Smart Glasses for Blind and Illiterate People

Keerthi Kiran A, Pranay Teja B, Laxmi Priya P Department

of Electronics and Communication EngineeringNational

Institute of Technology, Warangal

Email: bp942050@student.nitw.ac.in

Abstract – The human eye is an organ which gives us the sense of sight. Blindness is œof the important issues to ponder in current environmental situation as it has major effect on human health. There are currently 284 million people in the world who are visually impaired and 39 million blind people. Internet of Things (IoT) has transformeddevices to remote monitoring in the health care sector round the clock. An assistive technology involve development of devices to cater the needs of people suffering with different disabilities. The proposed prototype in this paper deals with Smart glasses which are developed to assist the blind people to read and translate the typed text which is written in English language. Prototype is designed using Raspberry pi board and appropriate sensors to capture the text in the picture and translate into speech using Optical Character Recognition technology (OCR), OpenCV with Tesseract, Efficient and Accurate Scene Text Detector (EAST) and Text to Speech technology (gTTS). This prototype is also embedded with GPS and GSM modules to provide a panic button for enabling the visually impaired person to inform care takers in an emergency.

Index Terms – IoT, Blindness, healthcare, OCR, EAST, OpenCV, Tesseract, gTTs.

I. INTRODUCTION

In our lives, there are people suffering from many diseases and illnesses. According to the Indian Journal of Ophthalmology [1], there are 4.95 million blind people in India including 0.24 million blind children and 70 million people with Visually impaired problems. However, only 1% of the blind people in India [2] are literate which is far lower than the regular literacy rate (77.7%). Since there are not many special schools for people who require special care, and that most of them are either private or expensive. Hence, the majority of blind people and people with vision difficulties didn't pursue education.It is observed that a few people gain basic knowledge from their parents just by studying at their home. Most of the people generally believe that blind people and vision impaired people cannot live alone and need help all the time. With the advancement in the technology, different cyber physical devices are being developed to aid people suffering with disabilities to lead a normal life independently. The main reason for the Implementation of Smart Glasses in this work is to convey that all blind people and vision impaired people have a chanceto live a normal life and pursue their career with the help of for help all the times.

During Literature Survey, it is noticed that a few prototypes are developed to aid visually impaired people using various wireless technologies and sensors. Smart Glasses with google vision API [3] developed for a navigation system that makes use of wearing smart glasses and a sensor to continuously take photographs of the environment. This system would solve direction-finding issues for people who are blind or visually impaired. Smart glasses with Arduino pi 2 [4] developed to assist visually impaired people by text recognition technology that can help reading the text from the hardcopy materials. The goal is to help blind people for carrying out their daily tasks utilizing the benefits of wearable devices.

Arduino based Smart Glasses [5] is a low-cost solution using ultrasonic sensors, gas sensors and LDR to detect the obstacles, smoke and darkness, respectively and alert the user by activating the buzzer. The Smart Glasses designed with a combination of Google Glass [6] and the educational application called English Today created the new educational chances for persons who are blind or deaf. The English Today application is used by the students on mobile phones (Android and iOS). EPSON BT-300 [7] is a wearable smart glass developed using YOLO Technology with Intel Atom X51.44GHZ CPU to detect the objects and inform the user about the objects using GTTS module.

Smart Glasses using ATMEGA 382p Microcontroller [8] developed to overcome the traveling difficulties. After receiving information from the environment, it is translated into audio and passed to the blind person through a headphone. Using this smart glass, visually impaired people can guide themselves in an indoor and outdoor environment hassle-free. Smart glass developed for public signs recognition systems [9] can detect and recognize the public signs in cities, and give corresponding voice note to the blind person. Smart glass developed using ESP32 [10] is to boost productivity and interconnect computing devices into our daily life by presenting the information right in front of the user's eyes. Smart glasses, developed using RGB-D [11] camera for glasses, built on VGG Very Deep Convolutional Networks (VGGNet) architecture. It provides Image Captioning to describe the environment and street segmentation for more accurate street view and the audio I/O periphery for the NLP dialogue system.

