JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Design and Development of Aural Companion: A Revolutionary Device for Visually Impaired Reading Assistance

[1]Hemanandhini N, [2]Kavinkumar N, [3]Nithin Nageshwaran K, [4]Sivaramakrishnan Rajendar

[1]Student, [2]Student, [3]Student, [4]Professor

[1]B.E Computer Science and Engineering,

[1]KPR Institute of Engineering and Technology, Coimbatore, India

Abstract: **Assistive devices play a crucial role in addressing the needs of individuals with visual impairments, facilitating their access to printed text. However, existing assistive systems encounter several limitations, including inadequate accessibility, high costs, and cumbersome interfaces. To overcome these challenges, this paper proposes a system called Aural Companion, an innovative device designed to revolutionize reading assistance for the visually impaired. By integrating state-of-the-art technologies such as Raspberry Pi, Python, and the Tesseract OCR engine, the system offers seamless conversion of printed text into natural audio. Unlike conventional systems, the Aural Companion boasts enhanced accessibility, affordability, and user-friendliness. Through the utilization of IoT principles and voice commands, users can navigate text effortlessly, empowering them to engage in reading activities independently. The findings demonstrate that the Aural Companion outperforms existing assistive systems in terms of accuracy, speed, and cost-effectiveness, heralding a new era of accessibility for individuals with visual impairments.**

Keywords – Raspberry Pi, Tesseract OCR engine, Google Text – Voice, Natural Language Processing

I.INTRODUCTION

The eyes are a crucial instrument for human observation and comprehension of the environment, allowing humans to carry out their jobs. Visual impairment creates several problems in the lives of visually impaired persons. As a result, it is critical to focus on the requirements of the visually impaired population. The World Health Organisation (WHO) divides distant vision impairment into four categories: mild, moderately visually impaired, severely visually impaired, and blind. Vision, the most dominant of our senses, plays a critical role in every facet and stage of our lives. We take vision for granted, but without vision, we struggle to learn, to walk, to read, to participate in school and to work [1].

Researchers study from several perspectives to assist visually impaired persons in leading regular lives. The digital era has significantly improed the lives of the visually impaired people by making life more convenient. Blind persons have an illness or an unintentional lesion to the optic nerve or eye that has rendered them blind in both or one eye. Most individuals believe that all people with vision impairments are blind. In truth, blindness is a form of visual impairment. People with vision impairment are not always blind, and some of them can see objects[2].

Blind individuals encounter a myriad of challenges in accessing various forms of information. Traditional printed materials like books, newspapers, and documents pose difficulties due to the lack of tactile feedback and limited availability of Braille materials. Additionally, the digital landscape presents hurdles as many online resources and documents lack accessibility features, impeding visually impaired individuals' access to information on the internet. Moreover, the cost associated with specialized assistive technology exacerbates the situation, making these resources inaccessible to some individuals with visual impairments.

Various technologies help visually impaired individuals access and interact with digital content. Screen readers like JAWS, NVDA, and VoiceOver convert text to speech, allowing blind individuals to listen on computers or mobile devices. Braille displays turn digital text into Braille, facilitating reading through touch. OCR software scans printed text, converting it into a digital format for screen readers. Accessible e-books improve reading with adjustable text size, customizable fonts, and screen reader compatibility. Additionally, mobile assistive apps convert text to speech, identify objects through image recognition, and provide navigation support for the visually impaired.

Assistive technologies have limitations affecting their effectiveness. Not all content is universally accessible, as websites and software often neglect accessibility. Complex layouts, graphics, and non-standard formatting pose challenges for existing technologies. Some devices rely on the internet, which is problematic in areas with poor coverage. High costs make specialized tools inaccessible for some. Users, especially those new to assistive devices, face a learning curve, adding to the challenges.

Keeping these factors in mind, the proposed systems aims to develop an innovative assistive technology named "Aural Companion," a wearable device designed to cater to the unique needs of individuals who are blind or visually impaired. The primary objective is to seamlessly convert printed text into natural audio, enhancing the reading experience for the visually impaired community. By integrating principles of Internet of Things (IoT), Natural Language Processing (NLP), and Optical Character Recognition (OCR), the device will employ a camera to capture text content. Users can seamlessly control reading activities by issuing voice commands, including starting, stopping, or replaying content. The Aural Companion aims to enhance accessibility and independence for visually impaired individuals in their reading experiences.

