi, A. ***et, al. [1]*** Now a day's written text is one among the outstanding communication medium to induce product names, sign boards message, eating house menu etc. so as to induce this data blind folks want some help. In recent years there's an enormous improvement in image process and embedded systems. during this project we have a tendency to planned associate economical system for blind folks to browse written text by combining image process and embedded systems. Camera acts as main vision in police investigation the label image of the merchandise or board then image is processed internally and separates label from image by victimization Matlab and at last identifies the merchandise name and by victimization serial communication transfer this data to ARM seven microcontroller. Microcontroller processes received data and pronounces product name or message by victimization voice module Gregorian calendar month 9600 and speaker. To isolate the article from complicated backgrounds, we have a tendency to 1st propose an efficient motion-based methodology to outline an area of interest (ROI) within the image. within the extracted ROI, text localization and recognition square measure conducted to accumulate text data. associate experimental result shows that our system achieves the state-of-the-arts and Self-Dependency for disabled persons raised.

R. U. Shekokar. *et, al. [2]* Today there square measure numerous camera based mostly technologies out there for blind or visually impaired peoples that enhance freelance living and foster economic and social self-reliance. numerous devices developed to create the lifetime of blind folks straightforward. variety of transportable reading assistants systems like barcode scanners, product symbol, indoor and outside object detection, road navigator are designed specifically for the visually impaired peoples. For planning these systems numerous algorithms were used. This paper provides survey on all algorithms and technologies employed in planning numerous systems for visually impaired peoples.

Ramesh Babu.Y. *et, al. [3]* A camera based mostly helpful text scanning framework to assist blind persons read text labels and products packaging from hand-held object in their daily resides is projected. To isolate the thing from littered backgrounds or different surroundings objects within the camera read, we have a tendency to propose associate degree economical and effective motion based mostly technique to outline a vicinity of interest (ROI) within the video by asking the user to shake the thing. within the extracted ROI, text localization and recognition square measure conducted to accumulate text data. To mechanically localize the text regions from the thing ROI, we have a tendency to propose a completely unique text localization formula by learning gradient options of stroke orientations associate degreed distributions of edge pixels in an Ad boost model. Text characters within the localized text regions square measure then binarized and recognized by off-the-peg optical character recognition computer code. The recognized text codes square measure output to blind users in speech.

Ch.Sridevi Reddy. ***et, al. [4]*** This paper advise a video camera-based helpful text browsing framework to help blind persons read text labels and merchandise packaging from hand-held objects at intervals their lives. To isolate the item from littered backgrounds or the other encompassing objects within the camera read, this 1st propose an {efficient good} and efficient motion based mostly methodology to outline a region of nice interest (Region of Interest) at intervals the video by asking the patron to shake the item. This method extracts moving object region with a mixture-of-Gaussians-based background subtraction methodology. at intervals the removed region of interest, text localization and recognition square measure allotted to urge text info. To mechanically localize the written language regions within the object region of interest, this advise a completely unique text localization formula by learning gradient choices that go with stroke orientations And distributions of edge pixels at intervals an Ad boost model. Text figures at intervals the localized text regions are known by video camera connected to computer in conjunction with application code. The recognized text codes square measure output to blind customers in speech. The proof-of-concept paradigm is additionally evaluated on the dataset collected victimization 10 blind persons to gauge the effectiveness from the system's hardware. Experimental results show this algorithmic program achieves the condition from the humanities.

III. PROPOSED WORK

We use the android studio to propose our work by Android Studio is an IntelliJ IDEA based IDE and declared by Google as the official IDE for Android application development. Android Studio is the official IDE for Android development.

The construct of projected system is that the plan of developing automaton app reader based mostly text reading system for visually impaired persons. This illustrates the text reading system for visually impaired users for his or her self-independent. the matter stresses the high importance of visually impaired system is that self -dependency of visually impaired users. This extends the work towards the event of simple assembling info, self-dependent. to realize the required result, framework combines a collection of various modules, like App reading device, TTS module and optical character recognition module. automaton is AN open supply and Linux-based software system targeted for mobile devices like smart-phones and pill computers. Applications square measure usually developed in Java artificial language exploitation the automaton computer code development kit (SDK). If used properly, the SDK, in conjunction with Eclipse (the formally supported IDE) and JDK (Java Development Kit) is capable to deliver trendy computer code for automaton devices.

SOFTWARE USED

Front End :ANDROID XML, JAVA
Back End : SQlite
Operating System : Windows 7/8/10
IDE : Eclipse, Android Studio

A Sequence diagram shows the participants in an interaction and the sequence of message among them. A sequence diagram shows the interaction of a system with its actors to perform allot part of a use case. Sequence diagram shows the interaction between the user and cloud for registering, uploading the images, converting image to text and text to audio. The order during which these interactions present itself. we are able to additionally use the terms event diagrams or event eventualities to talk over with a sequence diagram. Sequence diagrams describe however and in what order the objects during a system operate. These diagrams area unit wide utilized by businessmen and software system developers to document and perceive necessities for new and existing systems.

