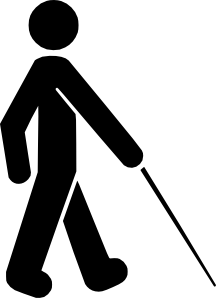
**Tell Me**

**(A Guidance system for a visually impaired)**



Abstract

**ABSTRACT**

One of the most common problems that many blind and visually impaired people face in their daily life is that they are totally unaware of their surroundings. They confront a number of visual challenges every day from reading the label on the frozen dinner to figuring out if they are at the right place. Along with this they are unable to identify things or people in front of them and what they are doing. In our society, these people are considerably neglected because of their disability factor and face great disadvantages while communicating. To overcome this problem and help out blind people I have designed this Android application which will facilitate them by making them more aware of their surroundings and improving their communication with the people around them. For this I have used Microsoft Cognitive Services APIs for face detection, gender and age, (Optical Character Recognition) OCR for text recognition and GPS for location.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  | Table of Content |
|  |  |  |  |  |

Table of Contents

[3 System Design 10](#_Toc470728248)

[3.1 Activity Diagram 10](#_Toc470728249)

[3.2 Sequence Diagram 11](#_Toc470728250)

**INTRODUCTION**

**1. Introduction**

The purpose of this document is to collect, analyze, and define the needs of blind people who are unable to access their surroundings and are at a distinct disadvantage because of that. The application I am going to design will help people of all age who have visual impairment of any kind.

**1.1 Overview**

The document discusses about the workings of the application and its impact upon the end users. The document also discusses how the application performs its tasks and the achieving of the end results.

I started developing this application keeping in mind the visually impaired as my end users. My application will not only help them in their everyday lives but also help them compete with their peers in a professional setting. I felt the need to develop this application because these people have to face far more obstacles in their everyday dealings than an average person.

The application that I have designed is an android application that uses the end user’s mobile camera application to capture the image. Once an image is captured, the application alerts the end user to the location. The application alerts about the location only once so long as the location doesn’t change. If there are any changes in the location, the application will intimate the user. Once an image is taken, the application first assesses it to identify whether there is any textual information or not. If there is text in the image, Text recognition API runs and textual information in the images will be read. The results achieved via Text Recognition API will be in JSON format. The application will make use of Microsoft Text to Speech services (function used is SpeakToAudio) to read out the information. This audio information will be a lot clearer if the end user is using head phones. If there are people in the captured image, then Face detection API will run and provide information such as the no. of faces in the image, their age group, gender, their moods and the activities they are involved in. If the image is mainly scenic then Image description API will describe the image.

**1.2 Purpose**

The purpose of this application is to facilitate people with visual impairment in their everyday lives and make it easier for them to carry out their activities independently. People with complete or partial visual impairment face challenges in their everyday lives. The purpose of my application is to minimize these obstacles and help them achieve their goals both in their domestic lives and their professional lives.

**2 Features:**

**2.1 Mobile Features:**

* NFC (Near Field Communication)
* Camera
* Speaker

**2.2 Technologies and Components:**

* Mobile Application Development
* Microsoft Cognitive Services
* Faces, Emotion and Computer vision API.
* OCR (Optical Character Recognition)
* Microsoft Text to Speech Recognition.
* VR Headset.



**2.3 Application Working:**

* The application will start by asking user to select from the two options;

1. Tap on the screen once to listen to the instructions of how the app will work.
2. Tap on the screen twice to skip the instructions and move on to the next step.

* The app will be running on Mobile device which will be placed in a VR classic headset which a blind person will wear on his/her head.
* The app will take the image on button (in VR classic) trigger. The mobile camera will be used for image capturing. It will then start processing the image and gather information such as face detection, age, gender, emotional state or the activity the subject is involved in.
* It will report back the gathered information to the end user (visually impaired person) in the form of voice generated messages that user can listen with headphones attached to VR headset. This entire information gathering is achieved by using Faces, Emotions, and Computer Vision APIs.
* In addition to all of this, the application will also be able to read textual information in a given image using Optical Character Recognition and extract the words it recognizes into a machine-readable character stream. The technology allows user to take images of text instead of copying the data thus saving time and effort.
* Another feature in this app will be to recognize the activity going on in the image.

e.g If the image taken is that of a man walking, the app will report back the activity to the end user.