The subsequent sections of the paper are structured as follows: Section 2 provides an overview of existing literature related to assistive technologies for the visually impaired, discussing various approaches and technologies employed in similar projects. Section 3 outlines the architecture of the proposed system, detailing the integration of hardware and software components. Section 4 elaborates on the design and implementation of the Aural Companion device, covering hardware specifications, software modules, and their integration. In Section 5, the experimental and system performance are finally shown and discussed. This is followed by a conclusion and recommendations for future directions in this field of study.

II. LITERATURE REVIEW

In recent years, there has been a concerted effort to develop assistive technologies aimed at enhancing the accessibility of printed materials for visually impaired individuals. This section reviews two notable research works in this field.

A. Handheld PDA-Based System for Visually Impaired Individuals

Peters, Thillou, and Ferreira (2004) proposed a handheld PDA-based system to assist blind people in their daily tasks. The system, developed by the Centre for Resources and Evaluation of Technologies adapted to the Handicapped (CRETH), follows a user-centred design approach. It combines users' involvement and engineers' efforts in a continuous process, facilitated by a specialist able to communicate with both parties. The system operates through a main loop, starting with the user taking a snapshot, followed by text/picture detection, optical character recognition (OCR), text-to-speech (TTS) synthesis, and feedback to the user. Each task is carried out by algorithms integrating both technical performance and user requirements. (Peters et al., 2004) [9]

B. Raspberry Pi-Based Device for Reading Assistance

Adusumilli et al. (2023) proposed a reading aid and translator device using Raspberry Pi to assist visually impaired individuals. With 1.3 billion people suffering from visual disabilities worldwide, reading poses a significant challenge. The device consists of a camera connected to Raspberry Pi, capturing printed text. Optical character recognition (OCR) software installed on Raspberry Pi converts the captured text to audio speech. The text-to-speech (TTS) engine within Raspberry Pi then reads out the captured text, facilitating access to printed materials. This implementation supports multiple languages, allowing users to choose their preferred language for audio output (Adusumilli et al., 2023) [10]

C. Comparison and Analysis

Both projects aim to address the challenge of reading for visually impaired individuals through innovative technological solutions. The handheld PDA-based system emphasizes a user-centered design approach, integrating user feedback throughout the development process. On the other hand, the Raspberry Pi-based device offers a portable and affordable solution for reading assistance. Both projects utilize the optical character recognition and text-to-speech synthesis to convert printed text into audio output. However, the Raspberry Pi-based device extends its functionality by supporting multiple languages, enhancing its usability for a broader range of users.

In contrast to existing systems, the proposed system overcomes common limitations. It eliminates costly Braille printing by utilizing a Raspberry Pi-based system for real-time OCR and TTS conversion. Additionally, it supports multiple languages, providing a more accessible solution for visually impaired individuals.

III.SYSTEM ARCHITECTURE

By combining hardware and software, Aural Companion creates an intuitive paradigm of interaction that empowers individuals who are visually impaired. When the user gives orders verbally, the device's microphone records the input. The NLP component then analyses the data to determine the user's intention. The OCR module then examines the text images that the camera modules have taken to extract textual content[4]. The extracted text is then subjected to text-to-speech synthesis using the Google Text-to-Speech engine, culminating in the user being provided with a natural-sounding audio output through the device's speaker. [3].The architecture displayed in Figure 1 illustrates the complete procedure.

