Device for Location Finder and Text Reader for Visually Impaired People

G. Aarthi[1], S. Gowsalya[2], A. Abijith[3], N. Subhashini[4]

[1,2,3]UG Student, Department of Electronics and Communication Engineering, Valliammai Engineering College, Tamil Nadu, India
[4]Assistant Professor, Department of Electronics and Communication, Valliammai Engineering College, Tamil Nadu, India

---***---

Abstract - *Scrutinizing is basic in step by step life for everyone. Apparently incapacitated individuals can peruse just by use of one of a kind applications by them like Braille tongue. The weakness of this system is that every thing does not give the substance in Braille. In this paper, they have proposed an assistive substance examining framework to help apparently blocked individuals to scrutinize compositions from various inquiries in their step by step lives. At first, we get the image of the required, pre-taking care of is performed on it. Pre-dealing with consolidates steps like dull scale and binarization, question of interest affirmation. In the proposed structure, we are making the use of OTSU estimation to change over the diminish scale picture into binarized one. The substance locale from the got picture are then isolated and seen by using optical character affirmation programming (OCR). The rule estimation in OCR to be explicit MODI is used here. This removed substance of various literary styles and sizes at that point can be seen freely and subsequently combined in a word giving its yield as sound using Text-to-talk using the SAPI libraries*

Key Words: — **OCR, Navigation, binarization, Pre processing, Image acquisition, Image recognition.**

1. INTRODUCTION

Perusing is critical and one perspective in our day today lives. Very nearly 314 million outwardly disabled individuals are there all around the universe [1], 45 Million are visually impaired and new cases being included every year according to inquires about done. Developing advances and late improvements in modernized vision, digit cameras, PDA and versatile PCs make it adaptable to help these people by creating picture based items that consolidate PC vision innovation with other existing business innovation, for example, optical character acknowledgment (OCR) stages. Printed content is one of the types of reports, receipts, bank proclamations, eatery menus, classroom presents, item bundles, drug bottles, pennants on street and so forth.

There are many favoring frameworks accessible right now however they have little issues in decreasing the adaptability for the outwardly impeded people. For instance, compact standardized tag perusers intended to support the visually impaired individuals perceive disparate items, it grants the

clients who are unsighted to entrance. Data about these items through discourse and Braille. Be that as it may, a major confinement is that it is exceptionally intense for unsighted individuals to discover the area of the standardized identification and to correct point the scanner tag peruser at the standardized tag.

There are frameworks like K Reader Mobile it keeps running on a phone and license the client to check mail, receipts, fliers, and numerous different things. Despite the fact that, the report to be scrutinize must be about kept undisturbed, set on a cognizant, dim surface (i.e., a no disarray foundation), and contain for the most part entry Furthermore, K Reader Mobile precisely observes dark print on a white foundation yet has issues in perceiving hued content or content on a shaded foundation. It can't peruse content with complex foundations. The principle objective is to advance such a framework, that will examine the writings from composite foundations effectively provide high accurate determination of face-to-face interaction distance. Most proposals for such services give low accuracy guarantees and incur high communication costs.

1.1 RELATED WORK

1.Yee-Ling Lu, Man-Wai and Wan-Chi Siu explains about Text-to-phoneme conversion by using recurrent neural networks trained with the real time recurrent learning(RTRL) algorithm [3].

2.Penagarikano, M.; Bordel, G explains a technique to perform the speech to text conversion as well as an investigational test carried out over a task oriented Spanish corpus are reported & analytical results also.

3.Sultana, S.; Akhand, M. A H; Das, P.K.; Hafizur Rahman, M.M. explore Speech-to-Text (STT) conversion using SAPI for Bangla language. Although achieved performance is promising for STT related studies, they identified several elements to recover the performance and might give better accuracy and assure that the theme of this study will also be helpful for other languages for Speech to-Text conversion and similar tasks [3].

4. Moulines, E., in his paper "Text-to-speech algorithms based on FFT synthesis," present FFT synthesis algorithms

for a French text-to-speech system based on diaphone concatenation. FFT synthesis techniques are capable of producing high quality prosodic adjustment of natural speech. Several different approaches are formulated to reduce the distortions due to diaphone concatenation.

5.Decadt, Jacques, Daelemans, Walter and Wambacq describes a method to develop the readability of the textual output in a large vocabulary continuous speech recognition system when out-of-vocabulary words occur. The basic idea is to replace uncertain words in the transcriptions with a phoneme recognition result that is post-processed using a phoneme-to-grapheme converter. This technique uses machine learning concepts.

