

Portable Monitoring and Navigation Control System for Helping Visually Impaired People

by

Mohit Sain

Thesis submitted

In partial fulfillment of the requirements

For the Master of Applied Science degree in

Mechanical Engineering



uOttawa

Department of Mechanical Engineering

Faculty of Engineering

University of Ottawa

© Mohit Sain, Ottawa, Canada, 2017

Abstract

Visual Aids for the blind people is an important subject. Apparently visually impaired individuals get impeded by certain hurdles in everyday life. This work proposes an indoor navigation system for visually impaired people. In particular, the goal of this study is to develop a robust, independent and portable aid to assist a user to navigate familiar as well as unfamiliar areas. The algorithm uses the data from Microsoft Xbox Kinect 360 which makes a 3D map of the indoor areas and detects the depth and estimates the relative distance and angle to an obstacle/human. To ensure the accuracy, Kinect tool is enabled with a colour camera to capture real-time details of surroundings which are then processed accordingly. Besides, the developed aid makes the user aware of environmental changes through a Bluetooth enabled headphones used as audio output device. The trials were conducted on six blindfolded volunteers who successfully navigated across various locations in the university campus such as classrooms, hallways, and stairs. Moreover, the user could also track a particular person through output generated from processed images. Hence, the work suggests a significant improvement for existing visual aids which may be very helpful in customisation as well as the adaptability of these devices.

Acknowledgements

This thesis has been a challenging and wonderful journey. I take this opportunity to firstly thank Faculty of Engineering, University of Ottawa for giving me this opportunity.

Such an endeavor would not have been possible without the much-appreciated help and support of every single person in my life has provided. I would like to express my gratitude to my supervisor Dr. Dan Neculescu whose continuous guidance, encouragement and support made it possible for me to finish this work and helped me improve different aspects of this thesis. I would also like to especially thank Dr. Natalie Baddour for her valuable comments and ideas for further improvement of this thesis work.

Also, I would like to express my thanks to Leo Denner who helped me with setting up the apparatus for my experimental work.

Further, I would like to thank my Father, his blessings made me reach here, he provided me all the strength, courage, guidance and the environment to grow which led my interest in the research field. Mom has been my backbone, my inspiration, she has provided me with all the strength. And further, I would thank my loving sister who has been there for me whenever I needed her.

I would also like to thank Amardev Khokhar, Amandeep Singh Hunjan, Armaan Sekhon, Christopher Lepine, Hartej Singh, Pranay Sharma and Karan Ghuman for all the support they have provided in my hard times and throughout my studies.

Also, I would like to express my thanks to my colleagues Arpit Ainchwar, Ali Reza Mirghasemi, Bilal Jarrah, Hamid Reza Fallah, Jasmeet Singh, and Vishal Koppula for their ideas, help and support in lab which played a big part in the completion of this thesis.

Finally, I would like to thank all my friends who helped me for testing the device, Arindam Banerjee, Gurjeet Singh, Randeep Singh, and Satbir Singh.

Table of Contents

Abstract.....	ii
Acknowledgements.....	iii
Table of Contents.....	v
List of Figures	vii
List of Tables.....	x
List of Abbreviations.....	xi
Chapter 1 Introduction	1
1.1 Motivation	1
1.2 Research objective	3
1.3 Method of approach.....	3
1.4 Thesis outline.....	5
Chapter 2 Literature Review	7
2.1 Background.....	7
2.2 Electronic Travel Aids	11
2.3 Sensory Technology and Computer Vision Algorithms.....	13
2.4 Summary of Literature Review.....	17
Chapter 3 Methodology.....	21
3.1 System Configuration: Microsoft Kinect	23
3.1.1 Microsoft Kinect.....	23
3.1.2 System: ASUS Rogue	31

3.2 System Operation	32
Chapter 4 Software Design and Implementation.....	34
4.1 Introduction	34
4.2 Kinect Interface and Working.....	37
4.2.1 Kinect Initialization.....	37
4.2.2 Skeleton Tracking.....	40
4.2.3 Connecting Kinect to Visual Studio.....	42
4.2.3 System Workflow.....	43
4.2.4 Software Interface	46
Chapter 5 Results and Discussions	48
5.1 System Testing.....	48
5.2 Test Course	50
5.3 Auditory Guidance	58
5.4 Results.....	59
Chapter 6 Discussion.....	74
6.1 Future Work.....	74
6.2 Conclusion.....	74
References.....	76
Appendix.....	85

List of Figures

Figure 2.1 Global causes of visual impairment [3].....	8
Figure 3.1 Project Design Methodology [52]	21
Figure 3.2 Kinect Sensor [54].....	24
Figure 3.3 Kinect horizontal Field of View in default range [55].....	25
Figure 3.4 Kinect vertical Field of View in default range [55].....	25
Figure 3. 5 Primary Kinect application categories [56]	26
Figure 3.6 Kinect Depth Data Processing	28
Figure 3.7 (a) Raw image from infrared camera showing emitted IR pattern as projected on a recliner chair, (b) Corresponding depth image [60]	29
Figure 3.8 Schematic representation of Triangulation method [61]	29
Figure 3.9 Person equipped with the system	32
Figure 3.10 Bluetooth earphones	33
Figure 3.11 Go Pro mounts	33
Figure 3.12 Power Source.....	33
Figure 4.1 System Configuration	36
Figure 4.2 Kinect Status	38
Figure 4.3 System Initialization	39
Figure 4. 4 Process flow that creates joint points from raw depth data [64]	40
Figure 4.5 Motion Tracking [56].....	41
Figure 4.6 Tracked skeleton with a total of twenty joints [65]	42
Figure 4.7 System workflow	43

Figure 4.8 Free Mode of guidance	45
Figure 4.9 Figure: Software Interface	46
Figure 5.1 User Starting point	50
Figure 5.2 User navigating using Follow mode	50
Figure 5.3 Kinect scanning QR code.....	51
Figure 5.4 User taking exit	51
Figure 5.5 User turn right after auditory guidance from QR code scanning.....	52
Figure 5.6 Navigating through corridors	52
Figure 5.7 Navigating in corridors using Follow mode.....	52
Figure 5.8 Kinect is providing audio information regarding elevators by scanning QR code.	53
Figure 5.9 User approaching stairs and further guided by QR code	54
Figure 5.10 User going up the stairs	54
Figure 5.11 Navigating through stairs using follow mode	54
Figure 5.12 User going downstairs after audio information form QR code	55
Figure 5.13 User Climbing downstairs using Follow mode.....	55
Figure 5.14 Navigating through hallway level 1	56
Figure 5.15 User receiving information about the destination from the QR code	56
Figure 5.16 Destination	56
Figure 5. 17 Follow mode guidance	57
Figure 5.18 Experimental Environment for guidance	58
Figure 5. 19 Minimal Audio Information provided to the user	59
Figure 5.20 Starting point at level two	60

Figure 5.21 Follow mode starting point at level two.....	61
Figure 5.22 Lab Exit	62
Figure 5.23 Right turn on the second floor	63
Figure 5.24 Elevator level two	63
Figure 5.25 Follow mode second floor	64
Figure 5.26 Stairs going up	65
Figure 5.27 Stairs going down.....	65
Figure 5.28 Follow mode stairs going up.....	66
Figure 5.29 Follow mode stairs down.....	66
Figure 5.30 Corridor level one	67
Figure 5.31 Navigating in hallway level one	68
Figure 5.32 Course endpoint.....	69
Figure 5.33 Table showing Prototype testing for blindfolded sighted participants	70
Figure 5.34 Follow mode using QR code in a lab.....	71
Figure 5.35 Follow mode using QR code in hallway	71
Figure 5.36 Follow mode with interference.....	72
Figure 5.37 Follow mode with another person in between user and the person being followed	73

List of Tables

Table 2.1 Summary of Literature 20

List of Abbreviations

SLAM	Simultaneous Localization and Mapping
ETA	Electronic Travel Aids
HCI	Human Computer Interaction
WPF	Windows Presentation Foundation
HRTF	Head Related Transfer Function
SIFT	Scale Invariant Feature Transform
GIS	Geographic Information System
GPS	Global positioning system
SONAR	Sound Navigation and Ranging
AGV	Automated Guided Vehicle
IDE	Integrated Development Environment
AMD	Age-related macular degeneration
VAS	Virtual Acoustic Space
WHO	World Health Organization
RFID	Radio-Frequency Identification

Chapter 1

Introduction

1.1 Motivation

Visually impaired people suffer from some form of deprivation, which affects them physiologically and psychologically. An estimate done back in 2007 recorded that half a million Canadians have significant vision loss and around 5.5 million have major eye disease which could lead to eye damage, which directly influences their quality of life. The National Coalition for Vision Health report indicates a potential crisis in Canada in eye health care. Moreover, vision loss is increasing at an alarming rate in Canada [1]. According to the World Health Organization in 2014, 285 million people are estimated to be visually impaired worldwide: 39 million are blind, and 246 have a low vision [2]. Considering the future, the reason to worry is that the rate at which it is increasing is around 50,000 people every year. The prevalence of sight impairment is expected to increase about 30 percent roughly, which could be leading to a crisis in vision health.

The blindness or partial sightedness may be due to various reasons such as Age-related macular degeneration (AMD), Birth, Injuries, diabetic retinopathy, glaucoma, cataracts, refractive error and other Medical reasons [3].

Some facts about vision loss in Canada [4]:

- \$15.8 billion: Total annual cost of vision loss in Canada in 2007
- \$30.3 billion: Projected total annual cost of vision loss in Canada in 2032

- Over 4 million: The number of Canadians with an age-related, blinding ocular disease in 2007 (will double as the number of aged doubles)

For the vision health, the predicament is due to:

- Aging population in Canada
- Scarcity of eye specialists
- Medical cost of vision care is high
- Lack if precautionary programs and knowledge in people
- Lack of Research and development in eye care

This experimental study is necessary to provide the visually impaired individuals with a device which can help then navigate indoors. Many people who got along with these visual aids found these to be aids to be helpful in day to day life. Blind people feel assistive technology trustworthy and helpful for navigating [5]. The device and the algorithms are brought together to provide better vision as compared to the conventional cheap sensors as those cannot produce the same quality output for various data, such as Depth data, face recognition and voice recognition which is the prominent part for the Kinect.

The exciting part about the Kinect Xbox 360 is that it is programmable using various languages such as C++, C#, and Visual Basic.NET [6]. Moreover, it has open source Application Programming Interface (API) for the developers and a lot more provided by Microsoft (Kinect for Windows) itself. These API's help us to build Kinect-based applications.

The generous and efficient way to help the user with the device which we have programmed is by attaching it to the chest mount, and the Kinect will get us the dynamic data which would be further processed in real-time, helping the visually impaired person with decent directions messages of the surroundings through the audio output.

1.2 Research objective

Being blessed with all the sensory organs, we usually start taking our body for granted. For people suffering from partial or complete blindness, even a little ray of light is precious. Everyday tasks include much navigation, in which we as rational human beings, don't suffer many hindrances, due to our eyesight. However, for a visually impaired person, accomplishing these simple tasks, without hurting themselves or people around them, can sometimes prove very challenging.

The visually impaired people strongly agreed that the assistive technology makes their tasks accessible and less stressful, that gives them more freedom to navigate in the surroundings [7].

1.3 Method of approach

So, in this thesis project, a more reliable and better quality yet an affordable prototypical device is designed that assists the user to navigate. Microsoft Kinect Xbox 360 device is the base for our research, to design it as the navigational tool. It is a robust system, which contains an infrared sensor, infrared projector, a microphone array, and

an RGB camera. Kinect is a boon for developers, allowing us to program this ideal device for helping people navigate indoors using computer vision algorithms developed in C#. All the information which we get from these sensors goes to the laptop where further processing takes place. The processed data from the Kinect helps to further guides the user with the aid of acoustic outputs from the computer to the individual to guide them through the surroundings. For example, while walking through a room, it will prevent the user from obstacles such as chairs, tables, walls, doors [8] and it will also help in the stairway by alerting them through auditory guidance.

The integrated sensors in the Kinect are so robust that it even works in low light in which most of our conventional sensors do not work. With this kind of device, we want to provide the user with a better understanding of the very surrounding. Moreover, it is more reliable, painless and inexpensive method to help navigate.

Specific scenarios experimented, to help the user navigate and the approach has polished test results using the Kinect.

Three main scenarios tested in our study are,

1. Navigate indoors such as in classrooms and laboratories, guiding the visually impaired person through obstacles such as tables, chairs, lab partitions, and cabins.
2. To able to detect the doors while in the hallways or corridors and to recognise stairway going up or down.
3. Follow a specific person out of three in the lobby, with audio guidance through Bluetooth headphones.

In the final part, we conducted some interviews with individuals (sighted human objects with blindfold) who were the part of experimentation and used our prototype and

answered various questions regarding our assistive technology and how they felt while testing it as the part of the project.