This paper presents the prototype of a reliable smart wearable for the visually impaired people, which scans any text image and convertsit into audio, enabling the users to hear through a headphone connected to the glasses. In addition, the proposed prototype is also embedded with a features to translate the audio into the required language depending on the user choice and panic button to help the user in an emergency to inform his care taker. From the literature survey, it is observed that no prototype was proposed for wearable glasses embedded with the above features. The proposed prototype can also be used by illiterate people. It enables a faster learning process, making the experience of the user memorable.

The Organization of this paper is as follows. Section II presents the design details and specifications of various modules used in developing the prototype. Section III presents the features and technologies used in the protype. The Hardware implementation is presented in section IV. Results obtained by testing the prototype under various conditions are discussed in Section V. Conclusions are presented in Section VI.

II. Proposed Prototype Design

In this proposed work, a prototype shown in Fig.3 is developed to aid the blind or illiterate people for converting the text to speech using Raspberry pi 3B+ model board (see Fig.1) and sensors (Fig.2). The block diagram of prototype designed using this minicomputer and sensors is shown in Fig.3. Raspberry pi 3 b+ is a minicomputer with arm processor that is used to for image processing functions that requires a higher clock speed performance, consists of Wi-Fi and audio output port. Ultrasonic Sensor (HC-SR04) is used to detect an object that is of range 2-400cm and Web camera (Logitech SC- 270) captures the object for text detection.

The raspberry pi board is loaded with Raspbian Stretch OS to process the image. The image captured by the webcam is send to pi board through USB. To extract the text from the image open cv tesseract and east libraries are deployed in the board for real time computer vision function. gTTS API is applied converting the text to speech. Earphone connected through the 3.5 mm headphone jack to output the generated audio from the image.

Fig. 1 Raspberry Pi 3 B+ Board

Fig. 2 Sensors

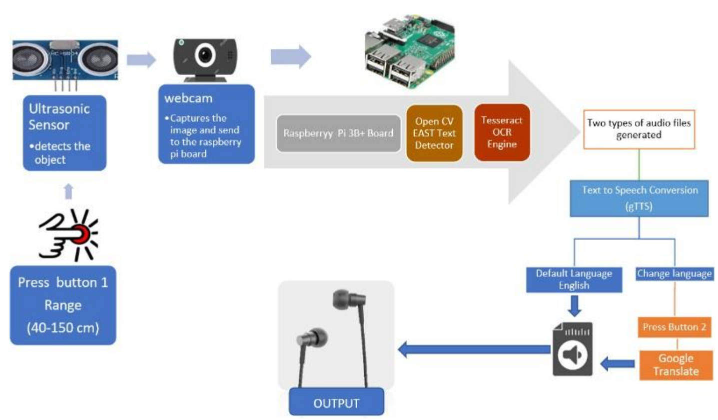


Fig.3 Block Diagram

To convert picture text to speech the user pressesbutton 1, the prototype gets activated, Ultrasonic sensor checks if any object is in the distance of 40-150cm. If anyobstacle is found within the range, then the web camera activates and there by captures a photograph. This photograph is processed using a minicomputer (Raspberrypi model 3B+) which uses many technologies to convert the picture text to speech. The prototype also has a translation feature which activates by clicking the button 2. The output speech is given through the headphone. Thefeatures and functionality of these modules shown in the block diagram (Fig.3) are explained in the following subsections

A. RASPBERRY PI MODEL 3B+

Raspberry Pi is a computer the spacing of a credit card. A power supply, monitor, SD card, keyboard, mouse, and operating system must be configured on it. The Raspberry Pi is an affordable embedded device that can carry out many important tasks. It may function as a basic PC, a portable computer for programming, a hub for improvised electronics, and more. GPOI (general purpose input/output) pins are present for controlling electronic parts. With a 64-bit four core CPU operating at 1.4GHz, dual-band 2.4GHz and 5GHz wireless LAN, Bluetooth 4.2/BLE, faster Ethernet, and PoE functionality through a separate PoE HAT, the Raspberry Pi 3 Model B+ is the newest device in the Raspberry Pi 3 line. The modularity conformity certification for the dual-band wireless LAN enables the board to be included into finished products with much less wireless LAN compliance testing, reducing both times to market and cost.