2. EXISTING SYSTEM

In Existing system, have endeavored to facilitate the weight on visually impaired individuals by proposing different strategies that changes over content to capable of being heard sounds. Tyflos [11] is two or three glasses that have cameras joined to the side, headphones and an amplifier. Voice orders can be utilized to order the client and direct the stage. A few directions incorporate "draw paper nearer," "move paper up," "move paper up, appropriate" from the gadget to the client, and "rewind section," "forward passage," and "volume up" from the client to the gadget. Regardless, the discourse client coordination probably won't work flawlessly in a boisterous domain, rendering it constrained to indoor use. Finger Reader [12] is one such gadget, a wearable ring with a camera which is available on the front. The voice UI probably won't work consummately in a bedlam encompassing; rendering is limited to indoor need. The proposed system helps blind persons to read product the project aims to implement a reading aid that is small, lightweight, efficient in using computational resources, cost effective and of course user friendly the processor based system can be equipped with a high resolution webcam the microcontroller-based system is easier to use when compared to the mobile one
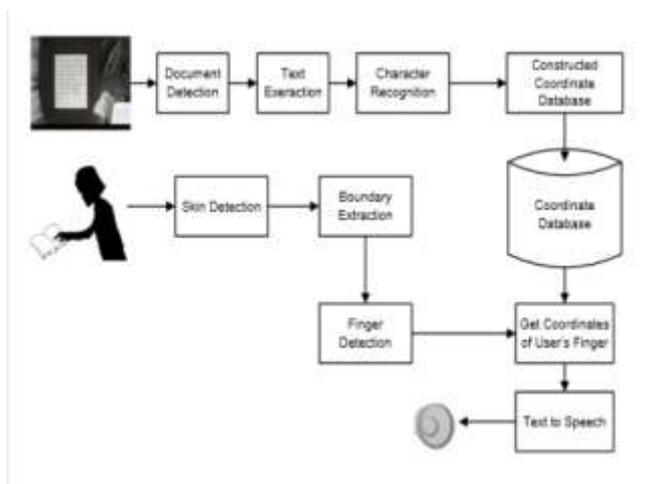
3. PROPOSED SYSTEM

However, the accuracy of the mobile in the conversion efforts is better, primarily due to the high resolution camera built in the device. Emerging technology and in future improvements of this project, this system can be provided with a good and high-resolution webcam contrasted with the one used in this project, and we anticipate, this will improve its certainty. We predict more work will be produced in this critical area of assistive technology, and project that future portable gadgets will have easy to use and built in mechanism as reading aids for the blind, similar, to the mobile-based solution presented here. Labels. Users should capture image and then system read out the text from image. It will be more appropriate for persons those are going through optical surgery. It can be appropriate for road side text recognition so that visually impaired person can travel alone. Along with this, emergency alert and obstacle detection are also to be provided for their better navigation. The proposed system will give capable result as analyzed to most of the existing systems
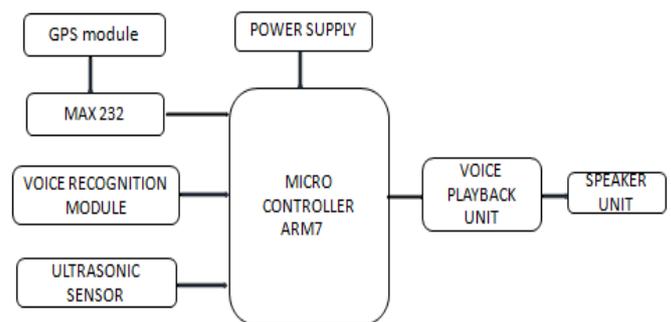
BLOCKED DIAGRAM

Fig 2: Block diagram for obstacle detection and emergency alert.

A. GPS MODULE

The GPS receiver gets a signal from a GPS satellite. The satellites transmit the exact time the signals are sent. So given the travel time of the **GPS** signals from three satellites and their exact position in the sky, the **GPS receiver** can determine your position in three dimensions - east, north and altitude.

B. GSM MODULE

The SIM card mounted GSM modem upon receiving the command from any cell phone send that data to the MC through serial communication. While the program is executed the GSM modem receives command stop to develop an output at the MC, the contact point of which are used to disable the ignition switch.

C. ARDUINO UNO
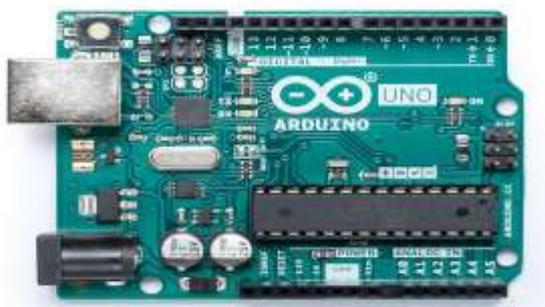
Arduino is an open-source electronics platform based on easy-to-use hardware and software. Arduino boards are able to read inputs - light on a sensor, a finger on a button, or a Twitter message - and **turn** it into an output - activating a motor, turning on an LED, publishing something online.