1.4 Thesis outline

This thesis consists of 6 chapters that walk the reader through different stages of work. It will cover the aspects such as system development, research part, hardware and software used for prototyping inclusive all the results.

Chapter 2 presents a literature review of the relevant work, done by the researchers and published on similar technology around the world. Furthermore, it also discusses the advantages and limitations of the various studies.

Chapter 3 provides information and introduction of the device used for the experimentation, and all other perspectives such as the reason behind why we choose it, the advantages, qualities, constraints and challenges of using it. It also describes the working of the device.

Moreover, the chapter also discusses the navigational strategy of the user with the device and presents the advantages of using a specific approach.

Chapter 4 describes the Kinect and the way it programmed for sensor fusion to get various data such as distances and depth data, and Color images for image processing. It also includes appropriate graphical representations showing the working of it.

Furthermore, this chapter gives the detailed explanation of the algorithms and codes which run the Kinect to help the user navigate. The programming language, C# (C sharp) contributes to building the platform for our device and further used for generating various results.

Chapter 5 details the experimental data and discusses specific scenarios. It also includes the review and implementation of using the Kinect sensor. Block diagrams used to support the results from sensor fusion and prepared in a way to make the reader understand the approach considering the navigation using the Kinect sensor.

Finally, Chapter 6 concludes the thesis by conclusions drawn and providing recommendations for the further research.

Chapter 2

Literature Review

2.1 Background

The impact of any disability on an individual fluctuates broadly, in the case of visually impaired people it varies as much as the variation in human personality itself. It is then hard, to sum up on innovative guides which may help them to build up some possible base from which to talk about the subject.

Blindness is sensory loss and reduced capability of vision. It makes even the simpler tasks, impossible. The effort of the researchers to serve blind people by assisting them in any possible way is impressive. Advancement in technology brought a boon in this field allowing them to use various sensors and devices to help them.

According to the survey done in back 2014, World Health Organization (WHO) revealed some shocking figures:

285 million people are estimated to be visually impaired worldwide: 39 million are blind, and 246 have a low vision [2]. However, the only positive thing which came out of this survey is that 80% of all visual impairment can be prevented or cured.

Globally the primary causes of visual impairment are shown in the pie chart below.

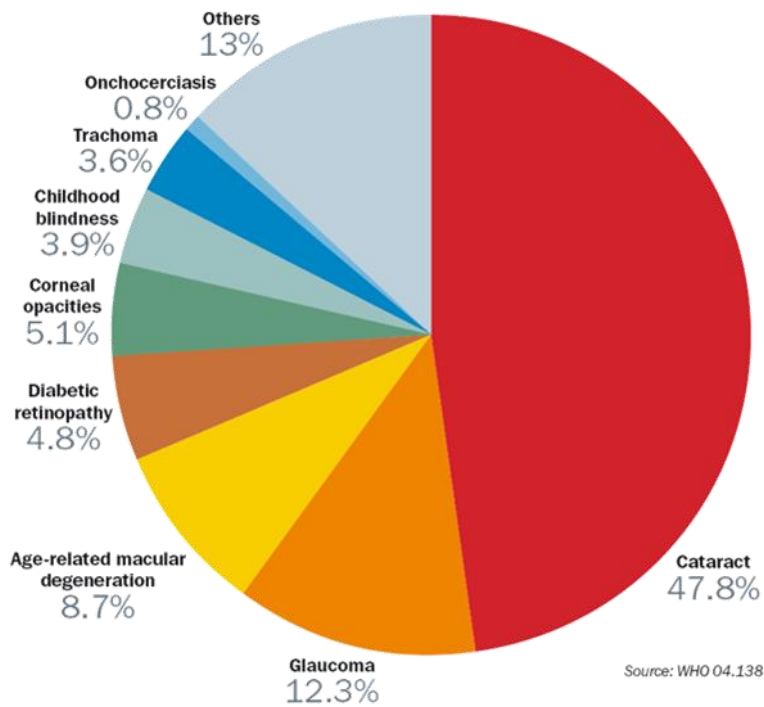


Figure 2.1 Global causes of visual impairment [3]

The literature review presented thrives for the research into visual aids for reducing the affliction of the blind people. In last three decades, many solutions have been proposed and were made available to the blind users through various sources such as white canes, laser canes, binaural sensing aids, Braille and guide dogs for blind people.

There has been some innovation in the history which have survived until today with minor changes in them, while some of those are yet to be thoroughly tested by a critical field. Recently the rapid advancement in the technology helped researchers in the field of

electronic and computer technology to adapt and solve the problem for blind people. There are some very fine, highly innovative aids now available to the area of blindness.

However, most of them are costly or not practically sound. Some people who got along with these visual aids found these to be helpful in day to day life.

The need for today world is to design reliable aids of low-cost, yet capable of delivering the right amount of information transfer to the blind person. Achieving this is tough, but looking at recent studies in this field it seems to be a highly rewarding and will be a test of one's technological stamina.

Blindness inflicts an individual's sense to perceive the surrounding environmental conditions. However, the use of hearing and touch sense by the visually impaired people help them through these day to day problems and reduces its effect onto their lives. The major aspect of blindness is locomotion control and guidance relative today fastly moving life for the people.

The primary requirement of any aid mobility is obstacle detection, and the cane can only provide safety for the upper body. Moreover, this technique is not foolproof against all kinds of environments with different kinds of hazards and obstacles.

The technological solution for the visually impaired people is achieved using computers and various electronic sensors. Otherwise, is an impossible task to provide a blind user with a device to help them navigate. A logical trail at the right moment might make undertaking conceivable.

To help blind people navigate, we need to detect the prompt environmental conditions for obstructions to travel. Moreover, it is required to explore obstacles and hazards which the regular aids cannot notice.

Furthermore, unforeseen obstacles in various routine tasks planning severely hinder the navigation of the blind individual. Moreover, exploring obstacles and hazards which the regular aids cannot notice. The circumstances above lead to user's unwillingness to travel and restrict themselves to a confined space despite having an aid [8]. Moreover, these aids are not foolproof and do not provide hassle-free navigation assistance against all kinds of environments with different kinds of hazards and obstacles.

The visually impaired people do not have the freedom to navigate without assistance, as the information regarding the environment is not within the sensing limits of laser canes and ultrasonic obstacle avoiders [9].

Navigation system for blind has been studied by various researchers to increase the mobility of the blind. However, they were only concerned with guiding the user along a predefined route [10]. The evaluated guidance performance was best with virtual display mode for this study. Following year, a group of researchers aimed to provide blind people as much information about their immediate environment, by capturing the form and the volume of the space in front of the blind person and send it back to the user in the shape of a sound map using headphones in real time [11]. These studies were based on the creation of virtual acoustic space (VAS) [12] giving the person more independence of orientation and mobility. Fundamentally, VAS is the perception of space using only sound. As researchers started working in this field they wanted a delimitation of observed capabilities. The study of the developed prototype in everyday life conditions, on exploring

how blind people learn to use new strategies to improve their perception of the environment and the exploration of the possible cortical brain areas involved in this process, using functional imaging techniques.

The noteworthy disadvantage for the blind is the lack of information regarding the surroundings where they can easily miss out on obstacles, landmarks, and their velocity, which are most necessary to a sighted individual to navigate through familiar or unfamiliar environments.

In the present generation, people addicted to social media and visual media tools and other applications, for the blind people seem to be an impossible task. Various tools for assistance have helped people suffering such as screen readers, braille terminals, screen magnifications and paper embossing. Nonetheless, these solutions remain limited when it comes to accessing and understanding visual contents. Another device attempted to aid in blind people for interacting with image contents and indoors navigations, presented in work [13] as an alternative to previous studies by using vibrating screens for a better understanding of the digital image contours. This approach allowed blind people to have a better understanding of various situations and contents.

2.2 Electronic Travel Aids

Electronic Travel Aids (ETA), have been used by researchers in the past [14] as an assistive device which transforms the environment surroundings into another sensory modality. These aids have proven to help visually impaired people navigate with high

confidence, physiologically and psychologically. These devices can detect obstacles in the path of the user. ETAs has three building blocks, the sensors, software interface, a feedback mechanism [15]. The sensors transmit the data to the system, which is further processed using the designed software and sends the user with the surrounding information, and a real-time feedback helps so that there are no hindrances in the user way.

Many ETAs reviewed in this study used various sensors for collecting environmental surrounding conditions for the user. Some of the sensors used are ultrasonic, pyroelectric, inertial devices, GPS, phone cameras, stereo cameras, sonar, depth sensors and so on. The data inputs from these sensors get processed using various processing tools such as programming languages on a computer, microcontroller, control box, remote server, and the blind users get only the useful or necessary information, by using audio outputs or vibration feedback/ tactile feedback.

Many other studies showed that blind users need aids to navigate and detect obstacles in their path to help perform wayfinding.

These devices have limited usability as they have lot many assumptions for the environment such as less crowded testing conditions, not many moving people, obstacles at the same place or familiar environments and can't navigate without actually touching the obstacle.

2.3 Sensory Technology and Computer Vision

Algorithms

Some research work used ultrasonic sensors to augment the performance of the guide cane [16], [17]. These sensors helped the user to detect barriers and steer accordingly, which proved better as compared to usual cane, as guide cane provided path easily and without much efforts. The use of laser and vision sensors [18], [19] enhanced user's confidence while navigating and was a reasonable mode for providing high information in real time using laser triangulation system. In 2011 [20] a group of researchers provided an overview of the literature available on assistive technologies with the focus on aspects such as assistance device for the daily life use an indoor/outdoor navigation in a dynamic environment. They also provided the list of solutions available for helping visually impaired people, such as navigation system, obstacle avoidance, and obstacle localisation.

Another related work presented an algorithm on Speeded-Up Robust Feature (SURF) [21] to assist blind people providing them information about a safe path to navigate by recognising movements of the object that interest user. This system specifically addressed the trajectory estimation at the pedestrian crossing to help visually impaired people. However, the functionality of the presented system which used GPS and Kinect camera was not precise and was limited by its purpose.

Many research groups made efforts to create electronic aids for the blind. After the advancement in the area of computers and electronics, a stereo vision based navigation

assistance was developed in 2010 [22]. In this study, a head-mounted robot vision for visually impaired is presented, which incorporates visual odometry and SLAM and tactile vibration motors give the output. However, this study was not foolproof as it was not useful in many realistic conditions and even the device used was inefficient and had many constraints.

The development in technology helped the blind people to overcome the difficulties that the dog and cane do not respond. Electronic and sensory substitutions contribute to transforming the data from the source into another sensory modality (auditory or tactile). The paper addresses that the recognition and localisation of objects are necessary to provide travel assistance for mobility to the blind. Which further helps in navigation and identifying objects. Fast and robust algorithms (Scale Invariant Feature Transform) contribute to recognising objects in a video scene [23]. The major drawback of this study was that it was not real-time.

The ability of white canes is insufficient to provide hassle-free navigation assistance to the blind person and moreover cannot detect all the obstacles. The inability of blind people to perceive their surrounding is the reason why researchers all around the world are coming up with various navigation systems. Another group designed and built a cost-effective navigation system for both indoor and outdoor environments with the purpose to assist blind people [24]. They used the ultrasonic sensors/ sonar for getting the range of the obstacle and also used a microcontroller for data processing. The feedback information is transmitted using voice and vibration means to the user.

The aim of most of the available systems is to provide help to visually impaired people without and secondary help from another person. In a survey done in 2014 [25],

proposed a new framework by overviewing the essential aspects which will help visually impaired individuals and additionally provided possibilities of some other capabilities which augments for better results. They also listed some challenges in various areas which still require further research and development. The evaluation and comparisons made in this study proved that for obstacle detection, image processing has a significant role.

Further, they also proposed a scheme for the following capabilities such as obstacle detection, object identification, path and door detection, feature extraction for various objects, and digital reading contents. All these capabilities mentioned above need have an excellent image processing techniques and gesture recognition.

Developing a computer-aided tool/vision system is another solution to assist the blind user and is still a developing area. The aim of most of the available systems is to provide help to visually impaired people without any secondary help. A survey done in 2014 [25], proposed a new framework by overviewing the essential aspects which will help visually impaired individuals and additionally provided possibilities of some other capabilities for better results. They also listed some challenges in various areas which still require further research and development. The evaluation and comparisons made in this study stated that for obstacle detection, image processing has a major role. Furthermore, they also proposed a scheme for the other capabilities such as obstacle detection, object identification, path and door detection, feature extraction for various objects, and digital reading contents. All these capabilities mentioned above need to have an excellent image processing techniques and gesture recognition. Indoor auditory navigation system [26] presented in this study assists blind and visually impaired people,

using computer vision and markers in the environment. The user navigates in the surrounding environment using a webcam attached to the system. Whenever web camera detects a particular marker, audio assistance provides the user, with valuable information that enables them to navigate independently in the environment.