B. A9G DEVELOPMENT BOARD

A9G Development board is a quad band GPS+GSM/GPRS module which is based on the RDA8955 Chip. Mobile SIM card or IOT is required to enable the device functions such as GPS, GPRS, SMS and Voice calls. This board adopts Micro SIM card (15mm x 12mm). This board comes with 29 GPIOs and works on Low Power Mode Current. The maximum download rate is 85.5 kbps, upload rate is 42.8 kbps.

C. ULTRASONIC SENSOR

Ultrasonic sensors' main function is to use ultrasonic waves to measure distance. Sensors that use ultrasound produce the waves and then catch the reflections. The ultrasonic sensor will then measure the distance to the item at this point. This ultrasonic sensor hasa range of 2-400 cm. To distinguish the text from the textpicture in Prototype the Ultrasonic sensor measures the distance between the camera and a subject. Since this is the necessary range to take a clean image, the distance should be between 40 cm and 150 cm.

D. BUTTON

One of the primary tasks of our developed prototype is to translate the whole text or some words of the subject by pressing a button that is connected to the glasses. For the purpose we will be using a manual switchor a button. User's eyes will be captured on camera for theproject. When the button is clicked, the camera will take aphoto in order to extract and recognize the text from the image.

E. WEB CAMERA

Web camera is a digital video camera used to capture image and video that is connected to a computer or computer network. The captured image is transferred tocomputer using USB, Bluetooth, or Wi-Fi protocols. It captures the Image at 720p/30 fps and widely used in virtual real-time interaction.

III. TECHNOLOGIES

The technologies and libraries used in the developed prototype are explained in the following subsections.

A. RASPBERRY PI OS

Raspberry Pi OS is a free Linux operating system based on Debian optimized for the Raspberry pi family ofmini or compact computers. It is optimized with ARM CPUs Pico sized microcontroller. It uses LXDE desktop environment with open box stacking. It is integrated with computer algebra system VLC, Wolfram Mathematica anda lightweight version of the Chromium web browser. Being deployed via APT it consists of 35,000 packages and bundles of pre compiled software.

B. OCR

The Optical Character Recognition Technology (OCR) procedure used to transform an image of text into a machine-readable text format is known as optical character recognition (OCR). The computer will store the scan as a digital image, for instance, if you scan a form ora receipt. The words in the picture file cannot be edited, searched for, or counted using a text editor. The image may be transformed into a text document with its contents saved as text data using OCR, though.

C. OpenCV

Open CV has been used for detecting the text in the image. It is a popular pre-built, open-source library (package) for CPU-only computer vision, machine learning, and image processing applications. Python is among the many programming languages that it supports.

D. EAST

Efficient and Accurate Scene Text Detector(EAST) technique that enables text detection in real settings, which achieves excellent accuracy and efficiency. This study employed three datasets: ICDAR

2015, COCO-Text, and MSRA-TD500. The study has demonstrated that this strategy produces more accurate and effective outcomes than earlier techniques.

*E. g*TTS

In order to convert the text into speech, the developed prototype uses Google Text to Speech (gTTS)), a Python library and Command Line Interface(CLI) tool to interface with Google Translate text-to-speech API.

F. TESSERACT

Tesseract with Open CV is a free and open-source OCR engine that complies with the Apache 2.0 license. For programmers, it may be linked topically or through theuse of an API to extract written text from photos. Many different languages are supported. Tesseract doesn't come with a built-in graphical user interface, however, there areplenty on the 3rdParty website. Tesseract is interoperable with a wide variety of frameworks and languages. It may be used in conjunction with the current layout analysis to identify text within a huge document or with an outside text detector to identify text from a picture of a single sentence line.

G. GOOGLE TRANSLATION

To transcribe text, documents, and websites from one language into another, Google developed Google Translate, a multi-language neural machine translator tool. It provides a website interface, an Android and iOS mobile app, and an API that aids creators in creating specific software apps and browser plug-ins.

