A picture containing text, clipart

Description automatically generatedA picture containing text, soup, dish, wheel

Description automatically generated

Cairo University

Faculty of Engineering

Credit Hours System

CCEN480 – GP1

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Visually Impaired Logo Safety Sign 1.2 mm rigid plastic 100x100 :  Amazon.de: DIY & Tools  SightGuide | GP 1 REPORT   |  |  | | --- | --- | | Ahmed Mohamed Ismail | 1180501 | | Mostafa Ashraf Ahmed | 1180406 | | Moaz Mohamed Mohamed | 1180528 | | Nader Youhanna Khalil | 1180477 | |

**Supervised by**

Assoc. Prof. Mona Farouk

**Submitted On**: 31/12/2022

Abstract

Even though many mobile devices today include accessibility features available for visually impaired and blind (VIB) users, many of these users are reluctant to use them. This is because the user interface is often designed for sighted people. This is caused by the fact that the main input and output methods in mobile devices are tactile or visual in nature. However, in recent years, there have been many innovative applications that assist VIB users in navigating their environment. Programmers have made use of technological advances regarding gyroscope sensors and vibration feedback to make communication possible. The proposed system relies on input images and videos provided by the user’s device camera to allow daily life navigation without the need to use such sensors. It makes communication between VIB users and their devices possible using speech/text conversion techniques.

Table of Contents

[1. Introduction 1](#_Toc123318876)

[1.1. Motivation and Justification 1](#_Toc123318877)

[1.2. Problem Statement 1](#_Toc123318878)

[1.3. Proposed Solution 1](#_Toc123318879)

[2. Market Feasibility Study 1](#_Toc123318880)

[2.1. Market Survey 1](#_Toc123318881)

[2.1.1. Survey 1](#_Toc123318882)

[2.1.2. Competitors 1](#_Toc123318883)

[2.2. Feasibility Study 1](#_Toc123318884)

[2.2.1. Market Analysis 1](#_Toc123318885)

[3. System Design and Architecture 2](#_Toc123318886)

[3.1. System Description 2](#_Toc123318887)

[3.2. System Block Diagram 3](#_Toc123318888)

[3.3. Modules 3](#_Toc123318889)

[3.3.1. Main Modules 3](#_Toc123318890)

[3.3.2. Secondary Modules 11](#_Toc123318891)

[4. Time Plan 11](#_Toc123318892)

[5. Task Division 11](#_Toc123318893)

[6. References 11](#_Toc123318894)

GRADUATION PROJECT REPORT

# Introduction

## Motivation and Justification

VIB users are often put at a disadvantage regarding their visually able peers. Technological advancements have always been concerned with providing better and easier to use solutions. These efforts have been largely directed toward the use of sensors, which can in many cases be unreliable.

Moreover, many of the applications that can be found in the market are not particularly easy to use. They often require some degree of tactile interaction, which VIB users will most probably not be able to provide. Some of these applications are designed to be used by sighted people alongside VIB users, which can come as impractical.

The before mentioned reasons led us to consider using AI and Machine Learning techniques to create a mobile application that can serve as an assistant to VIB people. We will be addressing these previous problems by rendering the contact between the application and the VIB user purely vocal as much as the desired features allow for it. In other words, the user will communicate with the chatbot through speech.

## Problem Statement

Our objective is to provide VIB users with a system that sufficiently fulfills their needs in daily life navigation. The system should provide all the functionalities needed for daily life tasks with ease of use.

## Proposed Solution

Our application SightGuide acts as a tool that can be used by VIB people in their daily life for navigation. Therefore, our goal is to facilitate daily life tasks which include identifying currency, clothes, people, products in markets, and other elements that may be in a scene and which prove essential for proper navigation. Our product also presents a text reader functionality that allows VIB users to perform the indispensable task of reading. The interface is proposed to convert speech to text and vice versa, to offer VIB users an application that can be used without the need for assistance.

# Market Feasibility Study

## Market Survey

### Survey

### Competitors

## Feasibility Study

### Market Analysis

# System Design and Architecture

## System Description

Our application targets visually impaired people, its focus is to help them in their everyday lives. It is a mobile app and since the users are practically blind the inputs to our system are voice and images captured from the mobile microphone and camera respectively. The proposed way to use the app is shown in Figure() which is a wearable mobility assistive system by using an Android smartphone and a sling pouch.