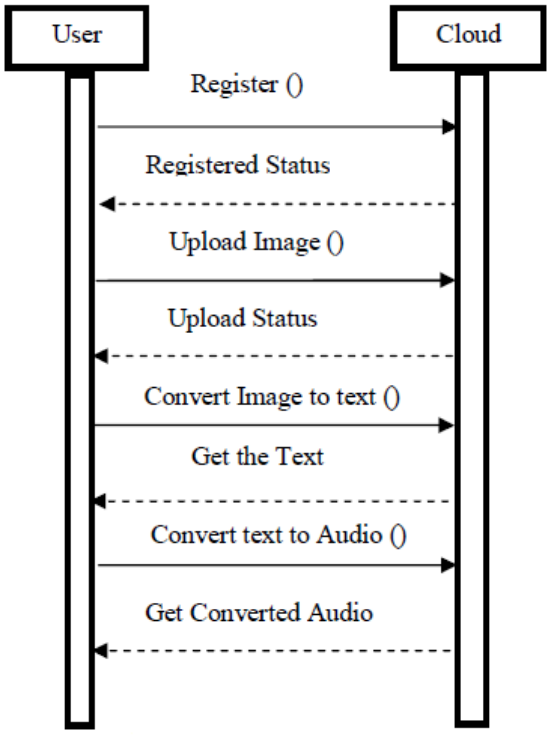

Fig 2. Sequence Diagram

OCR:

OCR is optical character recognition module is that the mechanical or electronic conversion of pictures of written, handwritten or printed text into machine-encoded text. it's a standard technique of digitizing written text so it is utilized in machine method like text-to-speech. OCR is optical character recognition module is that the mechanical or electronic conversion of pictures of written, handwritten or printed text into machine-encoded text. The input is given as text, employing a finger device mounted camera that captures text and sends the input text to the OCR method wherever the extraction of text to speech is been done. From the captured input text is metameric as word by word detection thereby to scan it as separate word. Boundary detection is finished by detecting words that square measure work within the boundary, if not it eliminates the text that is unfit to scan. The method of text extraction is dole out by matching with templates one by one and so forming a full word. The mentioned line or a word can be scan from the captured input text with a appropriate secret writing. once matching with the templates and displays it as a text and reads it aurally.

IMAGE CAPTURE MODULE:

The image capture module will catch the image captured by the Camera attached to the goggles. This will be easy for the visually impaired person to capture the image as the camera will be situated on the goggles. The image captured will be converted into grayscale and binarization

IMAGE CORRECTION MODULE:

This module will correct the image by reducing the noises by mean of filtering algorithm like median filter so that the text will be effectively recognized

TEXT EXTRACTION MODULE:

This module will recognize and extract the text. This will be achieved using OCR-Optical Character Recognition - is the Mechanical or electronic conversion of images of typed, handwritten or printed text into machine-encoded text. We will be using here MODI algorithm of OCR.

TTS:

TTS Module A text-to-speech (TTS) system converts normal text into speech other systems render symbolic linguistic representations like phonetic transcription into speech. A text-to-speech system is used to read each word as the user`s finger passes over it, and distinctive audio and/or haptic cues can be used to signal other events, such as end of line, start of line etc. It is composed of two parts: a front-end and a back-end. The front-end has two major tasks. First, it converts raw text containing symbols like numbers and abbreviation into the equivalent of written-out words. This process is often called text normalization, pre-processing, or tokenization the front-end then assigns phonetic transcriptions to each word, and divides and marks the text into prosodic units, like phrases, clauses and sentences. The process of assigning phonetic transcription to words is called text-to phoneme or grapheme-to-phoneme conversion.

AUDIO MODULE:

This module will get the extracted text as an input and it will read out the text using Text to Speech available in the mobile. If no text is recognized, then default audio output will be given. This will be done using SAPI libraries.

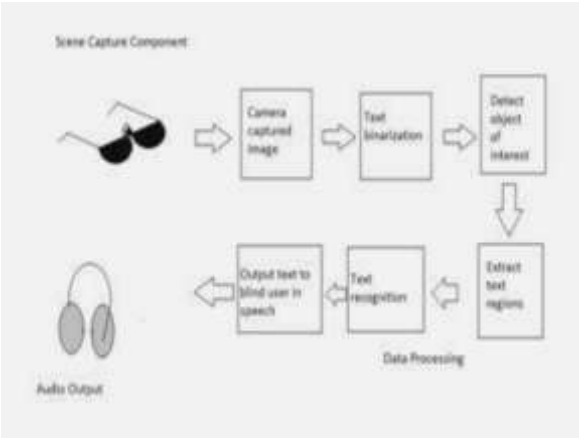

Fig 3 Architecture

System architecture shows the interaction between the user and application when it converts captured image to speech**.** It captures the image by user specs and convert it using text binarization algorithm. Then detects the object in the respective image and extract the text region in it. After recognize the text it transfer the text into the speech output as by audio module.