NAVI refers to systems that help or guide people; this system was designed by [27] after a review in this field [28]. This main idea behind this study is to make a person aware of the path and the obstacles in the path. The proposed system consists of sensors (depth and RGB sensors) embedded in shoes, control board and a response system (vibration and voice assistance). [29] An Ultrasonic Assistive Headset was developed for visually impaired and blind people. This headset guide user for obstacles using ultrasonic sensors, microcontrollers, voice storage circuit and solar panels. This device can be utilised both indoors and outdoors and can avoid obstacles quickly and accurately. [30] In their study Simultaneous Localization and Mapping Algorithm (SLAM) was used for visually impaired people for outdoor navigation. They used Android-based mobile phone having sensors such as an accelerometer, gyroscope, proximity and ambient light sensor. An application based on death reckoning SLAM algorithm is used for tracking and alerting user for obstacles. Vibration to audio signals was used to assist the blind person to follow the appropriate path. Another research presented by [31], is a wearable assistance system for helping visually impaired people. The system uses a stereo camera and sends acoustic feedbacks. The experimental study uses basic scene understanding, head tracking, and sonification that allows the user to walk in an unfamiliar environment and to avoid obstacles safely.

2.4 Summary of Literature Review

A summary of all the literature discussed above is depicted in Table 2.1.

Research Group	Research Focus	Components used	Uses
[32]	AGV, Kalman filter based model	Sensors, camera	Obstacle detection and navigation
[33]	GSM based navigation	GSM and GPS	Location-based services and voice feedback
[22]	SLAM, visual odometry, tactile cueing system and stereo vision	Stereo camera, vibration motor	Navigation assistance, obstacle detection
[34]	Computer vision based door detection algorithm	Camera and computer	Edges, corners for door detection
[35]	Stereo sound and image processing	Digital video camera fixed on the headgear, Stereo earphone, SBPS with chassis, NAVI vest	Obstacle detection and navigation
[36]	Kinect real-time image processing	Array of vibrotactile elements, helmet, Kinect sensor	Wayfinding system, collision avoidance
[37]	Stereo vision mapping	Sensors, acoustics devices, stereo vision camera	Obstacle detection
[38]	Stereo vision, Image processing, fuzzy inference	Wearable computer, Stereo camera as vision sensor and stereo earphones	Obstacle detection

[39]	Computer vision based approach, fiducial markers are used, image processing	Artoolkit markers, Kinect camera, audio information	Navigation guidance using auditory assistance (obstacles)
[21]	SURF (Speeded-up Robust Features) algorithm	Helmet, Kinect and GPS	Safe trajectory at a pedestrian crossing and Object recognition
[40]	SIFT, neural network	Video camera, sound systems	Object recognition, finding objects
[41]	RGB-D mapping, SLAM	Kinect	3-D modelling of the indoor environment
[42]	VizWiz:: Locate it, remote server, sonification, real-time computer vision	iPhone camera, Remote server, remote worker	Finding arbitrary things in the environment
[43]	Edge detection, corner detection	Video camera	Door detection
[44]	Depth sensing, OpenNI framework	Xtion Pro, headphones and laptop	Detecting humans and avoiding objects
[45]	Depth sensing	White cane, numeric keypad, ups battery, Kinect sensor	Objects, chairs, upward stairs
[46]	Basic image processing	Video camera and sound system	Obstacle detection
[47]	GPS-GSM based navigation assistant	Capacitive Touch Braille keypad, SONAR and smart SMS Facility	Navigation, audible message and haptic feedback
[48]	Depth based obstacle avoidance	LEDs, Asus Xtion sensor, computer	Navigation and Obstacle avoidance

[49]	Both sensing and image processing	Ultrasonic sensor and USB camera	Obstacle detection and obstacle identification
[23]	SIFT (Scale Invariant Feature Transform), Video recording, voice translation, image processing, visual distribution	Video camera	Recognizing objects in the video scenes and an auditory system to identify this information
[24]	Vibration and voice-operated navigation system	Ultrasonic sensor, microcontroller, voice renderings, vibration motor, sonar	Obstacle detection and its distance
[50]	Door detection method (depth and grey level images), image processing	Xtion depth sensor, voice output IC	The detailed configuration of stairs, steps, curbs, obstacles, etc.
[26]	Computer vision based approach, fiducial markers are used, image processing	Artoolkit markers, camera, audio information	Navigation guidance using auditory assistance
[28]	NAVI	RGB sensor, obstacle sensor, vibration motor, voice assist IC	Navigation, obstacle detection, distance
[29]	Ultrasonic Assistive headset	Ultrasonic sensor, microcontroller, voice storage circuit and solar panel	Obstacle avoidance
[30]	Augmented reality glasses, SLAM, sonification	Coloured camera and depth IR camera, micro projector, control box	Object distance, audio assistance, obstacle detection

[51]	Computer vision, voice recognition	GIS and GPS, mobile	Navigation based on static and dynamic data, optimised routes based on user preference
------	------------------------------------	---------------------	--

Table 2.1 Summary of Literature

The literature studied for this thesis tells us that there is much scope for improvement in case of visual aids for visually impaired or blind people. There have been a lot many solutions provided by various researchers to help address these problems, but much of those are just limited to specific area.

Chapter 3

Methodology

The need for today world is to design reliable aids of low-cost, yet capable of delivering the right amount of information transfer to the blind person. Achieving this is difficult, but looking at literature as discussed in chapter 2, recent studies in this field it seems to be a highly rewarding and will be a test of one's technological stamina. In order to design our product, an iterative process was used for its development methodology. The design process is divided into three stages: Problem specification, Iteration, and conclusion [15].

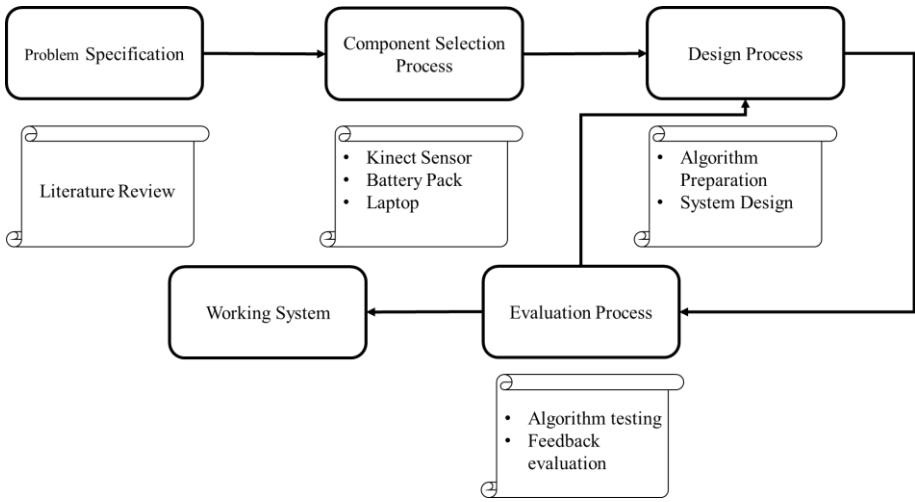


Figure 3.1 Project Design Methodology [52]

The first phase of this process was to specify the problem, in which an extensive literature review is presented in the second chapter where various aids have been evaluated and studied. To improve the final products lot of iterative measures were taken, such as the blindfolded sighted users were interviewed for their views and perspectives about how they feel after using this navigational aid. Moreover, then make changes to the design and the program accordingly and incorporate these into the final product.

After an extensive overview of the problem, the second phase was to select the best-suited components which are cost-effective, robust, non-intrusive. For this Microsoft Kinect sensor for Xbox 360 was chosen because of its affordability, open source accessibility and accuracy as compared to the conventional sensors.

Now the very next need for the system was to provide the Kinect sensor DC power supply as the prototype would be used for navigating. So, a portable power supply was used for its dynamic working. The third most important component was the laptop which was used to run the Visual studio in which the algorithms were written for running the Kinect sensor accordingly to help the visually impaired user. As the navigation process needs real-time image processing, therefore the laptop chosen for this was powered with latest specifications and a fast processor. The above-mentioned design process is integrated with the components and is evaluated with the help of blindfolded individuals.

This cycle which has design, development and testing phase helped to improve the overall system. The integration of all the components and user feedback led to the development of the working system. This small loop helped to find the limitations and other loopholes. This led to the implementation of QR code technology to augment the Kinect sensor output and provide the visually impaired better navigational guidance.

To help blind people navigate, the feedback from each stage helped improving the final system in a way to detect the prompt environmental conditions for obstructions to travel. Moreover, the final working system with the iterated algorithm is tested.

3.1 System Configuration: Microsoft Kinect

In this study, a computer vision system for navigation is proposed which is not limited to the sensor itself; it comprises of three main components. These components have their unique functionality, first Microsoft Xbox 360 Kinect sensor, used for collecting the environmental information (both depth images and RGB images). The second is the image processing algorithm written in C sharp language, performed on a laptop, and the final element, a Feedback system which assists the visually impaired person in navigation by providing directional information through auditory output using Bluetooth earphones.

3.1.1 Microsoft Kinect

Microsoft Kinect sensor was introduced for Xbox 360 gaming platform by Microsoft Corporation in 2010 as a gaming accessory for the console. It received high appreciation and sold more than 10 million devices by early 2011. Right after its launch the computer vision scientists and developers started using Kinect as a potential device for sensing purposes. The ease of using it and programming it as per needs provided with open source development kits it became a tool for non-commercial purposes and encouraging further interests in the device. Kinect software development kit (SDK) allowed developers to write Kinect based apps in C++/CLI, C#, or Visual Basic .NET [53]. The cost-effectiveness and qualitative depth imaging of the Kinect sensor gave it an upper hand

over the conventional sensors. The Kinect is also very well documented, with multiple SDKs available.

Technological innovations led to Microsoft Kinect for Xbox 360, a motion sensing device for a video game consoles [53]. Onboard it has depth sensor, IR emitter, RGB camera, multi-array microphone and a motorised tilt. RGB (i.e. red, blue and green) and depth streams use 8 bit and 11 bit VGA resolution video stream.

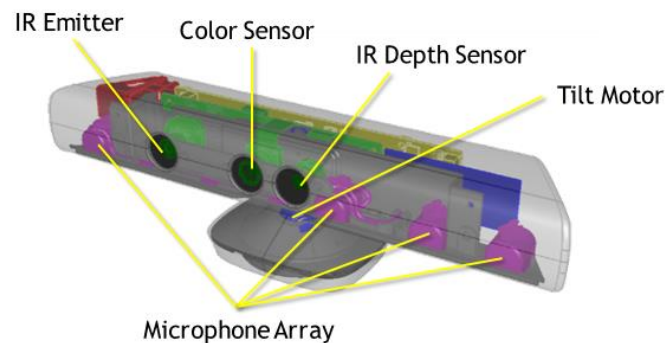


Figure 3.2 Kinect Sensor [54]

The colour sensor captures and streams the colour video data at 30 frames per second (FPS) with a resolution of 640×480 pixels or 1280×1024 at a lower frame rate. The field of view (FOV) for the camera ranges from 57 degrees horizontal and 47 degrees vertical. Kinect is capable of generating an image-based 3D reconstruction of an object or a scene. The processing is done using depth data with a stream resolution of 640×480 pixels. Kinect can capture a user standing between 0.8 meters and 4 meters which are the depth sensor range.

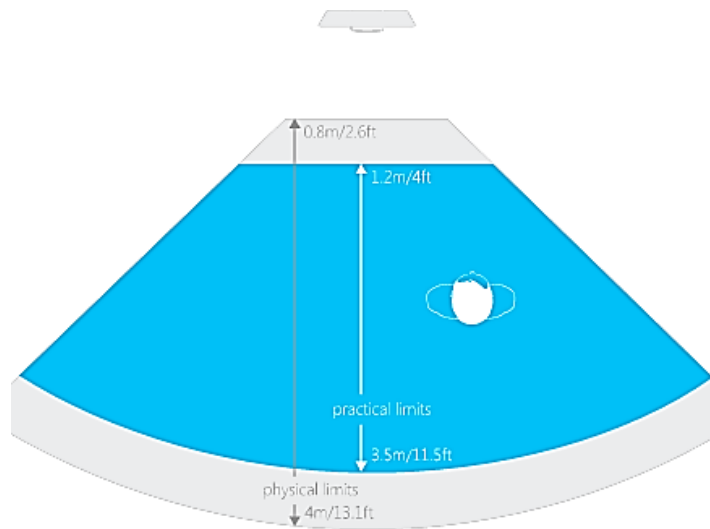


Figure 3.3 Kinect horizontal Field of View in default range [55]

In near range mode, Kinect can see people standing between 0.4 meters (1.3 feet) and 3.0 meters (9.8 feet); it has a practical range of 0.8 to 2.5 meters.