D.ULTRASONIC SENSOR

Ultrasonic sound vibrates at a frequency above the range of human hearing u**ltrasonic sensors**, like many others, use a single **transducer** to send a pulse and to receive the echo. The **sensor** determines the distance to a target by measuring time lapses between the sending and receiving of the **ultrasonic** pulse.

D. WEBCAM

Webcams are typically small cameras that either attached to the users monitor. It is used to capture the image while placing in front of the camera and recognizing the letters present in that image for conversion
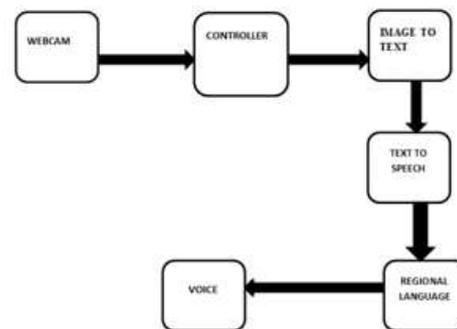
Fig 3: Block diagram for text to image conversion

4. SOFTWARE USED

Raspbian is a free working framework, in view of Debian, improved for the Raspberry Pi equipment. Raspbian Jessie is utilized as the variant is RPi's primary working framework in our undertaking. Our code is written in Python language (adaptation 2.7.13) and the capacities are called from OpenCV. OpenCV, which represents Open Source Computer Vision, is a library of capacities that are utilized for continuous applications like picture handling, what's more, numerous others [14]. Presently, OpenCV bolsters a wide assortment of programming dialects like C++, Python, Java and so on and is accessible on various stages including Windows, Linux, OS X, Android, iOS and so on. [15]. The variant utilized for our undertaking is opencv-3.0.0. OpenCV's application zones incorporate Facial acknowledgment framework, Signal acknowledgment, Human– PC collaboration (HCI), Mobile mechanical technology, Motion understanding, Object distinguishing proof, Segmentation and acknowledgment, Motion following, Augmented reality and some more. For performing OCR and TTS activities we introduce Tesseract OCR and Festival programming. Tesseract is an open source Optical Character Recognition (OCR) Engine, accessible under the Apache 2.0 permit. It tends to be utilized specifically, or (for software engineers) utilizing an API to separate composed, written by hand or printed content from pictures. It bolsters a wide

assortment of dialects. The bundle is for the most part called 'tesseract' or 'tesseract-ocr'. Celebration TTS was created by "The Center for Speech Technology Research", UK. It is an open source programming that has a structure for building proficient discourse blend frameworks. It is multi-lingual (bolsters English, American English and Spanish). As Festival is a piece of the bundle supervisor for Raspberry Pi, it is anything but difficult to introduce.

5. IMPLEMENTATION TEXT TOVOICE CONVERSION:

In project including sentence based examination for expressive of TTS framework determines the issue of making datasets and by creating yield as line by line recognition where the clients think that its hard to review their sentence. In this situation, the granularity of content under examination is normally resolved to be the sentence, as sentence are reasonably short printed portrayals, by seeing this as a sentence by sentence strategy. Unique normal highlights that are utilized to indicate the effect in content are portrayed which matches with the databases put away already and produces the yield aurally. It comprise of three modules Input as Text module, Database module also, discourse module. The info includes diverse highlights with various effects as lexical fillers, sentence splitter, catchphrase spotter and word sense disambiguate.

I) Lexical filler changes over the plain information content into a yield token stream. This module recognizes the conceivable full of feeling compartments (content words), valence shifters, for example, invalidation words and intensifiers.

ii) Sentence splitters parts the sentences in each line which catches the sentence as picture and process it aurally.

iii) Word sense Disambiguate settle the importance of full of feeling words as indicated by their specific situation. It utilizes a semantic similitude measure to score the feeling of an emotional word with the setting word utilizing word net oncology.

Picture procurement:

In this progression, the inbuilt camera catches the pictures of the content. The nature of the picture caught relies upon the camera utilized. We are utilizing the Raspberry Pi's camera which 5MP camera with a goals of 2592x1944.

Picture pre-preparing:

This progression comprises of shading to dim scale change, edge discovery, clamor expulsion, twisting and trimming and thresholding. The picture is changed over to dark scale the same number of OpenCV capacities require the info parameter as a dark scale picture. Commotion expulsion is finished utilizing respective channel. Watchful edge location is performed on the dark scale picture for better identification of the shapes. The twisting and trimming of the

picture are performed by the forms. This empowers us to recognize and extricate just that area which contains content and expels the undesirable foundation. At last, Thresholding is done as such that the picture looks like an examined record. This is done to enable the OCR to productively change over the picture to content.