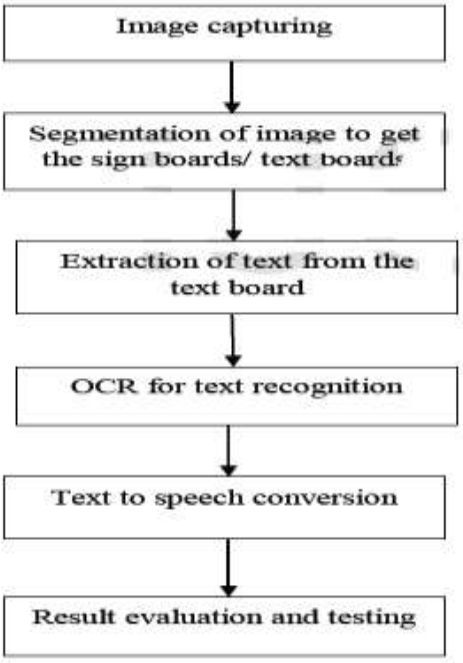

Fig 4.Flow Diagram

IV. Experimental Results

The texts within the kind of (jpeg, png, jpg, bmp etc) ar thought-about for the analysis. The image that is captured from the USB camera is splited in following conditions as delineated below for sleuthing corresponding text and matching it with templates prescribed within the below conditions.

- Text within the kind of Black and White
- Text within the kind of colored
- Text with image
- Text and image incorporate
- Text with totally different Font designs

Fig 5. TTS result

Fig 6.Bar Code Result

V. Conclusion

A text detection and recognition with speech output system was with success incontestable on mechanical man platform. this technique is extremely handy and helpful for the purblind persons. Compared with a computer platform, the mobile platform is transportable and a lot of convenient to use. this technique are going to be useful for purblind persons to access info in written kind and within the encompassing. it's helpful to know the written communication messages, warnings, and traffic direction in voice kind by changing it from Text to voice. it's found that this technique is capable of changing the sign boards and alternative text into speech.

VI. Future Enhancement

We have enforced text to speech conversion technique victimization by android. The results are with success verified and therefore the output has been tested victimization totally different samples. Our formula with success processes the image and reads it out clearly. this is often a cost-effective further as economical device for the visually impaired individuals. we've got applied our formula on several pictures and located that it with success will its conversion. The device is compact and useful to the society.

References

[1]. Yi, C., Tian, Y., & Arditi, A. (2014). Portable Camera-Based Assistive Text and Product Label Reading From Hand-Held Objects for Blind Persons. IEEE/ASME Transactions on Mechatronics, 19(3), 808–817.doi:10.1109/tmech.2013.2261083
[2]. R. U. Shekokar , A Survey On Camera Based Assistive Technologies For Visually Impaired Peoples,2016.
[3]. Ramesh Babu.Y, Vision Based Assistive System For Label Detection With Voice Output,2017
[4]. Ch.Sridevi Reddy, Image Processing Based Assistive System For Label Detection With Voice Output For Blind People, IJRRECS/August 2016
[5]. X. Chen, J. Yang, J. Zhang and A. Waibel, "Automatic detection and recognition of signs from natural scenes," In IEEE Transactions on image processing, Vol. 13, No. 1, pp. 87-99, 2004.
[6]. D. Dakopoulos and N. G. Bourbakis, "Wearable obstacle avoidance electronic travel aids for blind: a survey," In IEEE Transactions on systems, man, and cybernetics, Vol. 40, No. 1, pp. 25-35, 2010.
[7]. B. Epshtein, E. Ofek and Y. Wexler, "Detecting text in natural scenes with stroke width transform," In CVPR, pp. 2963-2970, 2010.
[8]. Y. Freund and R. Schapire, "Experiments with a new boosting algorithm," In Int. Conf. on Machine Learning, pp.148–156, 1996.
[9]. N. Giudice and G. Legge, Blind navigation and the role of technology, in The engineering handbook of smart technology for aging, disability, and independence, A.A. Helal, M. Mokhtari, and B. Abdulrazak, Editors. 2008, Wiley: Hoboken, N.J.
[10]. A. Huete, J. Victores, S. Martinez, A. Gimenez, and C. Balaguer. Personal Autonomy Rehabilitation in Home Environment by a Protable Assistive Robot. In IEEE Trans. Systems, Man, and Cybernetics, Part C: Applications and Reviews, Vol. 42, No. 4, pp. 561-570, 2011.