A Navigation Guide for Visually Impaired Person

Thesis submitted in partial fulfillment of the requirements for the award of degree of

Master of Technology

in

Computer Science and Application

by

Swati Singhal
Roll No: 601403028

Under the supervision of

Mr. Shatrughan Modi

Computer Science and Engineering Department
Thapar University
Patiala-147004

Certificate

I here by certify that the work which is being presented in the thesis entitled "*A Navigation Guide for Visually Impaired Person*" in partial fulfillment of the requirements for the award of degree of Master of Technology in *Computer Science and Application* submitted in Computer Science and Engineering Department of Thapar University, Patiala, is an authentic record of my own work carried out under the supervision of *Mr. Shatrughan Modi* and refers other researcher's work which are duly listed in reference section.

The matter presented in the thesis has not been submitted for award of any degree of this or any other University.

Swati

(Swati Singhal)

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.

Shatrughan Modi

(Mr. Shatrughan Modi)
Lecturer, CSE Department

Countersigned By

(Dr. Maninder Singh)
Head
Computer Science and Engineering Department
Thapar University
Patiala

(Dr. S.S. ~~Bhatia~~)
Dean(Academic Affairs)
Thapar University
Patiala

Acknowledgement

First of all I would like to thank the Almighty, who has always guided me to work on the right path of the life.

This work would not have been possible without the encouragement and able guidance of my supervisor Mr. Shatrughan Modi. I thank my supervisor for their time, discussions and valuable comments.

I am equally grateful to Dr. Jhilik Bhattacharya, Assistant Professor, Computer Science and Engineering Department, a nice lady and the sweetest Ma'am, a well-credited researcher in image processing. She always helped me and encouraged me whenever I faced a failure in research.

I am equally grateful to Dr. Maninder Singh, Associate Professor and Head, Computer Science and Engineering Department, a nice person, a well-credited researcher, who always encouraged me to keep going with work and always advised me with his invaluable suggestions.

I will be failing in my duty if I do not express my gratitude to Dr. S.S. Bhatia, Senior Professor and Dean of Academic Affairs, Thapar University, for making provisions of infrastructure such as library facilities, computer labs equipped with net facilities, immensely useful for the learners to equip themselves with the latest in the field. I am also thankful to the entire faculty and staff members of Computer Science and Engineering Department for their direct-indirect help, cooperation, love and affection, which made my stay at Thapar University memorable.

Last but not least, I would like to thank my family whom I dearly miss and without whose blessings none of this would have been possible. I would also like to thank my close friends for their constant support.

Date: 30-June-2016
Place: Thapar University, Patiala

Swati
(Swati Singhal)

Abstract

Visually impaired people face problems in navigation in an unknown environment. There are many portable devices and technologies to enhance the blind persons movement. Technology has already replaced the traditional methods like simple guiding canes with GPS based voice help, intelligent canes with inbuilt detectors and QR codes to name a few. This advancement in technology helps visually impaired people to achieve independence. This research developed an android based application for assisting blind people in navigation in any indoor or outdoor environment using a pre-generated map, GPS for self localization and vision based algorithms. The work makes use of Google voice recognition to acquire user voice input for the navigator system. Visual data acquired using the mobile camera is processed and analyzed for destination environment survey. To implement this functionality, application creates a SVM (Support Vector Machine) and train data set onto four categories like “Close Door”, “Open Door”, “Person Present”and “Others”. To implement support vector machine, the application uses OpenCV android SDK. Experimental results further demonstrate the working of application. The application allows to identify destination's environment with an accuracy of 82.5 %.

Keywords: Visually impaired people, Navigation, GPS, Android application, OpenCV android SDK

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

[4] According to World Health Organization [5] there are approx 285 million people around the world who comes under the category of visually impaired and the number is increasing every year. Out of this estimation 39 million people are blind and the number of people suffering from low vision is approx 246 million. This is the problem which does not depend upon the age and approx 82% of visually impaired people are older than 50 years. There is an estimation according to which someone goes blind in every five seconds and in every minute a child goes blind. The main cause of blindness is refractive errors like cataract which is the high cause of visual impairment in many countries especially in middle and low income countries and approx 90% of blind people live under low income category. Government is working to reduce blindness and running many campaigns due to which there has been a reduction of blindness caused by infectious diseases. These are some of the facts which are strong enough to show the need of technical systems that help the blind and visually impaired people.

1.1 Need of Technical System for Navigation for Visually Impaired People

Some visually blind people take help from organizations which appoints mobility trainer to guide them in a building but number of blind people are increasing so this idea seems no longer feasible. There is also other option that is use of cane or trained dog inside or outside a building. There are indoor localization system named as "conventional" that can guide a user to a destination by using some visual signs of regions inside a building. But guidance and detection of obstacles for blind people need the use of non visual type of communication like audio or any kind of vibration. Any system that targets the blind people should fulfill their needs like system must be user friendly, works on non visual communication, accurate etc.

In this thesis we are proposing an Android Application for the Navigation of

Visually Impaired People. Section 1.2 gives a brief overview of Android OS and its architecture. Section 1.3 explains how GPS is used for localization. Section 1.4 gives an introduction of Google Voice Recognition which is used for voice input from user. Section 1.5 introduces OpenCV which is used for obstacle detection and end scene scrutiny using Canny Edge Detector Algorithm and Support Vector Machine.

1.2 Introduction to Android OS

An operating system created by Google for mobile devices known as Android. The Android have its own SDK (Software Development Kit) which develops applications for its platform. There are many development tools and libraries included in Android SDK. The SDK works on JAVA and it requires a virtual machine, to run Android applications, known as Dalvik. Google provides Android Studio that can be used to create Android applications and alternatively Eclipse can also be used but while working with Eclipse it requires to add some ADT (Android Development Tools) plug-in. The compiled code is bundled with resource files to android package and having *.apk extension.

1.2.1 Architecture of Android Platform

There are four layers in Android platform as shown below in figure 1.1. Lets discuss these layers one by one briefly.