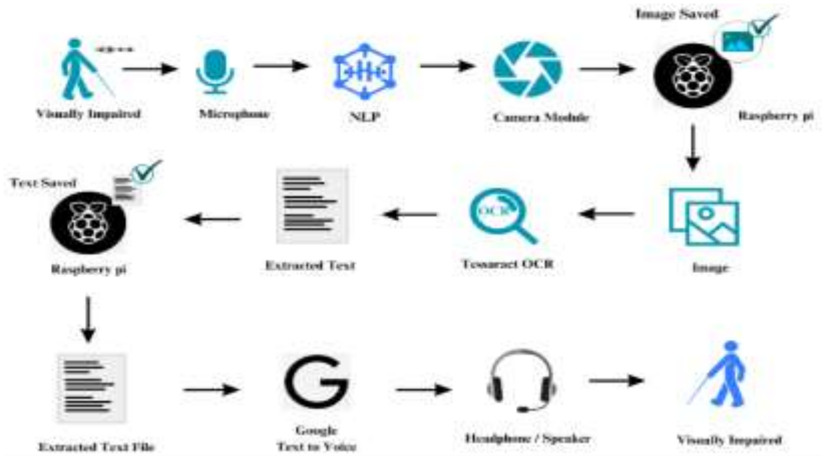

Fig. 1.Architecture of the Proposed System

IV.DESIGN AND IMPLEMENTATION

A.Hardware

1) Raspberry Pi Model 3 B+ - A system-on-a-chip (SOC) that combines multiple functional components into a single chip or chipset, the Raspberry Pi is a SOC. The Raspberry Pi foundation developed the Raspberry Pi, an ARM-based single-board computer the size of a credit card. In addition, the Raspberry Pi's SD card should contain the operating system and all programme files required by the device[6]. Formatting the SD card is a prerequisite to installing the OS. To connect to a network, an Ethernet cable is placed into the designated slot.

2) Pi Camera –The Raspberry Pi Camera Board is a particularly designed add-on module for Raspberry Pi hardware. It connects to Raspberry Pi hardware using a special CSI interface. The sensor's native resolution in still capture mode is 5 megapixels. In video mode, it can record in up to 1080p at 30 frames per second. For mobile projects, the camera module is a fantastic alternative due to its small size and light weight.

3) Headphones or Speaker, Microphone – The user can speak and hear the text from the image using this for input and output.

4) SD Card – An SD card is used as additional storage for the Raspberry Pi. Due to the limited memory capacity of the Raspberry Pi, the SD card serves as a means to store the operating system and program files required for the device. The SD card is formatted to install the OS and provides expanded storage capacity for data and applications.

B.Software

1) Tesseract – Tesseract OCR stands as a pioneering open-source OCR engine, renowned for its accuracy and versatility in extracting text from images. Originally developed by Hewlett-Packard Laboratories and later maintained by Google, Tesseract OCR has undergone significant advancements to become one of the most widely used OCR engines worldwide.

2) Google Text-to-Speech – The advanced text-to-speech engine known as the Google Text-to-Speech was created by the company and is well-known for producing speech output that sounds natural and excellent. Thanks to this technology, printed text can be read aloud on computers and other devices with human-like tone, pronunciation, and rhythm.

C.Integration

The integration that involves in configuring the Raspberry Pi with all the necessary software, connecting the camera module and microphone to the GPIO pins, and setting up the system to utilize these components effectively. The OCR software, Tesseract OCR, reads text from images captured by the camera [5], while the eSpeak NG TTS engine converts this text into spoken words. Finally, the audio feedback is relayed through the headphones or speaker connected to the Raspberry Pi.

V.EXPERIMENTATION AND DISCUSSION

A. Screenshots

In this section, the images captured during the experimentation process is shown which provides visual insight into the performance of the system.

Fig. 2.Original Image

The Figure 2 depicts an example of the original printed text that the system processes.

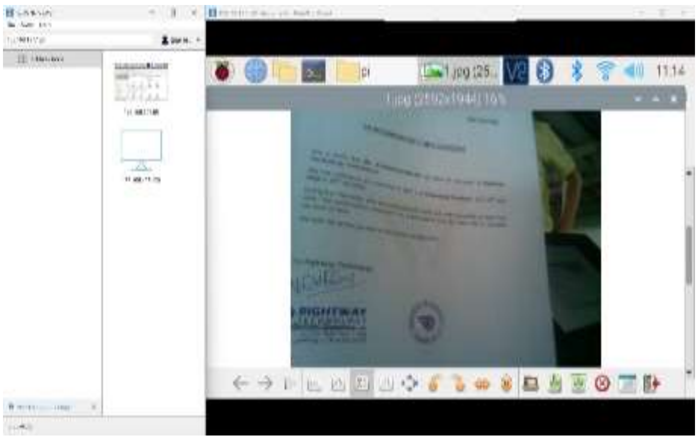

Fig. 3.Captured Image in Raspberry Pi

The Figure 3 shows the text captured by the Raspberry Pi camera module. The captured picture serves as input for the Optical Character Recognition (OCR) process. The Figure 4 shows the complete setup of the proposed system including Raspberry Pi, Camera and Microphone

