ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

[Computer Stream]

TITLE: OPTICAL CHARACTER RECOGNITION AND TEXT TO SPEECH SYSTEM TO ASSIST THE BLIND

GROUP MEMBERS

NAME	ID No
1. Adam Abera	UGR/0936/12
2. Ebisa Adugna	UGR/ 7063/12

ADVISOR: Dr. Menore Tekeba

Signature: _____

Date of Submission: January 29, 2023

Abstract

People with partial and full blindness have been using traditional devices like walking sticks or canes to interact with the world until very recent times. However, in this modern age, new devices have been designed to solve the problems that the traditional devices could not solve. They provide innovative solutions to the problems of the visually impaired by allowing them to read text, get spoken directions and identify objects. Although these devices provide solutions, they come at the sacrifice of cost, complexity and not being able to be used by the fully blind. The problem of finding a device that can read text for the partially and fully blind by giving them an audio version of the text is evident. Not only this but the device needs to relatively cheap, simple to implement and repair, and able to be used by both the partially and fully blind. This project aims to design and implement the device that solves this problem. This project is planned to be implemented using the following methodology. The system will be designed and then the components will be purchased. Afterward the code that will run on the device is written and upload to the components when they are delivered and connected according to the design. When the previous tasks have been completed, the device is rigorously tested and all the errors and bugs will be ironed out. The implementation will use a microcontroller to capture images which will then be processed on a smartphone or computer. The smartphone or computer will use Optical Character Recognition and Text to Speech to generate an audio version of the text found in the images. As such, this project aims to implement this device with the previous steps at a relatively low price and within the span of four months.

Contents

List of Figures and Tables

Acronyms

AI	Artificial Intelligence
GPS	Global Positioning System
CINVESTAV	Center for Research and Advanced Studies of the National Polytechnic Institute
MIT	Massachusetts Institute of Technology
OCR	Optical Character Recognition
Wi-Fi	Wireless Fidelity (Wireless Internet Access Technology)
IDE	Integrated Development Environment
API	Application Programming Interface
HTML	Hyper Text Markup Language
JS	Javascript

Introduction

People who live their lives with full or partial blindness usually struggle to interact with the world the same way a person with normal vision would. Most of their struggles stem from them not being able to read text that is not in Braille and identify objects around them. Over the course of human history, people with full or partial blindness have used several tools to help them interact with the world. A good example of the devices used even to this day is the walking stick or cane. This has been used by people with full or partial blindness to identify their surroundings by using the walking stick/cane to assess what is on the ground around them or if there are obstacles or holes. This however is a relatively mature device that people with impaired vision have been using for a long time. Now, new technology has given rise to products that better help people with impaired vision.

Some of the modern devices that have been developed or are being developed are trying to expand on what the walking stick/cane could do and also add new ways to help people with impaired vision. The Assisted Vision Smart Glasses developed by Stephen Hicks [1] from the University of Oxford allows people with partial blindness to allow them to walk around their surroundings freely. This device allowed people with partial vision to see things by showing them items in their surrounding by using dark and bright colors. This was done because people with partial blindness could still distinguish brightness and darkness. However, this device only works for people with partial blindness but not for people with full blindness.

One other device was the AI Glasses developed by Center for Research and Advanced Studies of the National Polytechnic Institute (CINVESTAV) in Mexico [1]. This device helps people with partial or full blindness to identify objects in their environment, read text and can give spoken directions with the help of GPS technology and Machine Learning. This device achieves a lot but its price is relatively expensive.

A slightly different device called the Finger Reader developed by MIT Media Labs [1] aimed to assist people with impaired vision read text. The device would be worn on a finger and when the user moves it over printed text. The device reads aloud the text that it recognizes and uses vibrations to let the user know when they strayed off the line in the text. This device is very useful but because it is worn on the finger, it is not convenient to use when the text is out of reach or if the text's font is too large.

A relatively new device called the Envision Glasses developed by Envision [2] which aims to aid people with impaired vision similar to the previous devices mentioned. It allows users to read text thanks to its text to speech functions, find objects, get descriptions of their surrounding and also use AI to scan text in documents to get summaries. This device is a competitive offering but it is relatively expensive.