- **Application Layer:** There are many built in applications in Android OS like calendar, contacts, email client and many more. All applications are created using JAVA as its programming language. This layer is used to run the developed applications in android.

- **Application Framework:** The applications can be developed in the framework of Android. The components can be reused in the framework and reusing of components of existing applications in new application is also allowed, this reusing makes work easier for Android Developer. Some systems and services are accessible to every application such as:

  - **Views:** Like text boxes, Google maps, lists, grids, buttons all are example of views and used to build the application.

  - **Content Provider:** Content Provider can be used to access and share the data among different applications.

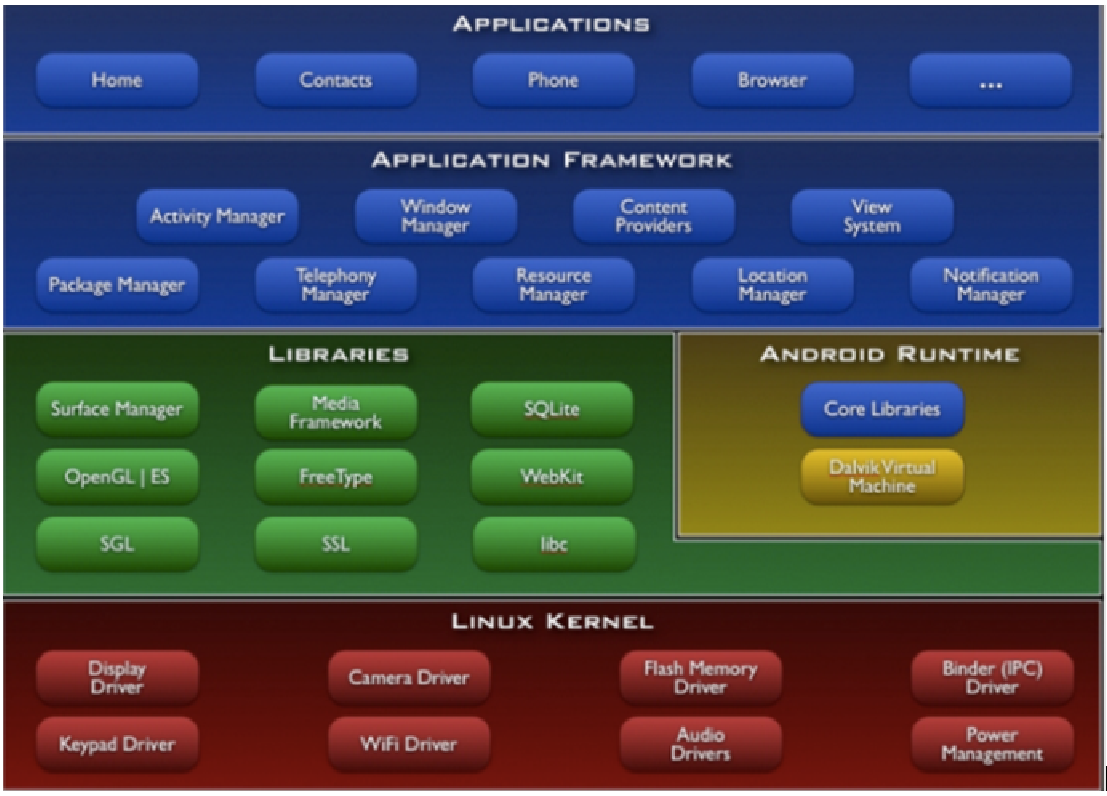

Figure 1.1: Architecture of Android [1]

  - **Resource Manager:** Some resources such as graphics and layout can be accessed by resource manager.
  - **Notification Manager:** Provides current status and alert messages.
  - **Activity Manager:** The applications whole life cycle is managed by activity manager.

- **Libraries:** The libraries of Android are developed in C/C++. there are many libraries such as the System C library which is used to import applications on Linux, the Media library which is used for the support of audio/video and image, the Free Type library used for the support of image rendering, the LibWebCore library is a web browser engine, the SQLite library is relational databased engine and many more.

- **Android Runtime:** This layer consists of two components: the libraries and the virtual machine called as Dalvik. These two components are available for any android application at run time.

- **Linux Kernel:** This layer is used to provide many system services to android. Many services like memory management, network stack, process management, security etc. are provided by this layer to android. It also works as a link between hardware and software stack of android.

1.2.2 Android Application Structure

As each application should be independently run in virtual machine thats because android provides independent virtual machine to every application. Each application is provided permission to use required components for a certain job. Android system works on this principle which makes android a very secure system. All the needed permissions should be mentioned at install time. There are four components of android application. Each component possess its own life cycle and has different use to build the android application. The components are defined as below.

- **Activity:** There are many activities in an application out of which one is main activity. This component works as a user interface as it allows user to interact with application. There can be an inter-activity communication and Activity class is the super class of all the activities of an application.
- **Service:** It works in the background. For example user wants to play music in the background then application calls service component. The Service class is the super class of all the services of an application.
- **Content Provider:** This component allow sharing data in different activities of an application. It also allows storing data and retrieving data. User is also allowed to create content and share data. The Content Provider class is the super class of all the content providers defined in an application.
- **Broadcast Receiver:** If any application sends any broadcast announcement then it is handled by this component. It does not directly interacts with user. The Broadcast Receiver class is the super class of all the broadcast defined in the application.

1.3 GPS

1.3.1 What is GPS

US placed the satellite based navigation system in the orbit that was formed from 24 satellites network and constitutes GPS system popularly known as Global

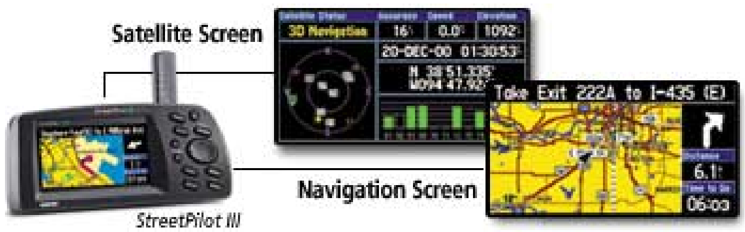

Figure 1.2: Working of GPS [2]

Positioning System. Initially it was started for military purpose and then in 1980 it was popularly used by civilians. It is free and available as inbuilt feature in every smart phone and can work in any weather conditions.