IV. Hardware Implementation

Fig. shows the implementation of the smart glasses designed for visually impaired people and Fig. shows the controller designed to operate smart glasses. This controller is programed to read the start signal issued by the user to send a signal to camera get the image of text captured by the camera when user presses. It may be noted in the image of the prototype that a camera, and ultrasonic sensor are attached to glasses. The camera attached to the smart glasses should support minimum resolution of 720p to capture the text available within the range of the person, whenever the user presses a button. The device checks the distance, when the distance of the objects(text) is in the capturing range then the image is captured using the webcam and saved into the device. The Open CV EAST detector detects the text position in the image and organizes in the orderly fashion. The detected text information is passed to the Tesseract OCR to extract the text from the image. The text containing small error causes the misspelling of words. In order to overcome this issue and language translation, google Translation is employed. After successful text detection the text is then converted to speech of required language(default English) using the gTTS Library and the audio file is generated. The Audio file is played over the headphones attached to Prototype.

The prototype proposed for Smart Glasses to aid blind people are developed using raspberry pi 3B+ model board and appropriate sensors to convert image text to speech is shown in fig.4. The sensors are mounted onto the wearable glasses to capture the real-time images instantly and convert the text to speech. The Raspberry pi board and power bank is enclosed in a case to make it compact and easily wearable over the arm. The raspberry pi is connected to sensor over the USB wire to webcam and power bank of power supply of 5V and 2Amp, GPIO pin to ultrasonic sensor. The Raspbian Stretch flashed to SD card using Win 32 Disk Imager. The SD card being inserted to the board Python 3 lib using the Board connected to HDMI, Keyboard and Mouse. The two PushButtons Attached to Raspberry Board one being used to capture the image and other being to translate the language.

Fig.4 (Smart Glasses)

Fig. 4.1 (Control Unit)

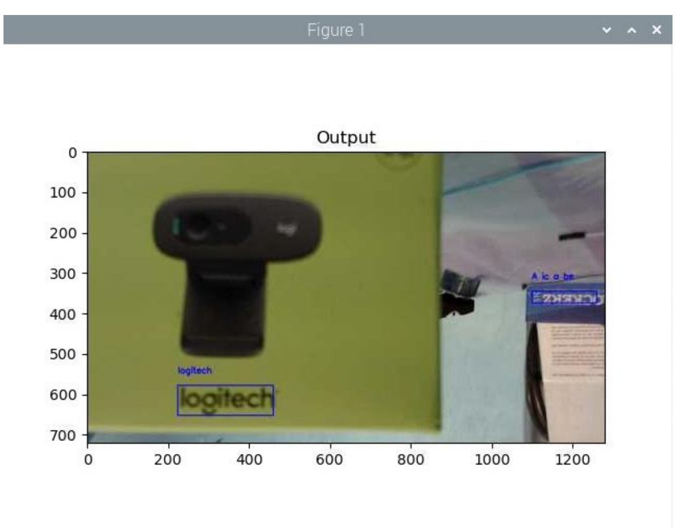
Fig. 4.3 (Person wearing Smart Glasses)

V. RESULT AND DISCUSSION

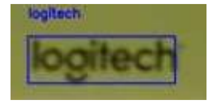
The prototype of Smart Glasses developed in this work is shown in Fig 4. This device is practically found to be portable and easy to maintain. The Glasses designed with the control box holds all the circuitry. The control box has the buttons which specifies respective operations. By pressing Button-1 it converts the picture text to speech. Button-2 translates the speech from English to Hindi.

BUTTON-1:

When Button -1 is pressed the text detected in the picture is given as speech in the headphone connected to the board.