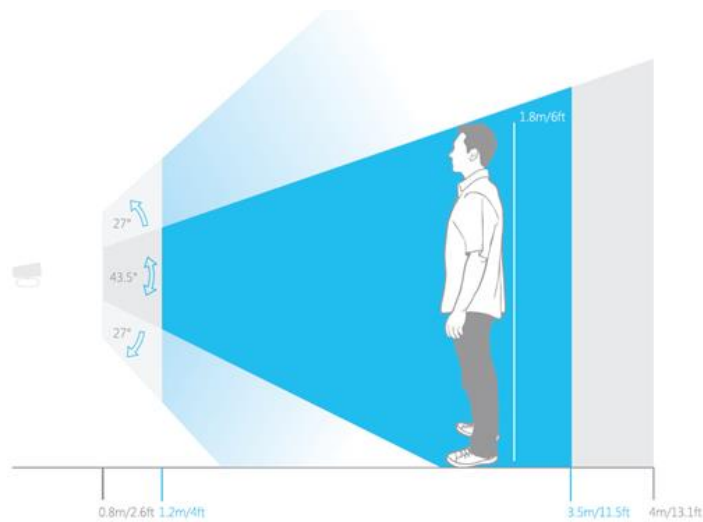


Figure 3.4 Kinect vertical Field of View in default range [55]

The sensor has been used by various researchers for visually impaired or blind aids. Interestingly, applications pertaining healthcare system have captivated the most research and advancement interests.

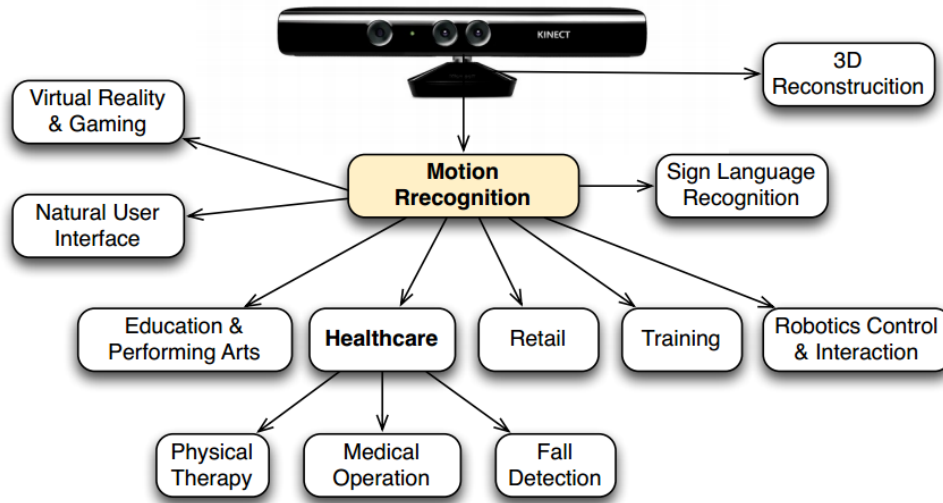


Figure 3. 5 Primary Kinect application categories [56]

Leading application areas of Kinect is demonstrated in Figure which ranges from healthcare, education, retail, training and gaming to robotics control, natural user interface, sign language recognition, and 3D reconstruction which has a substantial impact on 3D printing advancement. It has proved to be a more accurate and precise for navigational purposes.

3.1.1.1 The Kinect Technology

Depth and colour data are recorded at the same time at a maximum rate of thirty frames per second by the Kinect sensor. A coloured point cloud including approximately three hundred thousand points in each frame is resulted from combining the depth and colour images data. An elevated point density and also a total real-time point cloud of an indoor area can be acquired by recording the consecutive depth data. An examination of the systematic and random failures of the data is required in order to fully apply the mapping potential of the sensor. In order for the depth and color data to be aligned properly, the systematic failures should be firstly addressed which is associated with the recognition by a mathematical model of depth measurement and its calibration parameters [57, 58].

The Kinect depth image data is displayed on the output screen by the process shown in figure 3.6. [59] The Prime sense chip sends a signal to the IR projector to start emitting an invisible electromagnetic light onto the object or the scene. It also sends the signal to IR depth sensor to initialize and capture the depth stream and this information is sent back to the chip where the frame by frame depth stream is created for the display.

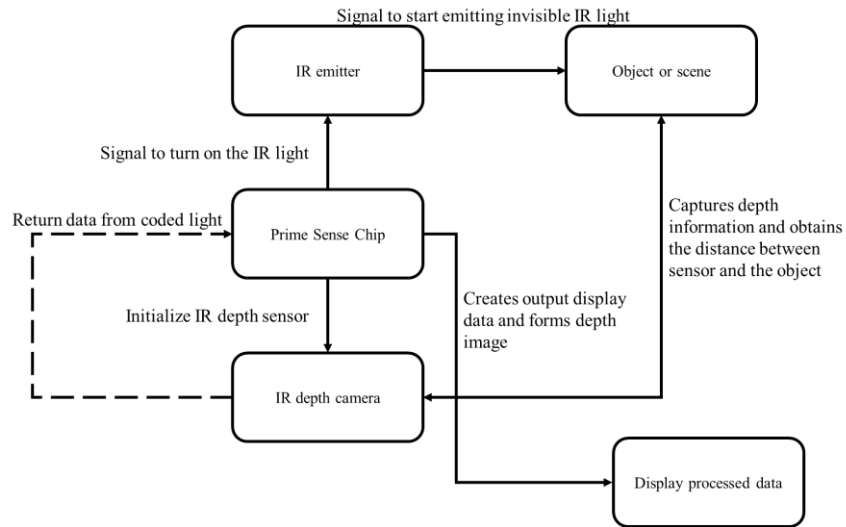


Figure 3.6 Kinect Depth Data Processing

A Kinect camera uses triangulation method for measuring the depth of objects. The IR projector projects the laser pattern which gets reflected by the object in the sensing range and IR camera triangulates it for depth map by recapturing the projected speckles.

Following the projection of the speckle onto an item, the speckle's location will be moved in the baseline path between the projector and the perspective centre of the infrared laser. A disparity image is resulted from measuring all speckles by a simple image correlation process. The distance from the sensor can be retrieved from the disparity image for every pixel. Figure 3.7 demonstrates the measurement of depth from the speckle pattern.

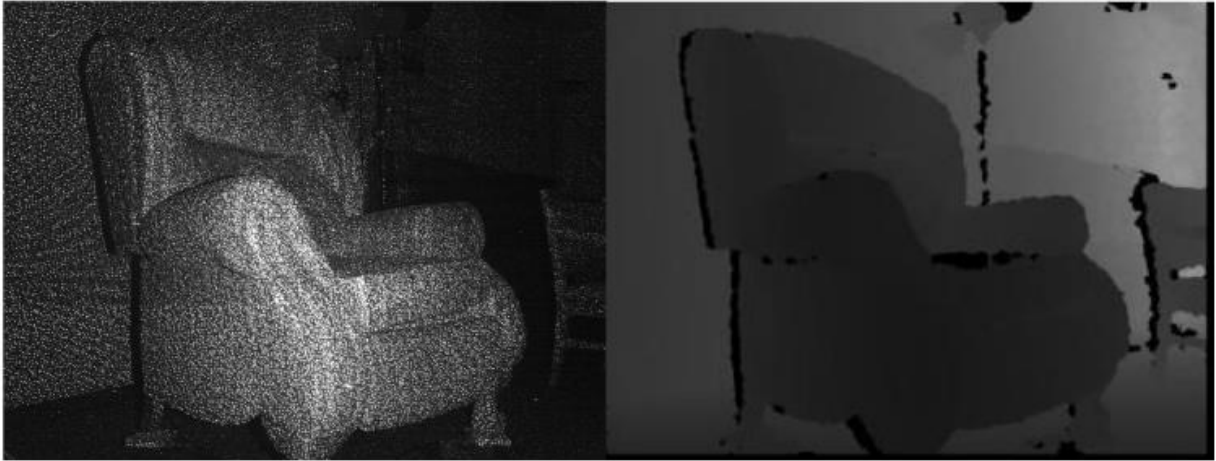


Figure 3.7 (a) Raw image from infrared camera showing emitted IR pattern as projected on a recliner chair, (b) Corresponding depth image [60]

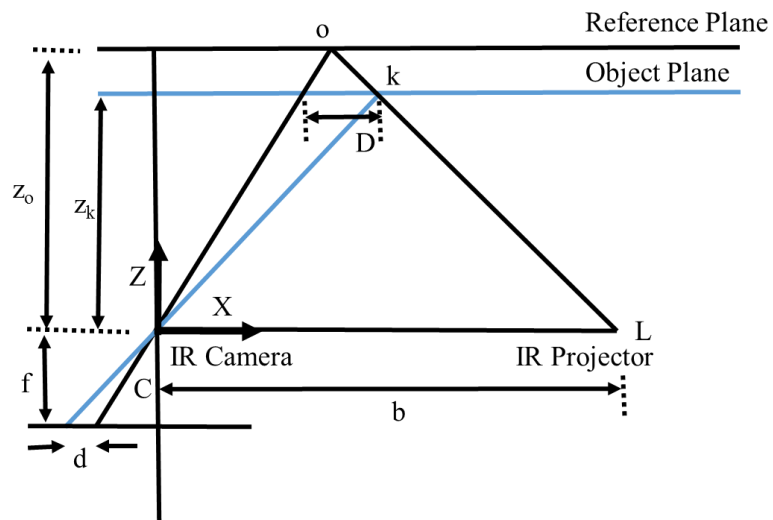


Figure 3.8 Schematic representation of Triangulation method [61]

Figure 3.8 demonstrates the three-dimensional coordinates of the item points; a depth coordinate system with its origin at the perspective centre of the infrared sensor is defined. The Z-axis is perpendicular to the image plane towards the object, the X-axis which is perpendicular to the Z-axis is in the direction of the baseline b (the distance between the infrared camera center and the laser projector), and the Y-axis is perpendicular to both the X and Z axis leading to a right-handed coordinate system.

The ratio of disparity D and depth distance d may be obtained as [61, 62]:

$$\frac{D}{b} = \frac{(z_o - z_k)}{z_o} \quad (1)$$

The coordinate system has its origin at the centre of the IR camera. Z and X axis are perpendicular to each other. b is the baseline between the IR camera and the IR projector, z_o is the assumed position of the object on the reference plane, and z_k denotes the depth or the distance of point k in object space. In Equation 1, D is the displacement of k in object space or the disparity of object's position between the reference and the object plane.

Further, the ratio of intrinsic parameters and depth parameter is given by.

$$\frac{d}{f} = \frac{D}{z_k} \quad (2)$$

Where d is the depth distance/ observed disparity, and f is the focal length of the Infrared Camera.

By substituting D from equation 2 into 1 and expressing z_k in terms of other variables, z_k is obtained:

$$z_k = \frac{z_o}{1 + \frac{z_o}{f \times b} d} \quad (3)$$

The above equation 3 is the fundamental mathematical model for obtaining depth of the witnessed disparity provided if the constant variables z_o , f and b can be identified in the calibration process.

3.1.2 System: ASUS Rogue

Asus Rogue laptop is used as the processing platform for navigational aid. It houses a powerful Intel i7 processor which runs Windows 10 which is compatible with running Microsoft Visual Studio 2015 and very much capable of running Kinect. It can run the heavy algorithm for real-time image processing without any visible time lag. The laptop is carried in a backpack to carry over the shoulder by the user. Figure 3.9 shows the blindfolded user equipped with the navigational aid.



Figure 3.9 Person equipped with the system

3.2 System Operation

When the system runs the algorithm, Kinect sensor starts capturing the depth and RGB data within the vertical and horizontal range of the sensor. This data is then sent back to the laptop for image processing in real time without any noticeable time delay and provides useful directional feedback to the user through the connected Bluetooth earphones figure 3.10. The Kinect sensor, shown in Figure shows the person equipped with the kit, where the sensor is mounted on the chest using GoPro chest mount right at the centre which makes it robust, portable and stable shown in figure 3.11. The Kinect is

powered by chargeable 6000 mAh Li-Ion, 12V DC portable battery pack shown in figure 5.12 which can power the system for almost 8 to 10 hours. The image processing is being performed on the Laptop in the backpack.



Figure 3.10 Bluetooth earphones



Figure 3.11 Go Pro mounts



Figure 3.12 Power Source

Chapter 4

Software Design and Implementation

As discussed in the previous chapter, Asus laptop is used as the computing system which runs Microsoft Visual Studio 2015. Visual Studio is an Integrated Development Environment (IDE), used for developing various applications such as websites, web pages, web applications and mobile apps. This software development platform from Microsoft is provided with API, Windows Forms, Windows store, Windows Presentation Foundation (WPF) and Microsoft Silverlight.