1.3.2 How GPS works

In a day, GPS satellites around the Earth twice in a particular orbit. Information signals are continuously transmitted by satellites to Earth and these signals are received by GPS receivers. GPS receiver uses trilateration technique to estimate the exact location of user. Basically the difference between the time signal was transmitted by satellite and received by receiver is calculated so that it can be estimated that how far the satellite is. The receiver follows the same procedure with more satellites and then calculates the exact position of user.

A GPS receiver can find the 2-D position (latitude and longitude) and movement of track of user by receiving signals from at least 3 satellites and it finds the user's 3-D position (latitude, longitude, altitude) by receiving signals from atleast 4 satellites. After calculating the position of user the GPS unit finds other required information like speed, bearing, track, trip distance, distance to destination, sunrise and sunset time and more as shown in figure 1.2.

1.3.3 How does the GPS works on Android device

All Android devices possess a GPS chip. Generally GPS systems receive signals and they usually do not send data to satellites. Because device requires a very large antenna to send a signal. But there are some systems such as OnStar which sends device's position back to the servers. An application of it is Google servers which receives device's position information which is helpful in calculating directions.

GPS Almanac technique calculates the expected position of satellites in approx 7 days in future, that's how it provides future data so that calculating the position of device can be speed-up. Gps satellites broadcasts the information of Almanac and device can download all the required information in only 15 minutes. There are some servers which provides all the information of Almanac in an efficient way. Android devices also uses those servers to get its information.

Some Android smartphone consists of AGPS chips (Assisted GPS). Devices consists of AGPS calculates position more faster and accurate information than only GPS enabled devices. There are three ways in Android device to calculate the position of device.

- With the help of WI-FI.
- With the help of network provider.
- With the help of GPS chip in Android Device.

Out of above mentioned three ways GPS is the most accurate one which makes the GPS used in navigation systems because their accuracy is very important. All of these methods are available in LocationManager of Android.

1.3.4 Android API

There is an API in Android which can directly access the information from device's GPS. The current state of GPS satellites can be obtained by this class. There is also one more class that is LocationManager which gives periodic updates of the device's geographical location.It takes some user's permissions also.

Public final class GpsSatellite extends object

<uses-permission android:name="android.permission.INTERNET" />

1.4 Google Voice Recognition

It is a product of Google which recognizes the speech input of user. It is a multilingual product and its API is available for Android [6]. It streams the real time data and works on the concept of see, hear and translate. It is based on the technique of neural network and it also works in noisy environments. Google provides this as a free API. The received voice is translated into text by this API. This service receives direct voice from user via microphone and sends all the data and language parameter online and it gives the text as output.

1.5 OpenCV

To sense the world visually, eyes and brain are both used by humans. Computer Vision is a technology which focuses to give same sense and capability to machine.

Extraction, analysis and understanding of useful information from images is the main aim of Computer Vision. There are various application areas of computer vision like robotics, augmented reality, character recognition, biometrics, remote sensing, security and surveillance and many more. The computer vision hierarchy contains three levels, for our work we only needs low level and middle level vision.:

- **Low Level Vision:** Feature extraction like edge, corner etc.
- **Middle Level Vision:** It works on the features of low level vision and recognizes object, analysis motion, constructs 3D images etc.
- **High Level Vision:** Interprets the information provided by Middle Level Vision and also directs the tasks performed by other two levels. It interprets such kind of information like activity, behavior, intention of a scene.

Open Source Computer Vision Library also known as OpenCV is machine learning and computer vision library. The goal of OpenCV is to give an infrastructure for applications of Computer Vision and to increase the use of machine perceptron at high level. It is licensed by BSD. There are approx 2500 optimized algorithms in OpenCV library and contains all the state of art algorithms of Computer Vision. The applications of OpenCV algorithms lies in many fields like face detection, object detection, and identify movement in an image, removal of red eyes, create 3D point cloud and many more. There is a very large community of OpenCV of approx more than 47 thousand and the estimated download count of OpenCV has already cross the 7 million counts. Now OpenCV used as commercial puprpose, research purpose and in government sector.

The interface of OpenCV is available in many languages like C, C++, Java, MATLAB, Python and it supports all operating system Windows, Android, Linux and Mac OS. It takes all the advantages of SSE and MMX instruction because it works for vision of real time applications. It is written in C++.

To achieve the object detection problem the proposed system extracts Canny Edge Feature of data set and uses Support Vector Machine to train and classify data. In this section a brief introduction of Canny Edge Detector and Support Vector Machine is described.

1.5.1 Overview of Canny Edge Detector

John F. Canny was the founder of Canny Edge Detector in 1986. It is also popular as optimal detector. The aim of canny edge detector is low error rate, good localization (minimum distance between detected edge pixel and real edge pixel) and minimal response (minimized to one edge receive only one detector response).

- The Gaussian filter is used to filter any noise. Size 5 Gaussian kernels example is shown below:

$$K = \frac{1}{59}\begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix}$$

- Procedure is followed to find the images intensity gradient.
  - Pair of convolution masks is applied in both directions x and y.

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

  - After that calculate the gradient strength and direction. Round the direction to any one of the possible angle from 0, 45, 90, 135.

$$G = \sqrt{G_x^2 + G_y^2}$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right)$$

- Apply non- maximum suppression so that those pixels which should not be the part of the image can be removed. And output comes only in thin lines which are also called as candidate lines.

- The last step is Hysteresis. Two thresholds as upper and lower are used by canny edge detector.
  - If pixel gradient > upper threshold then accept pixel as an edge.
  - If pixel gradient value < lower threshold, then reject the pixel.