The device that is going to be implemented in this project aims to assist people with full or partial blindness in a similar way to the other devices. The device will be used like glasses and will be used to identify the text around the user. When the user wants to know what the text around them states, they will use their smartphone to get an audio version of the text which they can listen to. The device will be relatively cheap to manufacture or implement since it will use readymade components to function. It will use a cheap microcontroller which is fitted with a camera to detect the text around and uses a Wi-Fi connection with a smartphone to relay the

text to be processed. The smartphone will then use Optical Character Recognition (OCR) to read the text from the images the camera of the microcontroller provides. After this, the smartphone will use Text-to-Speech to read aloud the text that has been recognized beforehand. This device will use this simple methodology to assist people with full or partial blindness. The advantages of this device will be the low cost. The device will use a low cost microcontroller and the user's smartphone which the user will already have and will not be an additional cost. Another advantage that this device will offer is that it will help people not only people with partial blindness but people with both partial and full blindness. An additional advantage will be the simplicity of repair and maintenance. Since the device is relatively simple, it will be simple to either find replacement parts or maintain the existing components. All these advantages would make this device suitable for the average person and will allow it to be used by many more people than the devices discussed above. Taking all these advantages into account, this project aims to design and implement this device.

Problem Statement

Traditional devices that help assist people with full or partial blindness do not allow their users to fully interact with their surroundings as they do not have the capacity to read text in a versatile manner. Modern devices designed to mitigate the issues of the traditional devices are usually too expensive for the average person to purchase, only help people with partial blindness and not people with full blindness, or are too complex that the manufacturing and maintenance of them is expensive and not simple. This project aims to design and implement a device that solves the previously mentioned issues by being relatively cheap, able to be used by people who have partial or full blindness and relatively simple to implement. The device will be relatively cheap because it will use readymade/off-the-shelf components which are easy to acquire and can be obtained for a low price. The device will allow users who have full blindness in contrast to other devices which only help people with partial blindness. The device will also use readymade components like stated previously and this will allow for the device to simple to implement/manufacture, maintain and purchase replace parts. The reasons stated above reflect the need for the device that will be designed and implemented in this project.

Objectives

General Objectives

- The general objective of the project is to design and implement a device that allows users with partial or full blindness to recognize the text in their surrounding by scanning the text and hearing an audio recording of it.

Specific Objectives

- **Design the Device**: Design the circuit of the device and specify what components will be used. The way that the different components work together will be decided and the design will be finalized.
- **Programming and Implementation**: Code that will control the device that was designed will be written. Then, the code will be uploaded to the components for the device to start working and start accomplishing its tasks.
- **Testing and Debugging**: The device will be used in different scenarios and conditions to test its efficiency and success/failure rate. After testing the device, changes will be made to improve the performance and efficiency of the device, if needed or possible.

Methodology

The methodology of this project will be divided into separate phases which will require their own amount of time and different resources to accomplish their objective. The methodology is divided into the following phases.

a) **Device Design**:

The circuit of the device will be designed and all the components required will be identified. After the device is designed, the list of all the components needed will be finished. The deliverables of this phase will be:

- System Circuit Design
- List of all required components

b) **Purchase of Components**:

After the list of all required components has been finished, the components will be purchased. The components will be waited for until they are delivered.

c) **Programming (Code Implementation)**:

In this phase, the necessary libraries required for the code needed for the project will be identified and chosen. Afterward, the code for that will run on the microcontroller (ESP32Cam) and communicate with the other devices (smartphone or computer) will be written. The code will be written in HTML and JavaScript. The code will utilize TesseractJS for OCR and use JavaScript's SpeechSynthesis API for converting text to speech/audio. The deliverables of this phase will be:

- List of libraries and programming language used.
- Code that will run on the device

d) **Implementing the Device**:

After the components are delivered and the code is finalized, the components will be connected based on the design in the first phase and then the code written in the previous phase will be uploaded to the components. The Arduino Uno will be used to upload the code from the computer to the ESP32Cam. The code will be hosted as a webpage by the ESP32Cam and this webpage will be accessed on a smartphone or computer's browser. Images from ESP32Cam's camera will be taken and then shown in the webpage. Then the OCR and text to speech conversion will be done within the browser. The deliverable of this phase will be a functioning device with the code uploaded to it.