Microsoft Visual Studio provides opportunity to developers to build applications using different programming languages such as C, C++, and VB.NET (via Visual Basic .NET), C# (via Visual C#) [63]. This software is also available for free for Community edition.

4.1 Introduction

Specific scenarios experimented, to help the user navigate and the approach has test results using the Kinect.

Three main scenarios tested in this study are,

1. Navigate indoors such as in classrooms and laboratories, guiding the visually impaired person through obstacles such as tables, chairs, lab partitions, other individuals and cabins.

2. To detect the doors and name the classrooms and labs by their names or numbers while in the hallways or corridors and to recognise stairway going up or down.

3. Follow a specific person out of three in the lobby, with audio guidance through Bluetooth headphones.

For the testing mentioned above scenario's, the system is designed for two different modes of guidance in accordance to the needs of the visually impaired person. In the Normal Mode of guidance, the user can roam freely indoors and make their way to their destination, where they would be informed about obstacles (both on ground or hanging), persons in their way, as well as stairs. Moreover, if in some case they do not receive a precise information, they are backed up with Quick Response Code (QR) which are put on at various locations in the building premises. These codes are readable much faster and can store a significant amount of information. Moreover, the user can get information such as stairs going down/up and some stairs, elevator and level information. The other mode of guidance is Follow Mode; in this mode, the visually impaired person can follow a particular person for navigational help and it would assist will not be altered even if anyone is in the range of the sensor.

Figure 4.1 shows the configuration of the system [50]. The algorithm used for image processing converts the data from the depth and RGB images, pixel by pixel, into various surface features as shown in the system configuration. It is then further processed and segmented into separate regions, and then it looks for the scene entities in these areas. Then these scenes are divided into left, centre and right region to assist the user which way to go avoiding all kinds of obstacles. QR code algorithm helps to scan various codes using in our study which store all the relevant data which the Kinect sensor can miss out; such as the depth data for stairs or number of stairs, elevator information, lab and classroom numbers.

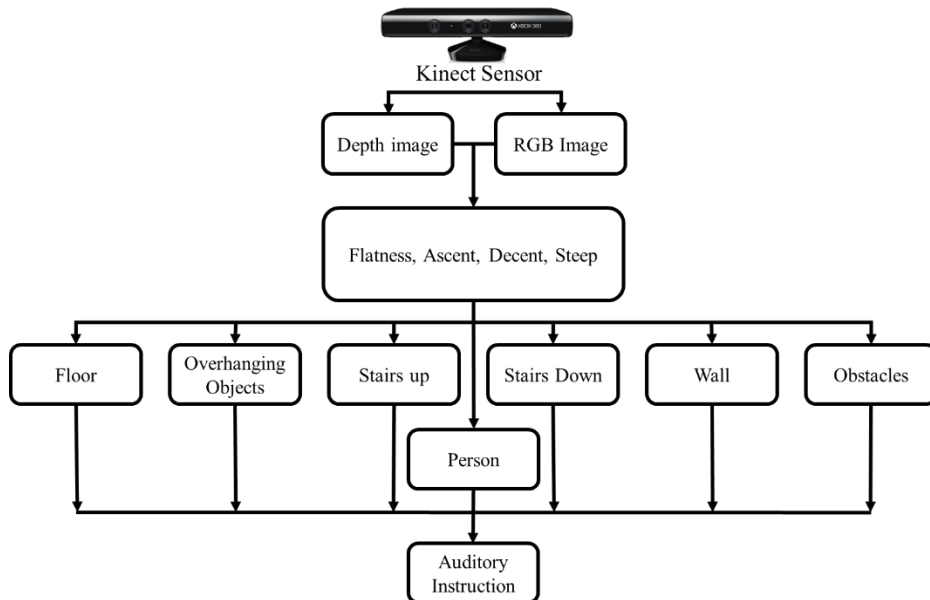


Figure 4.1 System Configuration

In the Follow mode, the depth camera is used with the help of image processing and skeleton tracking based approach to follow the nearest person to the camera. The sensor will not take account of other people passing through and will guide the user accordingly.

The complete experimental process until the directional feedback including Information collection and processing is a real time which makes it effective electronic aid. The final design uses C sharp libraries from Microsoft, OpenCV image processing libraries (Emgu CV), Zxing QR and Microsoft Kinect SDK. By framing all the above open source libraries, we accomplished this system, which helps visually impaired user to navigate. Avoiding all the obstacles and letting the user know about the environment. This all is done by initialising Kinect, and once it starts the process, it must capture and store all the depth and colour data and process its pixels to useful information.

4.2 Kinect Interface and Working

The task of the Kinect applications is to detect and initialise the Kinect device. Once, it starts the sensor the algorithm application must initialise and start capturing depth data stream and colour data from the surroundings. The SDK provided by Microsoft creates the interface between the Kinect application and the application. The sensor is accessed by calling the driver using API from the application. APIs help to directly talk to the sensor hardware, and processing the data captured by the sensor.

4.2.1 Kinect Initialization

The project is created in Microsoft Visual Studio 2015, using C# language in Windows Forms Applications (WPF). To use Kinect libraries, it must be referred in the

solution explorer. Most of the class are provided as part of SDK libraries, which help in Kinect operations. Kinect sensor working is simply initialisation (start the sensor), operation (colour stream, depth stream, skeleton tracking) and un-initialisation (stop the sensor).

The Kinect status is presented in the flowchart in the figure below which explains the different status of the device in various scenarios.

Once the device is connected and the power is turned off, it will show the NotPowered status. Similarly, unplugging the device from USB port will return the Disconnected status. If you plug it back in or turn the power on, it will first show the Initializing status before changing to the Connected status.

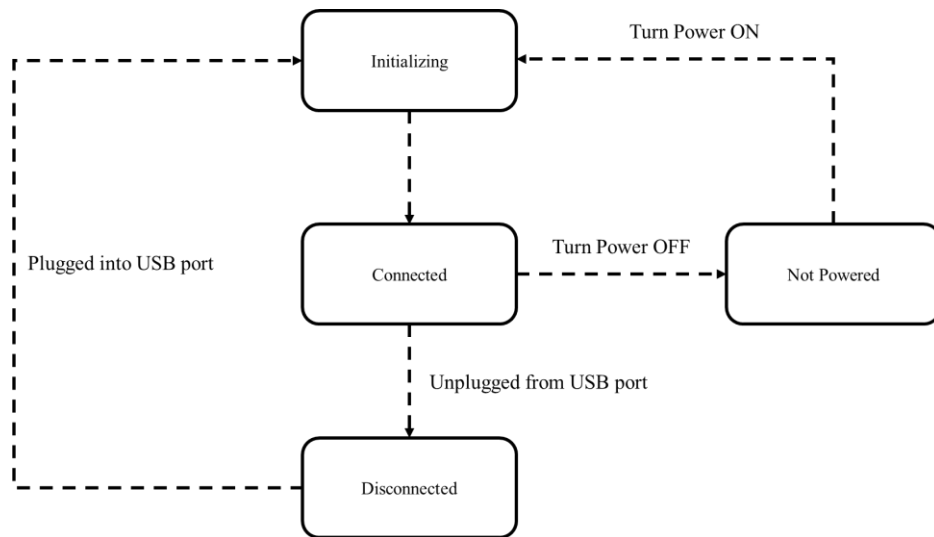


Figure 4.2 Kinect Status

Now, when the program starts, it makes sure that the Kinect is connected to the computer or laptop. If the system does not find any sensor connected, it will return the

program with error and exits. And if the Kinect is found the program proceeds normally. Then, various SDKs helps to connect to the inbuilt sensors for further use. The program now starts running various data streams using the Kinect Libraries, OpenCV to record the surrounding environmental data. As the processing of data is real time, so each stream always ready to indicate changes when new data is available. The data streams contain various surrounding elements such as obstacle (chairs, tables, partitions), floor, walls, doors, and people. Image processing pixel by pixel helps to differentiate between all the above-mentioned elements. Once the streams are initialised successfully it, then send the audio feedback else if does not initialise with success it will go through error and exit.

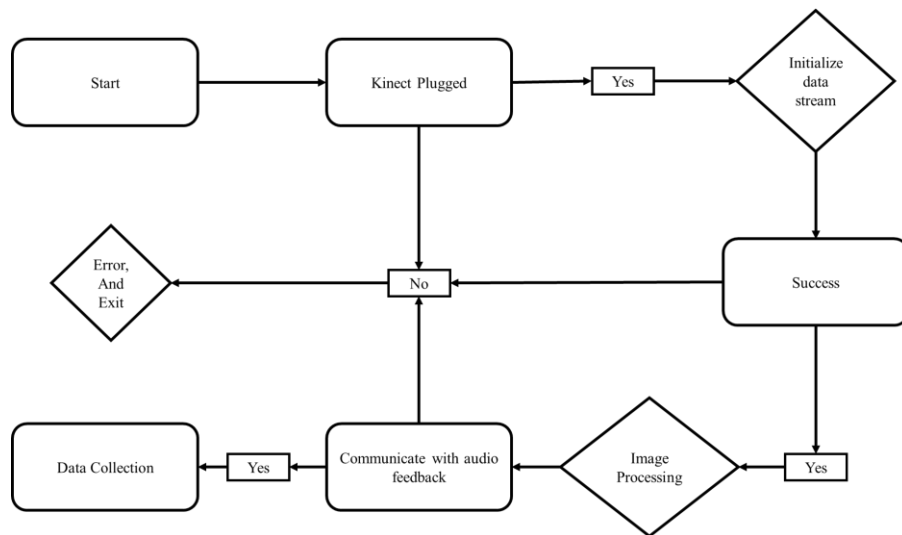


Figure 4.3 System Initialization

4.2.2 Skeleton Tracking

The integrated sensor on Kinect is so powerful that it can help tracking human movements and their skeletal which eventually helps our system to allow follow a specific person.

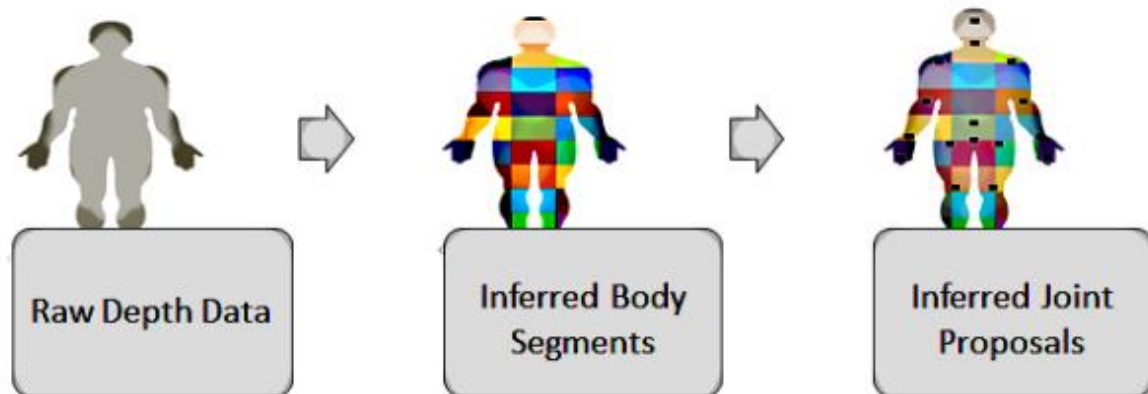


Figure 4.4 Process flow that creates joint points from raw depth data [64]

It uses a rendering pipeline process to process the depth data and matches it with the decision labelled data to generate the inferred body segments as shown in the above figure 4.4. So, once all the parts are identified based on the labelled data the sensor now starts identifying the body joints. And finally, it tracks the human skeleton and body movements [64].

The Kinect sensor uses the IR sensor that can recognise up to six people within its sensing range and skeleton application helps to track two users and their movements.

Kinect SDK comes with a skeleton estimation application from which we can obtain a stream of skeleton frames directly in real-time. The figure below shows the motion tracking analysis with Kinect sensor.