Figure 1.3: Output of Canny Edge Detector

– If lower threshold < pixel gradient < upper threshold then its acceptation is based its connection with upper threshold. The recommended thresholds are upper: lower :: 2:1 or 3:1.

The output of canny edge detector is shown in figure 1.3.

1.5.2 Overview of Support Vector Machine

A Support Vector Machine (SVM) is also known as separating hyper plane. It works on the concept of supervised learning. The input of algorithm is labeled as training data and its output is an optimal hyper plane.

For example if the problem is given the set of 2D points then its output should be to find the separating line which divides the 2D points into two classes. The problem is described in the figure 1.4. In this figure there are many lines which can be the solution of the problem but the aim is to find the optimal linear line. The output should be the line which is not too close to the points because it may contain some noise, but output should be the line which correctly classifies all the points and also as far as possible from all the points.

The SVM algorithm finds the hyper plane which gives the highest minimum distance from all the training data set. This distance is also termed as margin. So the main aim of SVM is to find the optimal hyper plane which separates the

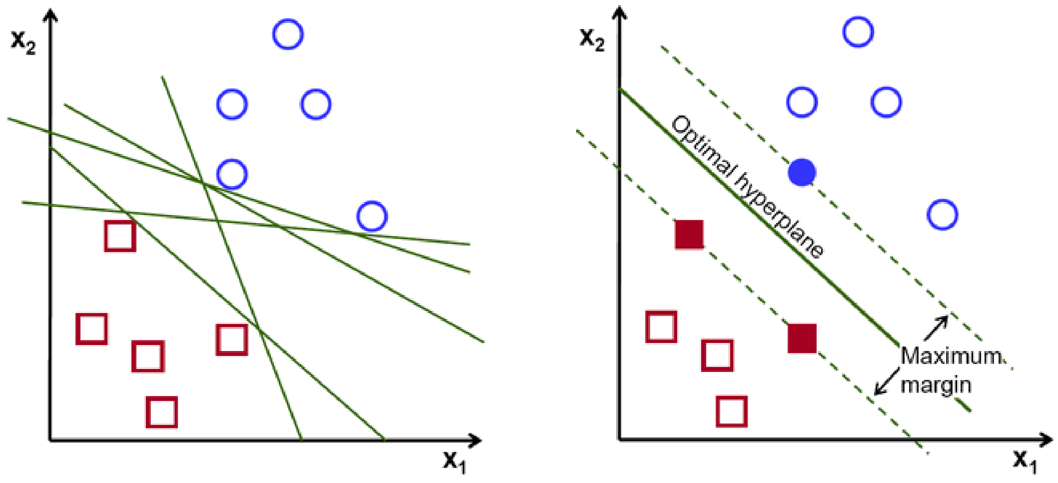

Figure 1.4: Output of Support Vector Machine [3]

margin and also margin should be as highest as possible. Its output is shown below in figure 1.4.

1.6 Structure of Thesis

The chapter 2 provides the detailed literature survey. Chapter 3 describes the problem statement followed by chapter 4 which describes the complete proposed method. Chapter 5 describes the experimental results and chapter 6 tells the conclusion and future work of the proposed system.

Chapter 2

Literature Review

There are many techniques developed to provide the both indoor and outdoor navigation for visually impaired people. While a large area has already covered by these technologies , the devices are not feasible for every user. Some devices are complex and heavy, some are environment dependent, some are robot assisted. Different Navigational techniques are described in this section.

2.1 Marker based Algorithm

In [7], [8] QR codes are placed on the floor area which makes the whole scenario working on pre-defined paths. The system needs android smart phone with QR scanner and web cam so that QR codes can be read through smart phone. This whole scenario contains six modules. First module is QR scanner through which application reads QR codes. After that according to the information present in QR code an instruction is processes from SQLite database and that instruction contains all the required information of device current position. QR codes can also be used as localization information as mentioned in [9]. ZXING library has been used for code scanning because it is helpful in scanning large quantity of QR codes.

Second module displays the information of all the nodes on the output screen of user and also pass the information in the form audio/speech instruction. Text to Speech module which is a third module gives an easy interface to visually impaired people. This module gives all the information of current location to user and also asks the destination from user in the form of audio. After collecting or finding the information of current location and destination the application passes the values to next module which is Path Generation Module.

User is also allowed to choose between shortest path or the optimal path in this fourth module i.e. Path Generation Module. For calculating optimal path in [7] and [10] the application uses A* algorithm because it is best to calculate number of turns for heuristic input. According to Current location and destination location shortest path can be find out but optimal path depends upon the turns between current and destination location.

After generating path the next module allows to store it into Trip Manager. This module allows user to have successful navigation as it already have generated path.

The next module, named as Deviation Detection Module, also helps user to have successful navigation because it detects deviation of user from the right path. If user is not visually impaired than only Trip Manger is sufficient to provide all the information but in case of visually impaired there are chances that user can deviate from the right path. Deviation is detected by comparing current location of user and current location according to generated path.

2.2 Obstacle detection by using OpenCV library

In [11], [12] and [13] the application uses OpenCV for detection of obstacles and this technique basically is divided into three modules.

First module is marker. QR codes are used to provide input for android smart phone. Initially very simple marker were used which are just a simple letter (L or R) inside a circle but they gave a lot of trouble in detection by using a function of OpenCV i.e. Hough Circle Transform. After failing the idea of simple circle, marker barcodes were used by application so that smart phone can take picture of barcode and decode the information present in it. But due to presence of many parallel vertical lines they were incapable of collecting and providing the information of stairs and obstacles. After that location fingerprinting were tried to take input by application. This technique finds difference between received signal strength (RSS), that was received by some defined wireless access point, and values of same locations which were already recorded in database. After calculating the difference it was checked that which position was most closely matched. This idea gave very accurate and efficient results but because of its time consuming nature it was not accepted. After that QR codes were used because they are comparatively cheap and can store large amount of information in it. As QR codes are not orientation dependent they can easily overcome the disadvantages of barcodes. QR code technique can also be combined with RFID for better results.