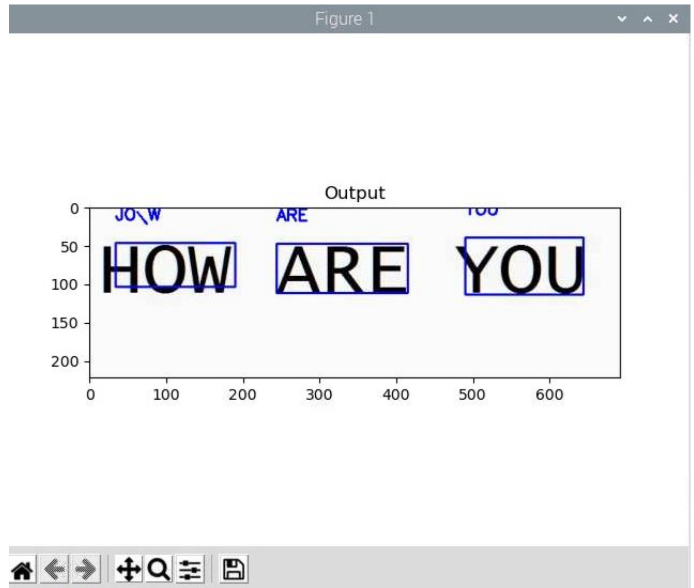



```
High Performance MPEG 1.0/2.0/2.5 Audio Player for Layer 1, 2, and
3.
```



BUTTON-2

When Button-2 is pressed the translation of speech happens from English to Hindi through gTTS.




```
आप कैसे हैं
completed2
completed3
High Performance MPEG 1.0/2.0/2.5 Audio Player for Layer 1, 2, and
3.
```

REFERENCES

[1]. Mannava, S., Borah, R. R., & Shamanna. (June 2022). B R. Current estimates of the economic burden of blindness and visual impairment in India: A cost of illness study. *Indian Journal of Ophthalmology*, 2141-2145. doi:10.4103/ijo.IJO_2804_21

[2]. G Murthy, S. V. (2005 Mar). Current estimatesof blindness in India. *The British journal of ophthalmology*, 257–260. doi:10.1136/bjo.2004.056937

[3]. Rajendran, P. S., Krishnan, P., & Aravindhar, D. J. (2020). Design and Implementation of Voice Assisted Smart Glasses for Visually Impaired People Using Google Vision API. *4th International Conference on Electronics, Communication and*

Aerospace Technology (ICECA), (pp. 1221-1224). doi:10.1109/ICECA49313.2020.9297553

[4]. Ali, M. &. (2016). Smart Glasses for the Visually Impaired People. *International Conference on Computers Helping People with Special Needs*, (pp. 579-582). doi:10.1007/978-3-319-41267- 2_82.

[5]. P. Samuda, N. G. (2022). Arduino based Customized Smart Glasses for the Blind People. *econd International Conference on Artificial Intelligence and Smart Energy (ICAIS)*, (pp.1136-1141). doi:10.1109/ICAIS53314.2022.9742799

[6]. Maly, A. B. (2019). Smart Google Glass Solution Used as Education Support Tool. *International* Symposium on Educational Technology (ISET), (pp. 265-267). doi:10.1109/ISET.2019.00063

[7]. J. -Y. Lin, C. -L.-J.-C.-C. (2020). Smart Glasses Application System for Visually Impaired People Based on Deep Learning. Taiwan 2nd International Conference on Computing, Analytics and Networks (Indo-Taiwan ICAN), (pp. 202-206). doi:10.1109/Indo-TaiwanICAN48429.2020.9181366

[8]. Miah, M. R. (2018). A Unique Smart Eye Glass for Visually Impaired People. 2018 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE). Dhaka. doi:10.1109/ICAEEE.2018.8643011

[9]. Feng Lan, G. Z. (2015). Lightweight smart glass system with audio aid for visually impaired people. IEEE Region 10 Conference, TENCON, (pp. 1-4). doi:10.1109/TENCON.2015.7372720

[10]. S. Das, S. S. (2021). Low Cost Smart-Glass using ESP-32. 2021 IEEE 2nd International Conference on Applied Electromagnetics, Signal Processing, & Communication (AESPC), (pp. 1-6). doi:10.1109/AESPC52704.2021.9708478

[11]. Shiu, C. -H.-F. (2020). Smart guiding glasses with descriptive video service and spoken dialogue system for visually impaired. 2020 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan), (pp. 1- 2). doi: 10.1109/ICCE- Taiwan49838.2020.925