Fig. 4. The Complete Setup of the Proposed System

B.Input Processing

The system receives photos that are taken by the Raspberry Pi camera module.These images contain printed text that needs to be converted into audio for the user. The input processing involves the following steps:

Image Capture: The printed text is imaged by the Raspberry Pi camera module are

1) Image Preprocessing: Preprocessing techniques are applied to enhance the quality of the captured images, such as noise reduction and contrast enhancement.

2) Text Detection: The captured images are analyzed to detect the regions containing text using techniques like edge detection and contour detection.

3) Text Extraction: Once the text regions are identified, the text is first extracted from the image using the optical character recognition (OCR) techniques[7].

C.Output Processing

The output of the system is the synthesized audio of the extracted text. The output processing involves the following steps:

1) Text-to-Speech Synthesis: Extracted text is converted into audio using text-to-speech synthesis techniques.

2) Audio Playback: The synthesized audio is played back to the user through the device's speaker or headphones.

3) Feedback to User: The user may provide feedback or issue commands to the system, which are processed to control the reading experience.

D.Performance

1) Processing Time Per Image: OCR processing time per image is approximately 10 seconds, subject to updates, while text-to-speech synthesis time per image is assumed to be two seconds.

2) Total processing time per image:

Total processing time per picture = picture capture time + OCR processing time + text-to-speech synthesis time.

Total processing time per image = 3 seconds (capture), 10 seconds (OCR), and 2 seconds (synthesis) = 15 seconds.

3) The number of images processed per minute:

Capturing each photograph within 15 seconds results in a rate of 4 images per minute, with 60 seconds equivalent to 4 photos captured every minute.

4) Throughput:

The throughput of the proposed system, determined by the processing time for OCR and image capture, is four photographs per minute.

VI. CONCLUSION AND FUTURE WORK

The proposed Aural Companion offers an advanced solution for visually impaired individuals to access printed text independently and efficiently. By leveraging technologies like Raspberry Pi, Python, and the Tesseract OCR engine, the proposed system surpasses traditional methods like Braille and finger reading devices in accuracy and speed. Looking ahead, there are promising avenues for future enhancements and research. These include expanding language support, improving the user interface for easier interaction, integrating advanced features like object recognition and navigation assistance, miniaturizing the hardware for portability, and exploring integration with wearable technology such as smart glasses for a seamless reading experience. These advancements will further elevate the accessibility and usability of the system, benefiting the visually impaired community.

VII.REFERENCES

[1] [online] Available: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment/
[2] Wang, J., Wang, S., & Zhang, Y. (2023). Artificial intelligence for visually impaired. Displays, 77, 102391.
[3] B Nithya Santhoshi and H Rithika, "Image Text to Speech Conversion In The Desired Language By Translating With Raspberry Pi", the proceedings of IEEE International Conference on Computational Intelligence and Computing Research, 2016.
[4] S Aaron James, M Monisha, and S Sanjana, "OCR based automatic book reader for the visually impaired using Raspberry PI", International Journal of Innovative Research in Computer and Communication Engineering, vol. 4, no. 7, pp. 1111-1118, 2016.
[5] K. R. Arun, S Parthsarathy, Umamaheswari S Srilakshmi and D Velmurugan, "Hardware Implementation of Smart Reader for Visually Impaired People Using Raspberry PI", International Journal of Advanced Research in Electrical Electronics and Instrumentation Engineering, vol. 5, no. 3, pp. 2055-2063, 2016.
[6] [online] Available: https://www.raspberrypi.org
[7] [online] Available: https://en.m.wikipedia.org-OCR
[8] [online] Available: https://pypi.python.org-gTTS
[9] Peters, A., Thillou, C., & Ferreira, F , A handheld PDA-based system for assisting blind individuals in daily tasks , 2004
[10] Adusumilli, (2023). Reading aid and translator device for visually impaired individuals using Raspberry Pi.