Second module is obstacle detection. For obstacle detection the OpenCV library has been used. There are many ways to detect obstacles. One way is to use Canny Edge Detection and Hough line. The idea is to apply Hough Line algorithm on the output of Canny Edge Detection. Both of these methods are present in ImageProc Package. To detect obstacles the number of lines of output image is compared with a threshold value and if it is greater than that threshold value than obstacle is present. The second method to detect obstacles is homography estimation between two images so that some feature points can be matched between two images. The third method for obstacle detection is optical flow. This method extracts information of optical flow from image and also detects collision at real time. Video package of OpenCV consists the required functions. One of the two functions can be used for this algorithm i.e. calcOpticalFlowSF or calcOpticalFlowPyrLK.

The third module is Dead Recockning. This module uses the sensors of smart phone for dead recockning. There are geo-magnetic sensors and accelerometer sensors which calculates orientation angles via rotation matrix. The getRotationMatrixFromVector() and getOrientation Android API used for this purpose. The another way or technique used in dead recockning is to calculate footsteps of user and it is known as Pedometer. It is based upon the searching of peaks and trough mapped on waveform of time-domain. These two peak and trough are searched alternatively, when the first one is detected other one is searched and vice versa. After searching peak and trough their difference has been calculated for dead recocking.

2.3 Robot assisted navigational technique

In this technique [14], [15] and [16] the guiding dogs are replaced by Mindstorm Lego NXT robots. NXT is the mind of robot and is programmed intelligently. Android provides easy interface to NXT, because android and NXT are both written in java, so it is programmed for android devices. NXT controller is the main part of Mindstorm Lego and it is also known as the mind of the device. The distance, speed and rotation of movement is controlled by servo motors. There are many types of sensors like sound, touch, light, ultrasonic sensors [17] are present in NXT robots but the application uses only ultrasonic sensor. The vision of robot which allows to detect and sense obstacles is given by ultrasonic sensors.

The firmware LeJOS is used for NXT. LeJOS is very powerful firmware as it replaced the LabView firmware which is the original firmware of NXT robots. LeJOS has JVM, so it allows java programming. Eclipse LeJOS plug-in is also required because programming is done with the help of Eclipse IDE. The projects run on mobile device via Bluetooth or USB. The current location of user in the terms of coordinates is passed from android application to NXT robots via NXT projects.

After getting the location from android application NXT saves those coordinates and finds the distance and direction of movement in terms of angle between current and next location.

The process of navigation using NXT is divided into four steps. First step is the calculation of distance between two locations which is calculated by Haversine Formula.The Haversine formula calculates distance between two pairs of longitudes and latitudes. Second step is the calculation of angle between two locations. After getting the angle it is checked that whether the value of angle is greater than 180 or less than 180. The movement will be in right if it is less than 180 and it will be in left if it is greater than 180. Basically the greater than 180 means (360 - given angle), so it will move left if angle is negative. All the information related to movement is displayed on the screen of NXT. Third step is movement of motor of NXT robot. It uses two motors for movement and rotation. There is a class i.e. DifferentialPilot which controls the movement of motors and also controls the speed of motors. This class basically controls the two functions of movement of NXT robot first one is to rotate in place and second one is to movement in straight line. This function finds the distance between the locations of user, the current one and the next one and the rotate function finds the angle. Fourth step is the object detection. The object which comes in the area of less than 50 cm can be easily detected by NXT robots. The ultrasonic sensor has been used to detect obstacles by object detector listener. This listener contains a Feature Detector which helps to fulfill the aim of object detection. Basically this feature scans the real time scenario automatically in period of 500 ms and collects the data related to it.

The connection held between NXT robot and android is usually done by Bluetooth where android acts as sender and NXT acts as receiver. For sending data between this Bluetooth connection the DataOutputStream and DataInputStream classes are used in java.

2.4 Navigation by light or sound

Navigation by sound or Spatial audio navigation means to walk towards the sound. This proposed system describes well for smart phone application. In [18], the feasibility of this navigation has been identified on Android platform. Two users starts with opposite direction in an environment and both users navigate towards in the direction of each others and it is done by the dynamic mapping of the audio. The author in [18] describes its achieved precision and accuracy in a very well results. The system based on audio or visible light may need to wear some wearable devices [19] or may walk in virtual environment [20]. In [21] and [22] the author proposed an indoor navigation system which uses a technology of visible light communication. The developer puts LED light at navigational way which supports the navigation of blind man.

2.5 Navigation by using NFC (Near Field Communication) tags

The proposed application in [23] is named as PERCEPT and it was proposed for both indoor and outdoor navigation. The landmark tags of the present environment is saved in near field communication tags [24] and user is given as audio information related to environment after tapping on any landmark tag. The smart phone access the near field communication tags with the help of its sensors [25]. The structure of node and link has been generated by the application which presents the landmarks as nodes and the links between the landmarks as edges as mentioned in [26]. With the help of Near field communication tags the current node can be generated and also the weight between two links.

2.6 Feature Extraction for obstacle detection

The proposed application, in [27], extracts feature of given image. The feature it extracts as Canny edge [28] , texture of image, color of image and after combining [29], [30] all these features it checks the best match from the database. The database contains the extracted features of various images. To extract texture feature it uses Gabor Algorithm [31]. The useful information can be easily extracts by the texture of images. the method [32] to extracts Canny Map is applied on images which gives a binary image as output and the color statistics can give useful information if used as in two different spaces. The proposed method also termed as Content Based Image Retrieval based method. The method provided by the author is very time efficient and very accurate.

2.7 Magnetic maps of working environment

In this proposed system [33] a tool named as Indoor Atlas used to generate the automatic magnetic maps of working region. At each location it finds the unique magnetic location and then use that value to calculate the position in map. It works with high accuracy and it does not require any extra hardware equipment. After find all the location, it stores the value of positions in its database. The input should be given in the form of graphical map to the Indoor Atlas. After that the Indoor Atlas itself generate the unique magnetic points of the locations and these positions should be confirmed by walking on that route.