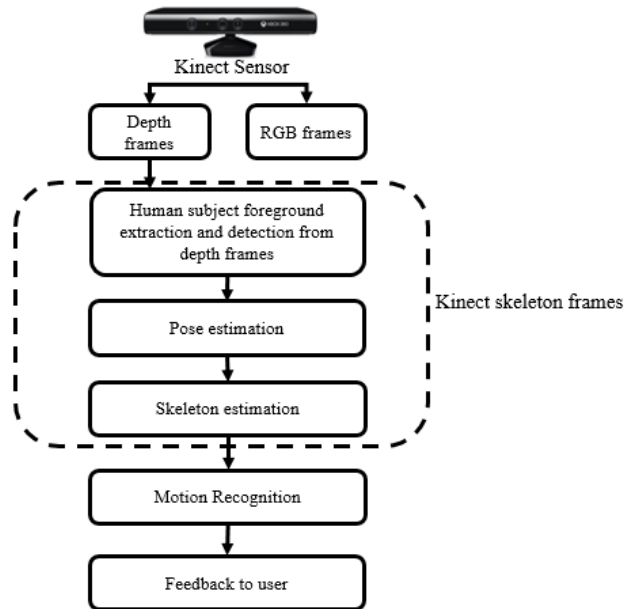


Figure 4.5 Motion Tracking [56]

Skeleton estimation process is shown in the above figure 4.5. Kinect sensor retrieves the depth data stream which contains one or more individuals in the sensor range. Now the process starts to get the Kinect skeleton frames, in which human subject foreground extraction is done from the depth frames. Then the processed data is matched against the trained model to estimate the pose, which helps to infer the position of skeleton joints and subsequently refined. Motion recognition now can be done once we have the information for the human skeletal. And further, the feedback is provided to the user.

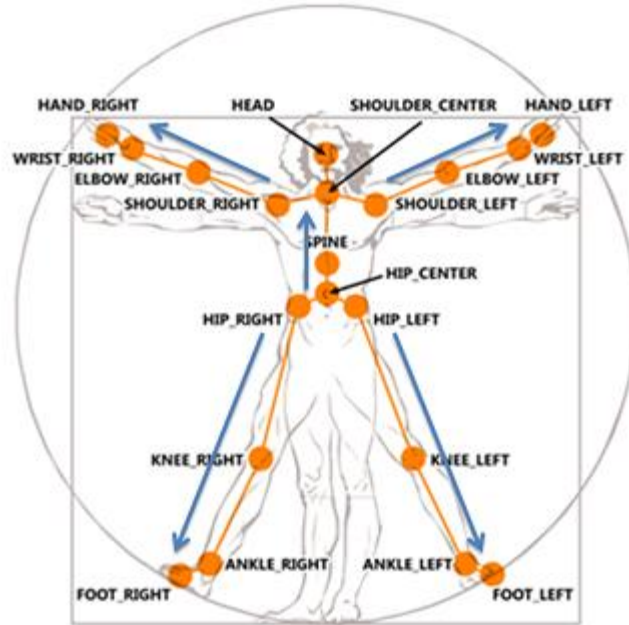


Figure 4.6 Tracked skeleton with a total of twenty joints [65]

The Kinect sensor can detect up to twenty joints from human's skeleton as shown in the above figure 4.6. For our study, we have used position only tracking state which only provides the information about the position of the user, but doesn't show the joints.

4.2.3 Connecting Kinect to Visual Studio

Now once the system starts collecting data. We need to create the reference of all the libraries used such as Microsoft Kinect, EmguCV, and Zxing. The program is written using C# by creating WPF file.

The algorithm is discussed in the appendix at the end of the thesis report.

4.2.3 System Workflow

First, the program is run on the laptop in Microsoft visual studio. This starts the flow; it checks if the Kinect sensor is plugged in and connected to the system. If in case this function ends with no which means that the Kinect sensor is not connected it will return with an error and exit. Provided that the decision be yes which means that the Kinect sensor is found, then with the help of SDK it will open a connection with the sensor.

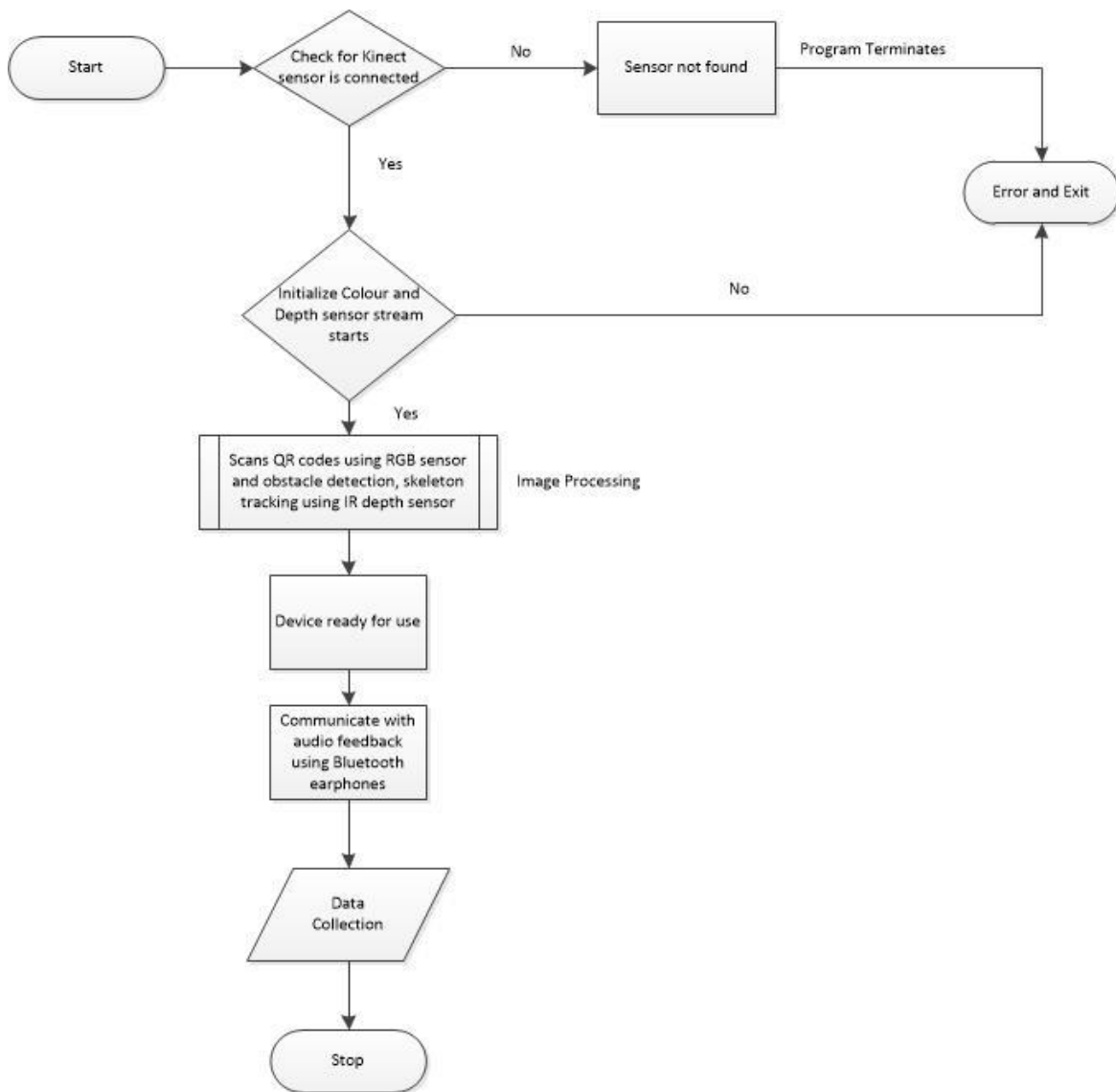


Figure 4.7 System workflow

Once the system is connected and can interact with the Kinect, it will send a signal to both RGB and IR depth sensor to start the streams for colour and depth frames as shown in figure 4.7. The RGB camera stream thus provides us with the QR code scanning for the audio output to the user which is done with the help of ZXing open source library which helps to scan the code and read the information stored. The Depth camera streams provide us with the skeleton tracking which is discussed in the above section, and it also helps to provide information about all kinds of obstacles. The image processing for the depth sensor is done using EmguCV which is an open source library which helps to find different contours, skipping objects by different width ranges, able to count people in the sensor range and take all the necessary decisions which help the user navigate. It also helps in the speech synthesis which further provides the user with the necessary audio information. Then finally both the streams are displayed on the laptop screen running at 640×480 at FPS 30 without any visible lag.

The figure 4.8, shows the flow process for the free mode of guidance. Once the user starts navigating using this mode the very first thing the program looks for is the QR code, if in case no QR code is detected by the colour camera then the depth camera will start and look for the surrounding obstacles and provide the user with auditory output necessary for navigation. And in case the camera detects the QR code then it will say the information stored using the Bluetooth earphones.

For the follow mode of guidance, before iterations were done, we assumed that there is no one interfering with the user and person being followed. In this case, the depth camera will help to locate the motion of the subject being followed using skeleton tracking and guide the user accordingly ignoring another human in the surroundings. However,

after the feedback from the blindfolded test subjects, another major iteration was done which improved the system. In this follow mode testing, instead of using the skeleton track we decided to follow a person with QR code at the back and the sensor will keep on following that QR even if there is any interference of another human which would be ignored by the system.

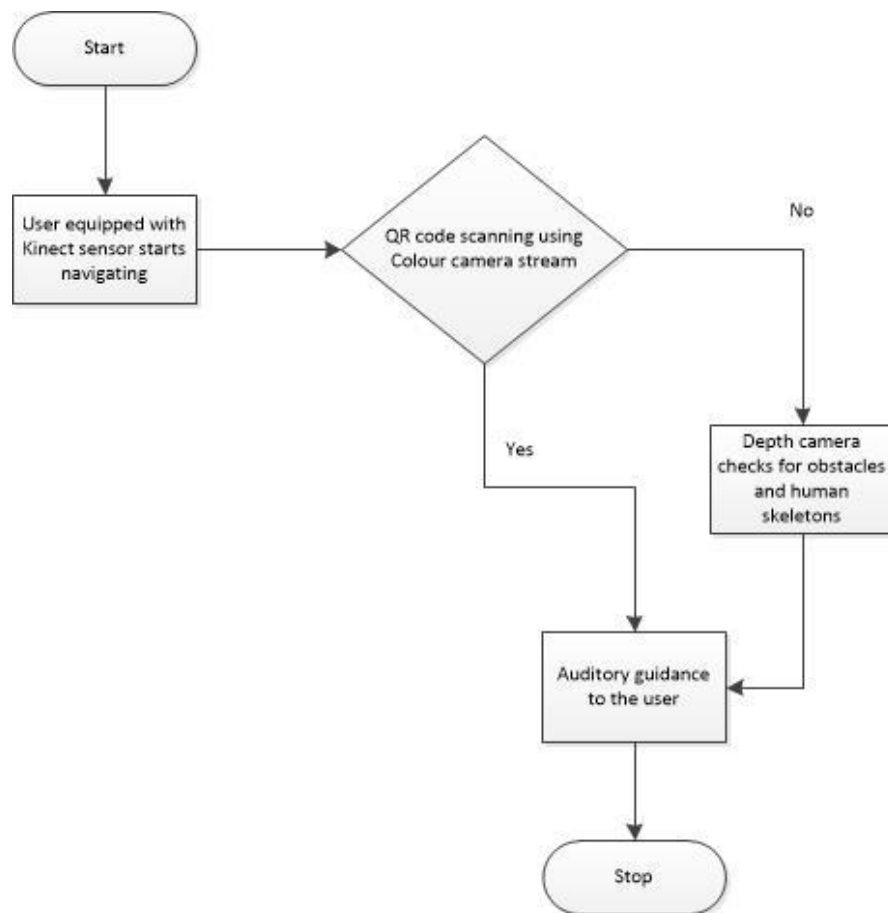


Figure 4.8 Free Mode of guidance

4.2.4 Software Interface

As discussed above the program is written in Microsoft Visual Studio 2015, using C# language as WPF application. The screen is divided into three different section displaying three different data streams as shown in figure 4.9. The first one is the colour data stream which is used for QR code scanning and better Image results. The second sections show the depth stream which gets us the distance of obstacles from the user. The third and the most important is the processed image data stream. Which helps the user to navigate based on the auditory guidance.

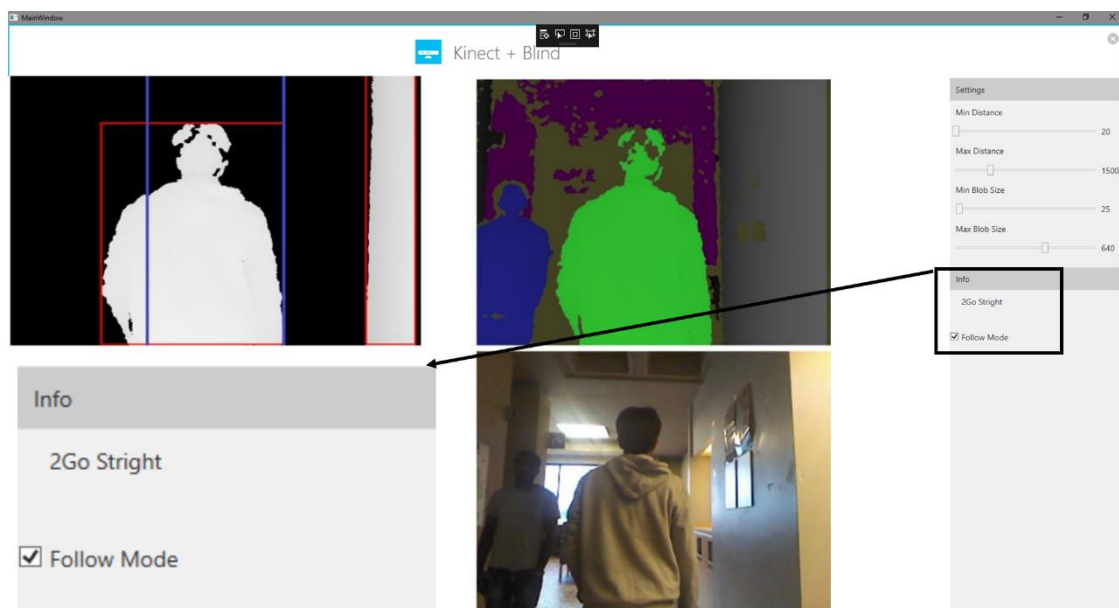


Figure 4.9 Figure: Software Interface

As we can see that the third section of the screen is further divided into three subsection columns wise, this is to make sure to help the user know which way to move. As these are three equal sections showing left, straight and right regions for navigation.