Picture to content change:

The primary square is the picture pre-handling modules and the OCR. It changes over the pre-prepared picture, which is in .png structure, to a .txt document. We are utilizing the Tesseract OCR. Content to discourse change: The second square is the voice handling module. It changes over the .txt document to an sound yield. Here, the content is changed over to discourse utilizing a discourse synthesizer called Festival TTS. The Raspberry Pi has an on-board sound jack, the on-board sound is created by a PWM yield.

6. SIMULATED OUTPUT

The proposed system consist of a OCR module, where a USB camera which captures the input given in text format and it is sent to OCR process which processes the text and convert it into a speech form. Captured image is sent to MATLAB, this is to fetch word-by-word segmentation from the input image and compare it with templates. The output of detecting text will be processed from each conversion of RGB TO GRAY, GRAY TO BINARY. The texts in the form of (jpeg, png, jpg, bmp etc) are considered for the analysis. The image which is captured from the USB camera is splited in following conditions as described below for detecting corresponding text and matching it with templates prescribed in the below conditions.

- Text in the form of Black and White
- Text in the form of colored
- Text with image
- Text and image merged
- Text with different Font Styles

7. CONCLUSION

The framework empowers the outwardly debilitated to not feel off guard with regards to perusing content not written in braille. The picture pre-handling part takes into account the extraction of the required content area from the complex foundation and to give a decent quality contribution to the OCR. The content, which is the yield of the OCR is sent to the TTS motor which delivers the discourse yield. To take into consideration compactness of the gadget, a battery may be utilized to control up the framework. The future work can be creating gadgets that perform object discovery and separating content from recordings rather than static pictures

REFERENCES

[1] Sanjib Das, "Speech Recognition Technique: A Review", International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 Vol. 2, Issue 3, May-Jun 2012.

[2] Ms. Sneha K. Upadhyay,Mr. Vijay N. Chavda,"Intelligent system based on speech recognition with capability of self learning" ,International Journal For Technological Research In Engineering ISSN (Online): 2347 - 4718 Volume 1, Issue 9, May-2014.

[3] Deepa V.Jose, Alfateh Mustafa, Sharan R,"A Novel Model for Speech to Text Conversion" International Refereed Journal of Engineering and Science (IRJES)ISSN (Online) 2319-183X, Volume 3, Issue 1 (January 2014).

[4] B. Raghavendhar Reddy,E. Mahender, " Speech to Text Conversion using Android Platform", International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622,Vol. 3, Issue 1, January -February 2013.

[5] Kaveri Kamble, Ramesh Kagalkar," A Review: Translation of Text to Speech Conversion for Hindi Language", International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064.Volume 3 Issue 11, November 2014.

[6] Santosh K.Gaikwad,Bharti W.Gawali,Pravin Yannawar, "A Review on Speech Recognition Technique",International Journal of Computer Applications (0975 – 8887)Volume 10– No.3, November 2010.

[7] Penagarikano, M.; Bordel, G., "Speech-to-text translation by a non-word lexical unit based system,"Signal Processing and Its Applications, 1999. ISSPA '99. Proceedings of the Fifth International Symposium on , vol.1, no., pp.111,114 vol.1, 1999.

[8]. Olabe, J. C.; Santos, A.; Martinez, R.; Munoz, E.; Martinez, M.; Quilis, A.; Bernstein, J., "Real time text-tospeech conversion system for spanish," Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '84. , vol.9, no., pp.85, 87, Mar 1984.

[9]Kavaler, R. et al., "A Dynamic Time Warp Integrated Circuit for a 1000-Word Recognition System", IEEE Journal of Solid-State Circuits, vol SC-22, NO 1, February 1987, pp 3-14.

[10] Aggarwal, R. K. and Dave, M., "Acoustic modelling problem for automatic speech recognition system: advances and refinements (Part II)", International Journal of Speech Technology (2011) 14:309–320.

[11] Ostendorf, M., Digalakis, V., & Kimball, O. A. (1996). "From HMM's to segment models: a unified view of stochastic modeling for speech recognition". IEEE Transactions on Speech and Audio Processing, 4(5), 360–

378. [12] Yasuhisa Fujii, Y., Yamamoto, K., Nakagawa, S., "AUTOMATIC SPEECH RECOGNITION USING HIDDEN CONDITIONAL NEURAL FIELDS", ICASSP 2011: P-5036-5039.

[13] Mohamed, A. R., Dahl, G. E., and Hinton, G., "Acoustic Modelling using Deep Belief Networks", submitted to IEEE TRANS. On audio, speech, and language processing, 2010.
[14] Sorensen, J., and Allauzen, C., "Unary data structures for Language Models", INTERSPEECH 2011.