Chapter 3

Problem Statement

In this particular chapter, the gaps which are there in the current work, the problem statement, the objectives which are to be achieved and method for achieving these objectives are discussed.

3.1 Gap Analysis

In the Literature Review, we have discussed the different techniques of Navigation for visually impaired people. In the existing work following gaps are there:

- The Marker Based Algorithms are valid only for pre defined paths. The paths should be covered with some markers, so there is a need of method which does not depends on QR codes.
- The Robot Assisted Navigational Techniques are hardware dependent. These technique suffer cost expensiveness, such as carrying expensive robots or intelligent sticks, so a method is required which does not require carrying heavy hardware device.
- The results of the navigational techniques which are dependent on light or sound are not quite good, an improvement can be done by implementing method which analyze the environment and provide correct output to the user.
- The technologies based on Near Field Communication (NFC) tags are easy to use but works only in a short range and these technologies does not provide security to hardware device, and there is a scope of improvement in method which can work in a large range and can provide security to device.

3.2 Problem Statement

Most of the work reported in literature for navigation purposes either suffer cost expensiveness, such as carrying expensive robots or intelligent sticks. Also, many methods are based on marker based algorithms which requires calibrating the scenes and hence cannot be used in unstructured environments.

Our aim is to develop an android application which can be easily operated on a mobile phone without the need of any other hardware equipment and can fulfill all the requirements of visually impaired people. This application will use pre-generated map, GPS for self localization and vision based algorithms for end scene scrutiny.

3.3 Objectives

- To study different techniques of navigation for visually impaired people.
- To propose a method of navigation using GPS for localization and vision based algorithms for end scene scrutiny.
- To implement the technique using the proposed method.
- To test and validate the output of proposed method.

3.4 Methodology

- Literature Review.
- Propose the method of navigation by using Google Voice API for taking input, GPS for localization, Dijkstra's Algorithm for finding optimal path, Support Vector Machine to classify images of destination.
- Implementation of the desired method is as follows:
  - Use device's GPS system to locate the position of user.
  - Allow user to choose the destination in voice command.
  - Apply Dijkstra's algorithm on source point and destination point and finds the optimal path for user.
  - Give navigational instruction to user till destination arrived.
  - Apply Support Vector Machine Classifier to check the destination's door is open or closed, and if it is open then whether someone is sitting inside or not.
- Test and validate the results of application.

Chapter 4

Proposed Approach

The proposed application is an Android based application and it works for pre generated paths. A map of Thapar University was generated for a floor with number of faculty cabins as a prototype region of navigation. Proposed system requires an Android mobile phone having OS version of at least 4.1 (Jelly Bean) and integrated camera.

4.1 Flow of Application

The flow of application is shown in figure 4.1 and it divides into five modules.

4.1.1 Localization

The application uses device's GPS system to localize the user. First it finds the positional coordinates (Longitude and Latitude) of user and then matches those coordinates with the database and find out the position of user. Once it gets the position of user it allows user to choose the destination.

4.1.2 Choose the destination

After getting current position, user can give voice command to choose destination cabin. The user can tap anywhere on the screen and tell the destination in own voice and Google Voice Recognition identifies user's destination.

4.1.3 Finding Optimal Path

After getting current position and destination point of user the application finds nearest point from current point and calculates the optimal path by applying Dijkstra's algorithm on the prototype map containing the information of faculty cabins. Then it displays the optimal path on the user's screen.

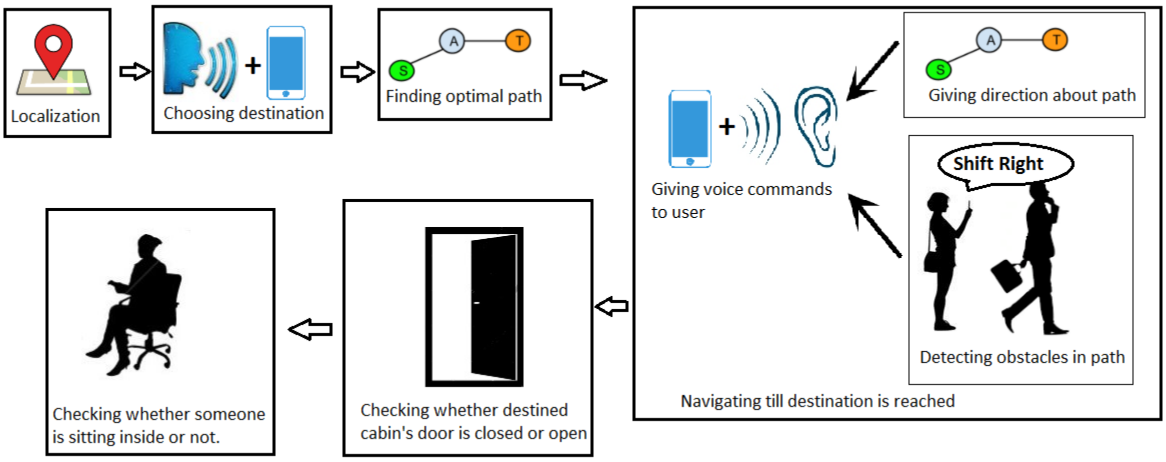

Figure 4.1: The Flow of Application

4.1.4 Navigation

After finding optimal path the application has to guide the user. So it first finds the distance and direction from current point to next point of optimal path by Haversine Formula and gives voice command to the user to move resulted distance in resulted direction. It continuously retrieves positional coordinates of user and checks whether next point has been reached or not.

4.1.5 End Scene Scrutiny

After arriving at destination, the application checks that the destined faculty room's gate is open or closed. After getting room is open it checks further that whether someone is sitting inside the office or not. To implement these functionalities, application creates a SVM (Support Vector Machine) and train data set onto canny edge feature of data set and after that classifies data set into four categories like "Close Door", "Open Door", "Person Present"and "Others". Then application extracts canny edge feature of current captured frame of destination and pass it into SVM (Support Vactor Machine) Classifier. To implement support vector machine the application uses OpenCV android SDK.