For the Image processing, the Microsoft WPF sample algorithm is used as the second project. The algorithm presented in this study is dependent on the above-mentioned project to get functions such as image viewer, joint mapping, Colour Viewer, Kinect audio viewer, Kinect depth viewer, Kinect Diagnostic viewer, and skeleton viewer.

The settings on the right column can be adjusted according to the surrounding environmental needs. The info tab shows the audio information sent to the user. As in the above figure, it shows the user going to the left, so the info bar sends audio guidance asking the person to go straight. The algorithm is programmed in such a way that in follow mode it will only track one human being which is being followed and will neglect all other passing by in the environment.

Chapter 5

Results and Discussions

5.1 System Testing

Experiments are performed indicating different scenarios. As described in the previous chapters, the Kinect sensor was used. In both the guidance modes, the decision making for visually impaired people improved and was accurate in detecting obstacles and humans. The algorithm was written in C sharp language in Visual Studio software. The data processing helps to get only the required information out of the large environment.

In the arrangement with an IR sensor, IR projector and an RGB colour camera contributed to help blind people for indoor navigation. Experiments were performed in the University building where there were not more than four humans in the surrounding areas or corridors to reduce the risk of unwanted injuries or distractions. Taking into account the assumptions during the testing the use of the system is limited in some areas.

Our system provides blind users with the ability to navigate unfamiliar indoor environments by communicating the presence of tables, chairs, other obstacles, walls, elevators, stairs, and people. A total of six users participated in this experimental study using both normal or free mode guidance and follow mode guidance. Two of the test subjects were from another university so they were unfamiliar with the test course environment while the other four being from same university. Each user carried out two

trials for the path specified for the system testing, which means a total of twelve trials were conducted for each mode. The obstacle test course is designed in such a way that the user does not get into any accident or injury. All the safety measures were taken while performing the experiments. There were few assumptions taken into account such as there would not be more than four people while navigating in corridors, QR codes were attached to most of the dead ends and labs to increase the accuracy of the system. The above-mentioned functionalities are intuitive for the user.

To determine the efficiency and accuracy of the final system, all the test objects were questioned about the physiological and psychological effect while navigating by entirely depending on the device. Moreover, they answered about the ease, and accuracy of using the assistive device. All the mentioned user experience were mainly assessed by seeing the user able to avoid the obstacles in their pathway.

The novelty in this work allows the user to follow another person for wayfinding guidance using the system freely. It communicates with the individual for obstacle avoidance, classroom or lab find, elevators, stairs to achieve proper or accurate navigation.

As the part of experimentation, the system is tested on the individuals with a blindfold (sighted people) and tests are conducted on the test course discussed in the next section. Sighted people are objectified for the testing because of higher availability. It is essential that this feature is tested as a measure of accuracy to determine the full functionality and success of our navigation device. The prototype testing and the results are further discussed in this chapter.

5.2 Test Course

The testing course consisted of various paths which are discussed below; each sighted subject was blindfolded and equipped with the device (Kinect, Laptop, portable power bank). The user starts from a lab on the second level/ floor on free mode guidance as shown in figure 5.1, and for the follow mode in figure 5.2. The lab is full of obstacles such as tables, chairs, partitions, exit door. The user was asked to make a way out of the lab avoiding all barriers and into the hallways.



Figure 5.1 User Starting point



Figure 5.2 User navigating using Follow mode

The exit door as shown in figure had QR code posted on it to help the user know about the way out as in figure 5.3 and 5.4. So, whenever the blind user is about to reach the wall or the exit door, instead of turning around mistaking it as the wall, it will scan the QR and guide the user by feedback saying exit. This is because the depth of the door

and the wall are same for the depth camera as it uses the IR projector to calculate the distance data.

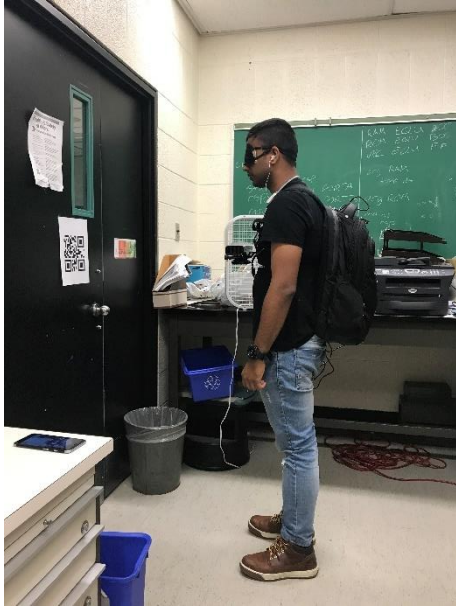


Figure 5.3 Kinect scanning QR code

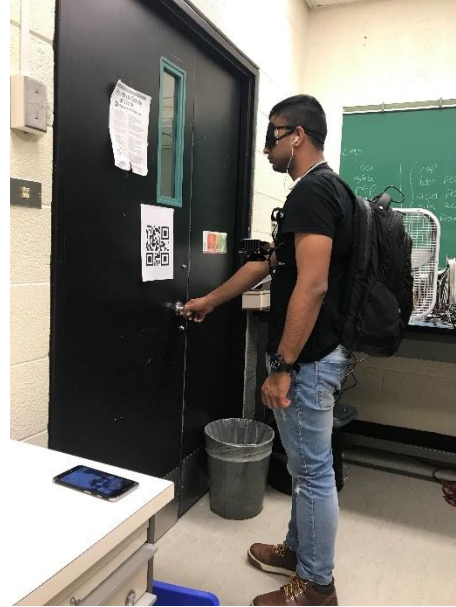


Figure 5.4 User taking exit

Once the user is out from the lab, they turn right towards the elevators and further go straight towards the stairway as shown in the figure. In the other scenario, it is assumed that all the other doors in the are open which helps the user to navigate better. As the user reached a dead end after leaving the lab on the second floor, audio information is passed over to turn right for the elevators and stairs by scanning the QR fixed onto the wall figure 5.5.



Figure 5.5 User turn right after auditory guidance from QR code scanning

The blindfolded person now makes the way through the corridor on the second level as shown in figure 5.6 for normal mode and figure 5.7 for the follow mode.



Figure 5.6 Navigating through corridors



Figure 5.7 Navigating in corridors using Follow mode

As shown in the above figure for the follow mode of guidance, that the blindfolded person is following another person for navigational inputs. However, there is another human passing by, in this case, the Kinect sensor is programmed in such a way that it would not track any other person except the one being followed. Figure 5.8 shows that the user is also provided with the information regarding the elevators on the same floor with the help of QR codes.



Figure 5.8 Kinect is providing audio information regarding elevators by scanning QR code.

According to the test course, the test object is supposed to walk through the hallway all the way to the stairs and then the QR codes which assist the user and give them feedback about the number of stairs and up or down as shown in figure 5.9 and figure 5.10.



Figure 5.9 User approaching stairs and further guided by QR code



Figure 5.10 User going up the stairs

Once the sensor sees the stairs approaching it will scan the QR codes which assist the user with feedback about the number of stairs and up or down. However, in the follow mode the blindfolded person is just following the other person as in figure 5.11.



Figure 5.11 Navigating through stairs using follow mode

From second level now the user must go down to the first level towards the destination per the designed testing course. The user can go down by using any of the two modes. For normal mode as shown in figure 5.12 in which the person receives audio guidance from scanning QR code and figure 5.13 for follow mode, where the object receives inputs as the sensor scans the current location of the person being followed.



Figure 5.12 User going downstairs after audio information form QR code



Figure 5.13 User Climbing downstairs using Follow mode

Further, the user must walk through the hallways which have many classrooms and labs on both sides indicated by QR codes as shown in the figure 5.14. As we can see in the figure below, there are some obstacles in the hallway such as garbage bins, humans. The system helps the user to avoid all those obstacles.



Figure 5.14 Navigating through hallway level 1

QR codes will further assist the user about their destination such as classroom or lab as could be seen in figure 5.15 where the user is provided with the auditory information saying the lab is on your left. All the scenario discussed here is for normal mode guidance.



Figure 5.15 User receiving information about the destination from the QR code



Figure 5.16 Destination

Figure 5.16 shows the user entering the destination. Now for the follow mode, the blindfolded person is tested on the same test course and is following a person to the destination as shown in the figure 5.17.



Figure 5. 17 Follow mode guidance

In follow mode guidance, the user will be following another person on the same path as mentioned above for normal mode. The user will just keep on following a particular person in the indoor environment. In this mode, we assume that there are not more than four individuals walking in the hallway and following that person the user will reach its destination by getting auditory feedback about every movement in the form of directional guidance.

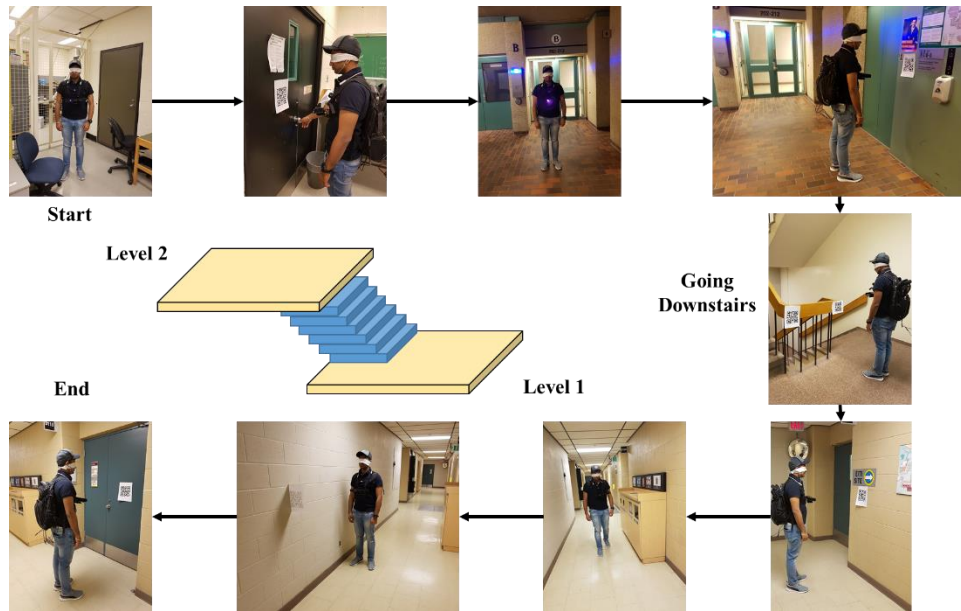


Figure 5.18 Experimental Environment for guidance

The above figure 5.18 shows the complete test course for the system testing.

5.3 Auditory Guidance

It was discovered from the literature review that auditory guidance is the best method to provide the user information. It helps the individual to create an image in the brain of the surrounding environment. The audio feedback is programmed using C sharp which processes all the available information into a narrow or precise piece of informational. This determines the user’s intuitive response to the directional audio inputs.