Figure : Proposed Wearable

The main interface provided to the user is a virtual assistant (chatbot) that takes commands from him via speech and then redirects the user to the corresponding module to do the required task which eventually responds to the user in the form of speech feedback.

## Diagram Description automatically generatedSystem Block Diagram

## Modules

### Main Modules

#### Virtual Assistant (Chatbot):

This module serves as an assistant that carries our everyday tasks via voice command. It recognizes the user’s voice and accomplishes the task requested from it.

Our application helps the user to search for things on the browser, set alarms, ask for the weather, and do any other everyday task. Also, it redirects the user to the other modules he requests, which are listed below.

Diagram

Description automatically generated

Figure :Mobile Voice Assistant Architecture

STT: Speech to Text, this is the process of converting speech signal into digital/text data.

TTS: Text to Speech, this is the process of converting digital/text data into speech.

Intelligent tagging and decision-making serve for interpreting the user’s request.

Image recognition: It is used for identifying places, and people within images.

Noise control: reduce and eliminate the background noise for voice clarity.

Voice biometrics and Security: Identity who is talking to it (customized assistant to your voice for security).

Speech compression: resize the voice data and send it to the information server in a succinct format.

Voice interface/feedback: the response that the user receives as feedback for his request.

However, there are ready-Made solutions for voice assistants like Siri, Google Assistant, Alexa, and Cortana so this module will not be implemented from scratch, it will be customized for our app purpose.

#### Scene Descriptor and Navigator:

Diagram

Description automatically generatedThis module describes the scene that the camera captures in from of the user so that he can imagine what the scene in front of him looks like. It also helps him avoid any obstacles that are less than 1.6m away so that he can travel safely in any place. It warns him that an object of a certain height is x meters away.

#### People and emotion recognizer:

Diagram

Description automatically generatedThis module is responsible for recognizing the people from the set of the user’s friends and family and notifying the user of their presence. Also, it recognizes their emotions from their facial expressions.

Face Detection: Viola Jones Algorithm

Face Recognition: DeepFace, DeepID, series of systems, VGGFace, FaceNet, FischerFaces/EigenFaces

Emotion Detection: Histogram segmentation, Feature extraction, fuzzy classifier-based emotion detection

#### Text Reader:

This module provides audio guidance to capture a printed page and recognizes the text along with its original formatting then speaks the document. We will be using CNN to implement this module.

Diagram

Description automatically generated

#### Product Identifier:

The product identifier module is very useful for visually impaired people. This module takes an image containing a product and its barcode as an input and outputs voice feedback telling the user what product he just captured. For the user to find the barcode, they will have to keep rotating the object slowly in front of the camera. When the module sees a barcode, it will start beeping. This is very helpful for shopping.

Diagram

Description automatically generated

#### Currency Detector:

A screenshot of a phone

Description automatically generated with medium confidenceThis module will help visually impaired people recognize currencies. It will also help them count the currencies. This module has five steps, preprocessing techniques for removing noises and preparing the image for the next operations, segmentation, and ROI extraction processes in the second and third steps for extracting the foreground currency from the background, applying ORB Algorithm in step four, and finally matching the results with the dataset.

#### Clothes Descriptor:

This module will be implemented from scratch. Its input consists of an image or a sequence of images. The output consists of whether clothes are to be found in the scene, and the type, color and texture if yes.

Clothes detection and extraction: Morphological operations + Segmentation techniques + CNN

Texture Analysis: LBPH algorithm + GLCM Features + Law’s Texture Energy Features

Classifier: Final classifier (SVM)

Diagram

Description automatically generated

### Secondary Modules

* Note Taker
* Event Scheduler

These two modules will be part of the existing application on the user’s mobile device. The chatbot will redirect the user to the corresponding application, convert speech to text, and then set the event or register the requested note.

# Time Plan

# Task Division

# Conclusion

In conclusion, our application has the goal of helping VIB users in their daily life tasks by enabling them to perform basic functions that are particularly difficult for them. Our product aims to use images and video footage provided by the user’s device camera in identifying scene elements and people that are necessary to know and locate for proper navigation. The product also aims to be used completely by VIB users without the need for assistance from sighted people.

# References

<https://www.cleveroad.com/blog/how-to-create-virtual-assistant-apps-like-siri-and-google-assistant/#:~:text=How%20to%20make%20a%20virtual,like%20Wit.ai%20or%20Jasper>.