Chapter 5
Experiments and Results

Initially the graph as shown in fig. 5.1 was displaying on the user's screen, where "Door"(marked as red) represents the initial location of the user. Application checks user's positional coordinates and finds that user's initial location was at Door. Then user was asked to choose the destination by tapping anywhere on screen. After tapping on screen, Google Voice Recognition screen was popped up and user chose destination cabin as CITM by giving voice command. After choosing the destination user was asked to verify it. Then application founds the optimal path for user and starts giving voice instructions. Navigation instructions are being displayed on screen as well as user was also getting them as voice. While navigation user was getting the information of movement as followed.

- Go 54 m in straight.
- Now you are passing through ME Research Lab.
- Go 22 m in straight.
- Now you are passing through the "way".
- Move forward 10 m and then turn right.
- Turn Right.
- Go 76 m in straight.
- Now you have reached your destined cabin.

The above scenario has shown in fig 5.2 in which green color shows the current location of user. After user was reached at destination, it was checked that door is open or closed, and if it is open then whether some one was sitting inside or not. The application's database contains 2000 images of four sets. Each set contributes 500 images to database. Set 1 contains the images of different type and color of doors like wooden doors, glass doors from which wooden doors are easily recognizable by application. Set 2 contains images of various open doors. Set 3

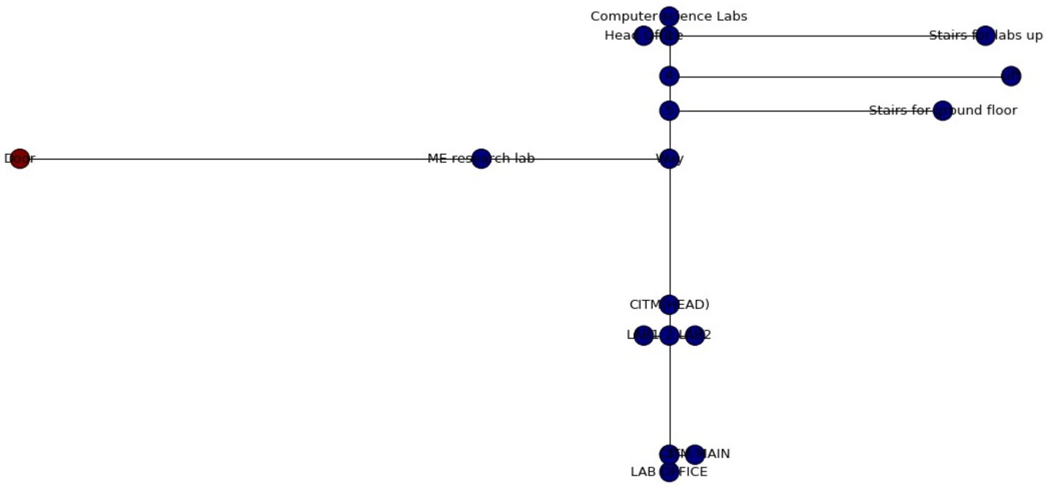

Figure 5.1: Graph of the prototype location.

Figure 5.2: Screenshots of user's screen.

Figure 5.3: Database of Images

contains images of someone sitting in cabin and set 4 contains images of open, empty cabin. The database contains images of different angles and orientation as shown in figure 5.3.

The application was trained for 1600 images, 400 images from each set, and was tested for 400 images, 100 images from each set. The confusion matrix and the results are as followed in table 5.1 and table 5.2 respectively.

Table 5.1: Confusion Matrix.

		Predicted			
		Close Door	Open Door	Person Present	Others
Actual	Close Door	87	5	0	8
	Open Door	10	78	0	12
	Person Present	0	3	84	13
	Others	7	12	0	81

Table 5.2: Recognition Results.

Sr.No.	Image Set	Images correctly recognized / Total no. of Images	Recognition Rate (in %)
1.	Close Door	87/100	87
2.	Open Door	78/100	78
3.	Person Present	84/100	84
4.	Others	81/100	81
	Total	330/400	82.5

Chapter 6

Conclusion and Future Work

6.1 Conclusion

An application is designed to provide navigation for visually impaired person. It is android based application and does not need any other hardware thats why it is very cheap and can be accessed by every visually impaired person. It allows visually impaired person to move in unknown environment independently without any guiding dog or white cane or any other hardware device. It allows to identify destination's environment with an accuracy of 82.5 %. This application has been tested on Samsung Galaxy Core on Jelly Bean OS.

6.2 Future Scope

It can be further extended as multiple language support can be added to application. Multi floor support can be added as right now this is working on single floor. Magnetic maps can be added in this application so that it does not limit to pre generated paths. Face recognition at destined cabins can also be added to this application. Database of images can be further extended as dynamic database so that it can learn from new scenarios. More features like finding optimal path in sense of crowded area can also be added to this application.

References

[1] E. Team, *Architecture of Android*. Available at http://elinux.org/Android_Architecture, Accessed on June, 2016.

[2] G. Team, *Working of GPS*. Available at http://www8.garmin.com/aboutGPS/, Accessed on June, 2016.

[3] O. D. Team, *Support Vector Machine*. Available at http://docs.opencv.org/2.4/doc/tutorials/ml/introduction_to_svm/introduction_to_svm.html/introductiontosvms, Accessed on June, 2016.

[4] Y. Freund, "Predicting a binary sequence almost as well as the optimal biased coin," *Information and Computation*, vol. 182, no. 2, pp. 73 – 94, 2003.

[5] WHO2016, *World Health Organization*. Available at http://www.who.int/mediacentre/factsheets/fs282/en/, Accessed on June, 2016.

[6] G. D. Team, *Google Voice Recognition*. Available at https://developers.google.com/voice-actions/interaction/, Accessed on June, 2016.

[7] A. Idrees, Z. Iqbal, and M. Ishfaq, "An efficient indoor navigation technique to find optimal route for blinds using qr codes," in *Industrial Electronics and Applications (ICIEA), 2015 IEEE 10th Conference on*, pp. 690–695, IEEE, 2015.