* The application will also be able to alert the user about current location once as long as the location doesn’t change.

**2.4 Application Flow:**

* The application starts by running the function **initTextToSpeech()** which put forth two options.

1. To listen to the instructions tap on screen once.
2. To skip the instructions tap on screen twice.

Microsoft Text to Speech service is used to run the function SpeakToAudio() which will read out the instructions for first time users.

* Tap on screen twice to quit instructions and move to camera activity.
* Take picture on cardboard trigger by calling function **onCardboardTrigger().**
* The app will tell you to **wait** (2 sec) after image is captured.
* Convert the captured image to a bitmap.
* Compress bitmap image to send it in an API call as a parameter
* **Text Recognition:**

Call text recognition API **doRecognize(text\_bitmap)** to read textual information in an image using OCR.

Get the results in JSON format and speak out the extracted results to the user in the form of voice messages using Microsoft text to speech services.

* **Face Detection:**

Call API **doDetect(inputStream)** to detect number of faces in a captured image, their age , gender and mood.

Speak out the results in the form of voice messages using text to speech services.

* **Image Description:**

Call **doDescribe(bitmap)** API to describe any activity going on in the image.

Speak out the results in the form of voice messages using text to speech services.

**SYSTEM DESIGN**

## 3 System Design

System design is the process of defining the architecture, components, modules, interfaces, and [data](http://www.answers.com/topic/data) for a [system](http://www.answers.com/topic/system) to satisfy specified [requirements.](http://www.answers.com/topic/requirement)

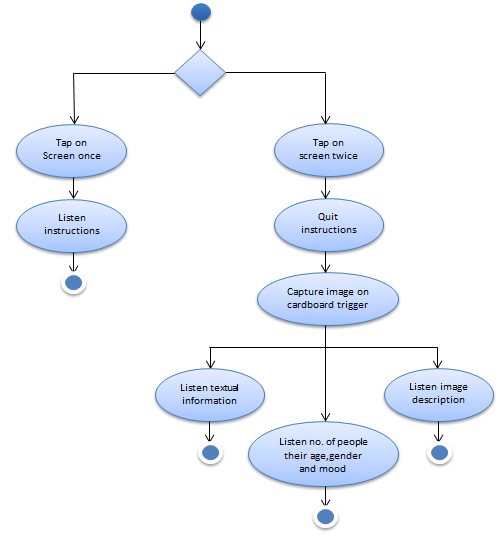
### 3.1 Activity Diagram

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modeling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. An activity diagram shows the overall flow of control.

Activity diagrams are constructed from a limited number of shapes, connected with arrows. The most important shape types

* + - Rounded rectangles represent activities.
    - Diamonds represent decisions.
    - Bars represent the start (split) or end (join) of concurrent activities.
    - A black circle represents the start (initial state) of the workflow.
    - An encircled black circle represents the end (final state).

Arrows run from the start towards the end and represent the order in which activities happen. Hence they can be regarded as a form of flowchart.

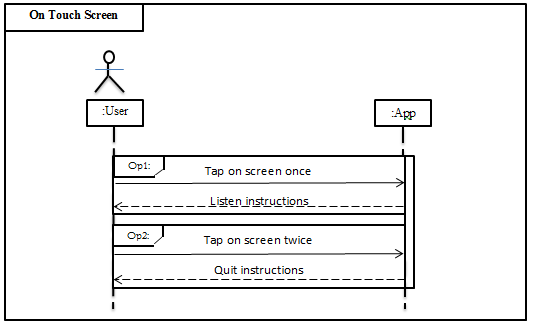


**Figure 3.1 Activity Diagram**

### Sequence Diagram

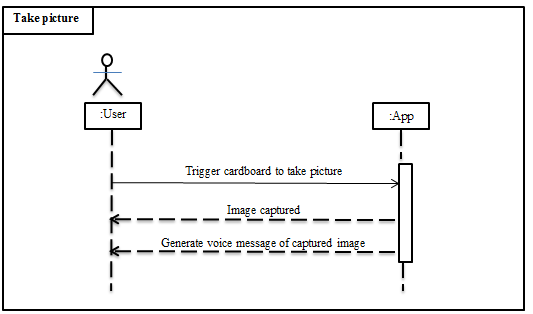
Sequence Diagram is one of the type of interaction diagram that are used to show the dynamic behavior of the system, that is how objects communicate with each other within a system. Here system is treated as white box. System dynamic design is made through Sequence diagram. In it, objects are assigned responsibilities by applying different design principles and patterns.