Figure 1: Uploading code

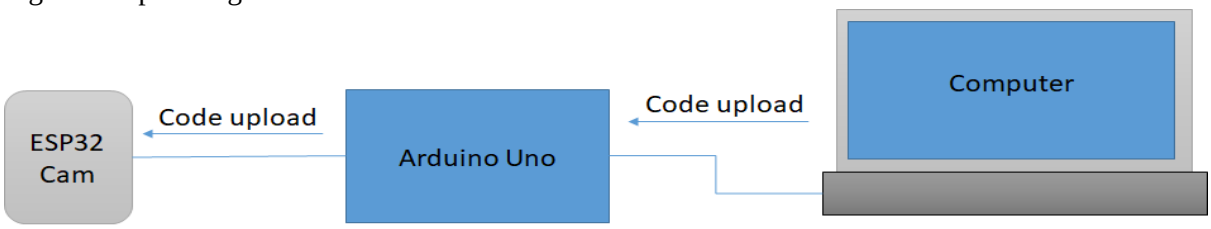
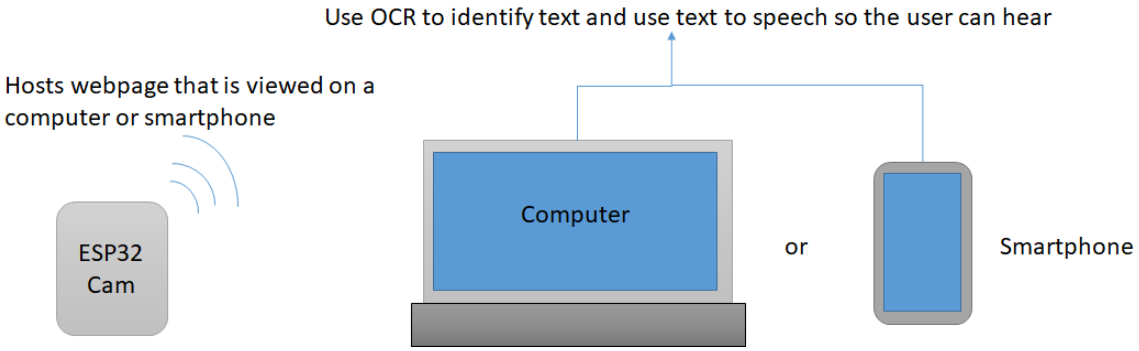


Figure 2: Device functionality



e) **Testing and Debugging**:

After all the components are connected and the code is uploaded to the device, the device will be used and tested. As the device is being tested, the success/failure rate of the device will be observed. Following the rigorous testing, the code will be debugged or refined to either solve problems, fix errors or improve the performance of the device.

f) **Documentation**:

After the device has been rigorously tested, and the code has been debugged and refined, the documentation phase will start. The final paper detailing the work done in the project will be prepared. The deliverable will be the final project document and will be submitted after finishing preparing it.

Resource Requirement and Estimated Budget

Successfully implementing the project will require a financial budget and time budget. The time budget is already shown in the next section below this but the financial budget estimate will be discussed here.

1. The microcontroller to be used for the OCR will be the ESP32Cam with Wi-Fi, Bluetooth and Camera module. It costs 1,590 ETB.

2. The microcontroller that will be used to connect the ESP32Cam board and the computer when the code is uploaded will the Arduino Uno. It costs 1,820 ETB.

3. Jumper wires will be needed for connecting the components. A pack of jumper wires costs 390 ETB.

Other resources will include the software used to write and implement the code. Those resources are listed as follows:

1. Visual Studio Code: This will be the code editor for writing code.
2. Arduino IDE: This will be the IDE used to write code and upload it to the microcontrollers.
3. HTML and JavaScript: The code will have HTML and JavaScript as the programming language.
4. TesseractJS: This will be the OCR engine that will assist with identifying the text in the project.

The following table shows the prices of the components and the total budget needed to purchase them.

Table 1: Cost Breakdown

No.	Name	Quantity	Unit Price	Subtotal Price
1	ESP32Cam with Wi-Fi, Bluetooth and Camera module	1	1,590 ETB	1,590 ETB
2	Arduino Uno	1	1,820 ETB	1,820 ETB
3	40P Conductor Male to Male Jumper Wire 20CM,40P Color Wires Ribbon Cable	1 pack (40 wires)	390 ETB	390 ETB
			Total Price:	3800 ETB

Project Timeline

The phases detailed above in the methodology section will be performed within the time frame of four months. The following Gantt chart shows the sequencing and duration of the activities.

Table 2: Work Plan Gantt Chart

Activity \ Weeks→	Month 1	Month 2	Month 3	Month 4
1. Device Design				
2. Purchase Components				
3. Programming				
4. Implementing the Device				
5. Testing and Debugging				
6. Documentation				

References

[1] Simon Hill, “5 amazing gadgets that are helping Blind people see”, June 20, 2014. [Online Serial]. Available: https://www.digitaltrends.com/mobile/blind-technologies/. [Accessed: January 22, 2024]

[2] Andy Klein, “Top 5 Assistive Technology Devices for People who are Blind or have Low Vision [2023]”, September 06, 2023. [Online Serial]. Available: https://www.letsenvision.com/blog/top-5-assistive-technology-devices-for-people-who-are-blind-or-have-low-vision-2023. [Accessed: January 22, 2023]