[8] S.-J. Lee, G. Tewolde, J. Lim, and J. Kwon, "Qr-code based localization for indoor mobile robot with validation using a 3d optical tracking instrument," in *2015 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, pp. 965–970, IEEE, 2015.

[9] H. Kobayashi, "A personal localization system using self-contained 2d barcode landmarks," in *2015 IEEE 24th International Symposium on Industrial Electronics (ISIE)*, pp. 882–887, IEEE, 2015.

[10] X. Cui and H. Shi, "A*-based pathfinding in modern computer games," *International Journal of Computer Science and Network Security*, vol. 11, no. 1, pp. 125–130, 2011.

[11] F. Li, C. Zhao, G. Ding, J. Gong, C. Liu, and F. Zhao, "A reliable and accurate indoor localization method using phone inertial sensors," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pp. 421–430, ACM, 2012.

[12] A. Mian, "Realtime face detection and tracking using a single pan, tilt, zoom camera," in *2008 23rd International Conference Image and Vision Computing New Zealand*, pp. 1–6, IEEE, 2008.

[13] M. Brady, "Vision-based detection of stair-cases stephen se department of computer science, university of british columbia vancouver, bc canada v6t 1z4 and,"

[14] A. Kulkarni, A. Wang, L. Urbina, A. Steinfeld, and B. Dias, "Robotic assistance in indoor navigation for people who are blind," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 461–462, IEEE, 2016.

[15] D. M. Alghasra and H. Y. Saeed, "Guiding visually impaired people with nxt robot through an android mobile application," *International Journal of Computing and Digital Systems*, vol. 2, no. 3, pp. 129–134, 2013.

[16] W. M. El-Medany, A. M. Abuaesh, and N. M. Abuaesh, "An efficient car like-mobile robot design," in *EUROCON 2009, EUROCON'09. IEEE*, pp. 996–1001, IEEE, 2009.

[17] P. Veelaert and W. Bogaerts, "Ultrasonic potential field sensor for obstacle avoidance (vol 15, pg 774, 1999)," 1999.

[18] J. Hinge, N. W. Jensen, and M. Poulsen, "Feasibility of spatial audio navigation using the android platform," *Itu. Dk*, 2012.

[19] J. Wilson, B. N. Walker, J. Lindsay, C. Cambias, and F. Dellaert, "Swan: System for wearable audio navigation," in *2007 11th IEEE international symposium on wearable computers*, pp. 91–98, IEEE, 2007.

[20] A. Sontacchi, M. Strauß, and R. Holdrich, "Audio interface for immersive 3d-audio desktop applications," in *Virtual Environments, Human-Computer Interfaces and Measurement Systems, 2003. VECIMS'03. 2003 IEEE International Symposium on*, pp. 179–182, IEEE, 2003.

[21] M. Nakajima and S. Haruyama, "New indoor navigation system for visually impaired people using visible light communication," *EURASIP Journal on Wireless Communications and Networking*, vol. 2013, no. 1, pp. 1–10, 2013.

[22] M. Nakajima and S. Haruyama, "New indoor navigation system for visually impaired people using visible light communication," *EURASIP Journal on Wireless Communications and Networking*, vol. 2013, no. 1, pp. 1–10, 2013.

[23] A. Ganz, J. M. Schafer, Y. Tao, C. Wilson, and M. Robertson, "Percept-ii: Smartphone based indoor navigation system for the blind," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3662–3665, IEEE, 2014.

[24] B. Ozdenizci, K. Ok, V. Coskun, and M. N. Aydin, "Development of an indoor navigation system using nfc technology," in *2011 Fourth International Conference on Information and Computing*, pp. 11–14, IEEE, 2011.

[25] L. A. Guerrero, F. Vasquez, and S. F. Ochoa, "An indoor navigation system for the visually impaired," *Sensors*, vol. 12, no. 6, pp. 8236–8258, 2012.

[26] A. M. Ali and M. J. Nordin, "Indoor navigation to support the blind person using weighted topological map," in *2009 International Conference on Electrical Engineering and Informatics*, vol. 1, pp. 68–72, IEEE, 2009.

[27] H. Salehian, F. Zamani, and M. Jamzad, "Fast content based color image retrieval system based on texture analysis of edge map," in *Advanced Materials Research*, vol. 341, pp. 168–172, Trans Tech Publ, 2012.

[28] F. Mahmoudi, J. Shanbehzadeh, A.-M. Eftekhari-Moghadam, and H. Soltanian-Zadeh, "Image retrieval based on shape similarity by edge orientation autocorrelogram," *Pattern recognition*, vol. 36, no. 8, pp. 1725–1736, 2003.

[29] D. A. Clausi and H. Deng, "Fusion of gabor filter and co-occurrence probability features for texture recognition," *IEEE Transactions on Image Processing*, vol. 14, no. 7, pp. 925–936, 2005.

[30] T. Ojala, M. Rautiainen, E. Matinmikko, and M. Aittola, "Semantic image retrieval with hsv correlograms," in *Proceedings of the Scandinavian conference on Image Analysis*, pp. 621–627, Citeseer, 2001.

[31] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using gabor filters," *Pattern recognition*, vol. 24, no. 12, pp. 1167–1186, 1991.

[32] R. Kaur, M. Verma, and H. K. Kalpna, "Classification of various edge detectors," *Department of Computer Science, RIEIT, Railmajra*, 2008.

[33] V. Prudtipongpun, W. Buakeaw, T. Rattanapongsen, and M. Sivaraksa, "Indoor navigation system for vision-impaired individual: An application on android devices," in *2015 11th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, pp. 633–638, IEEE, 2015.

List of Publication

1. Swati Singhal, Jhilik Bhattacharya and Shatrughan Modi, "*A navigation guide for visually impaired using GPS based localization and vision based obstacle detection and end-scene scrutiny*", Accepted in ICICT, August 2016.

Video Presentation

1. https://youtu.be/7NhgMnkZcFo