**3.2.1 On Touch Screen**



**Figure 2.2.1 SD: On Touch Screen**

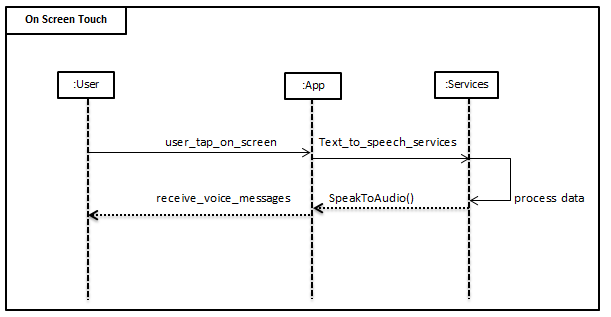
**3.2.2 Take Picture**



**Figure 2.2.2 SD: Take picture**

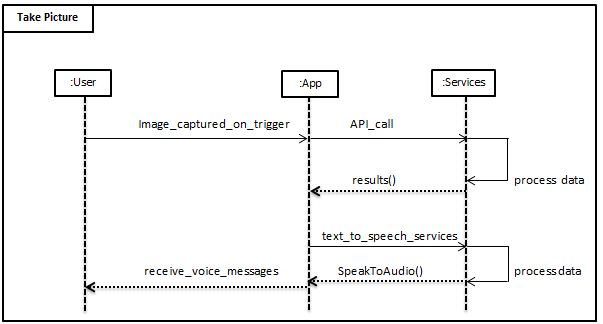
**3.3 System Sequence Diagram:**

**3.3.1 On Screen Touch**

****

**Figure 2.3.1 SSD: On Screen Touch**

**3.3.2 Take Picture**

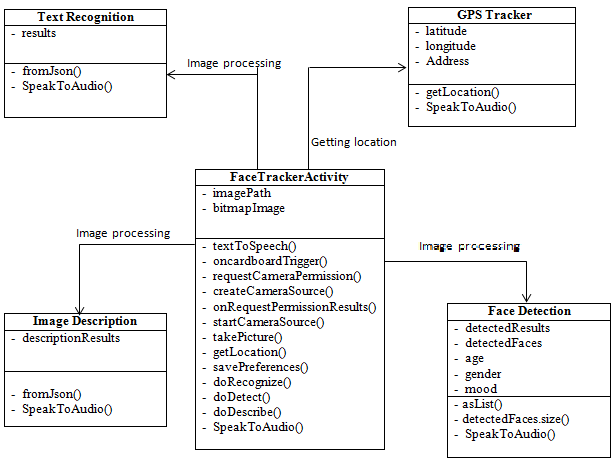
****

**Figure 2.3.2 SSD: Take picture**

**3.4 Class Diagram:**

A Class diagram shows the static structure of the system. It defines model elements such as classes, interfaces, and user-defined data types, their internal structure, and their relationships to each other.

* Relationships, or associations, are shown as lines connecting elements, and are annotated to describe the relationships and their cardinality (1...1, 1...\*, 0...\*, etc.).
* Inheritance (generalize/specialize), aggregation (comprises), and composition (has) relationships are also captured in this diagram.
* Class attributes and their data types are identified here, as are the operations and their return types.
* Visibility is indicated by +, #, or - for public, protected, or private. The class diagram plays a vital role in the transition from design to construction as it contains sufficient detail to begin the coding process. It is often used to partition responsibilities among the project team members, and to guide and measure the construction process.

****

**Figure 3.4 Class Diagram**