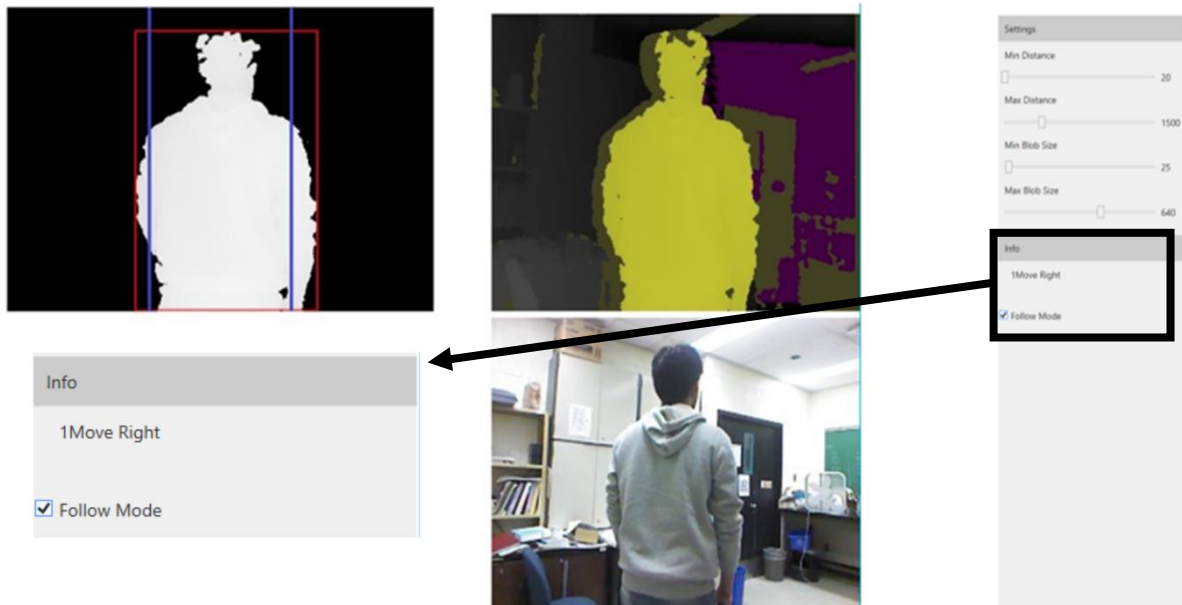


Figure 5. 19 Minimal Audio Information provided to the user

In the above figure 5.19, we can see that the audio information provided to the user by the Bluetooth earphones. The software interface shows the auditory info under the info tab. In the similar way, all the information being transferred to the user is shown under the same tab.

5.4 Results

Our System provides visually impaired users with the ability to navigate in an indoor environment. The obstacle test course as discussed above and provides an answer to the questions regarding navigation for the blind. A total of six sighted people were part of

the testing, due to easy availability and accessibility to test. This testing also helped in improving our device with further iterations. All this improved our system implementation.

As discussed in the previous chapters regarding the test course, now here in this section the respective results are discussed for the whole track. Firstly, we start from the level 2 from a lab which is full of obstacles which can be seen in the figure 5.20 below with normal mode guidance. In this mode whenever the Kinect sensor comes across any obstacle or human it will let the user know about it and guide accordingly for safe passing.

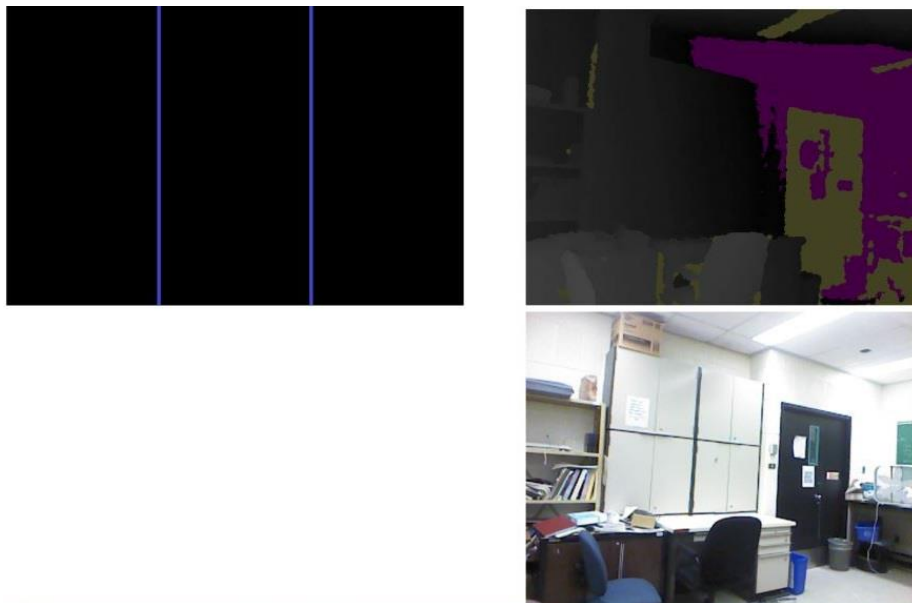


Figure 5.20 Starting point at level two

The blue vertical lines which can be seen in all the figures, tells the user in which way they to move according to the environment. Just in case if there is an obstacle or human sharing more space on the right those lines then the audio feedback will say move

left, and vice versa for moving right. And if in case the obstacle only covers the central area then it would say turn around until unless there is any QR code in the region.

Simultaneously starting from the same place and for the same testing course for the follow mode of guidance we can see the figure 5.21 below that the processed image which is on the top left shows a red box for the person being followed and provides the auditory feedback accordingly. The audio information provided to the blindfolded person works on the same concept described above using the blue lines.

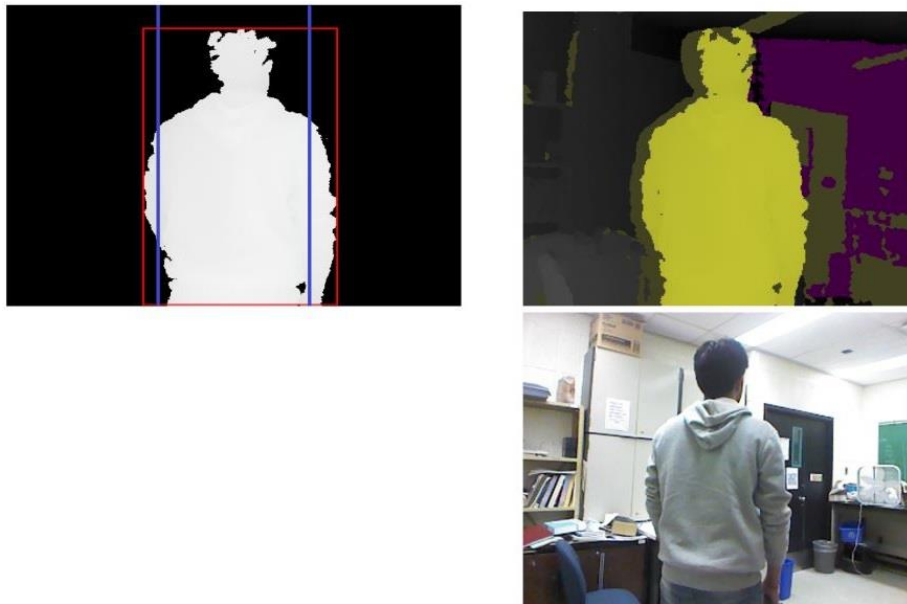


Figure 5.21 Follow mode starting point at level two

As discussed above if the Kinect sensor finds a dead end in front it will straight away let the user know to turn around only if it does not find any QR attached to that wall or partition. Now as per the testing environment the user was supposed to take the exit from the lab by avoiding all the obstacles on the way. Now since there was a QR code attached

to the door so in this case, the Kinect sensor will not inform the user to turn around or dead end, instead it would mention the data stored in that QR. In the figure 5.22 we can see the same, and when the person reaches the door, right away he is informed that this is an exit from the lab.

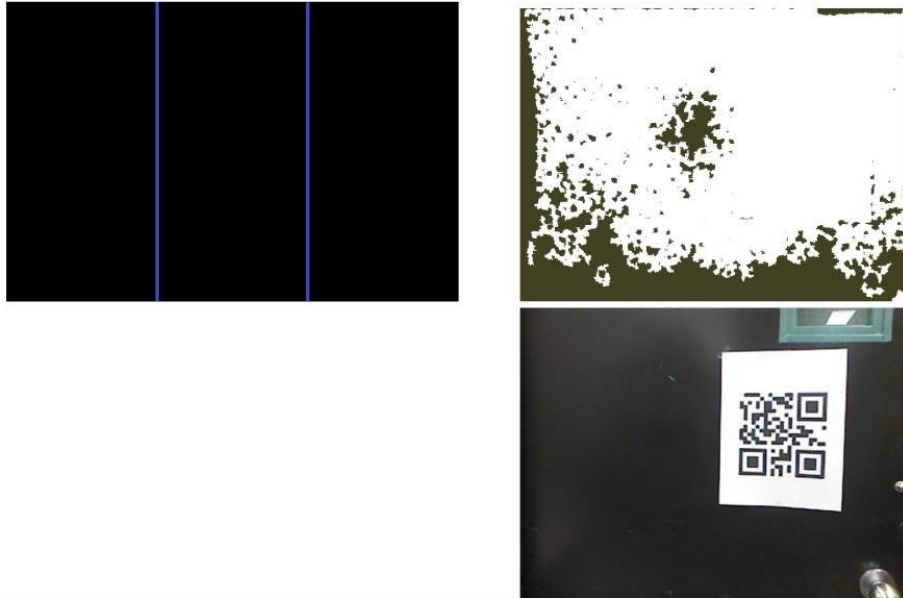


Figure 5.22 Lab Exit

Once the blindfolded person is out from the lab, he turns right in the corridor towards elevators and stairs. For guiding the user there is another QR code affixed to the wall as shown in the figure 5.23. While the red box still shows that there is a dead end but the QR direction is preferred according to the programming.

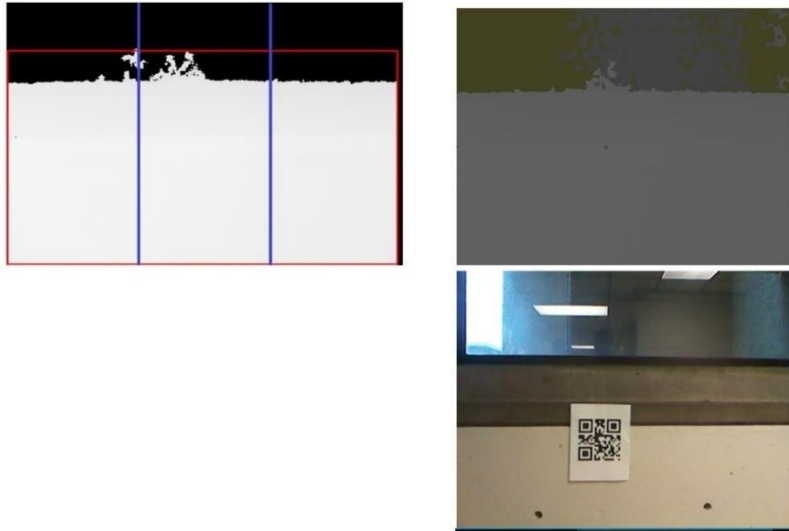


Figure 5.23 Right turn on the second floor

Similarly, the user is provided with the information regarding the directions to the elevators as shown in the figure 5.24 below.



Figure 5.24 Elevator level two

After the user turn right towards the hallway and further avoiding obstacles and humans as shown in figure 5.25 he or she moves towards the stairs going down.



Figure 5.25 Follow mode second floor

As we can see in the above image if there is another person passing by the person being followed, in this case, the Kinect sensor will not track any other person except the one being followed. The powerfully designed algorithm helps to achieve this concept, and it can be visible in the processed image which is the top left that there is no other human tracked by the sensor.

The assumptions for the hallway while navigating through the hallway is that the doors are assumed to be open during the tests being performed for the ease. The QR codes attached beside the railing on the stairway gives information for the stairs going up or down. And for the safety of the user, they are asked to use staircase railing while going downstairs.

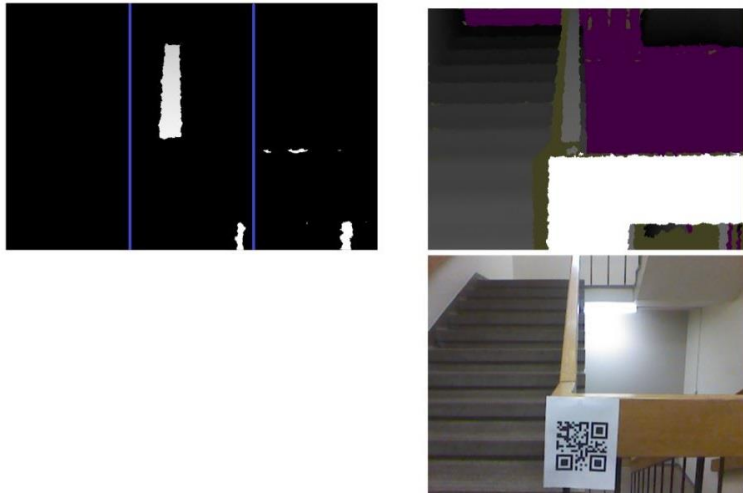


Figure 5.26 Stairs going up

Figure 5.26 and figure 5.27 show the QR codes affixed to the stairway railing that guides the user for the stairs going up or down. These help the user in normal mode guidance.



Figure 5.27 Stairs going down

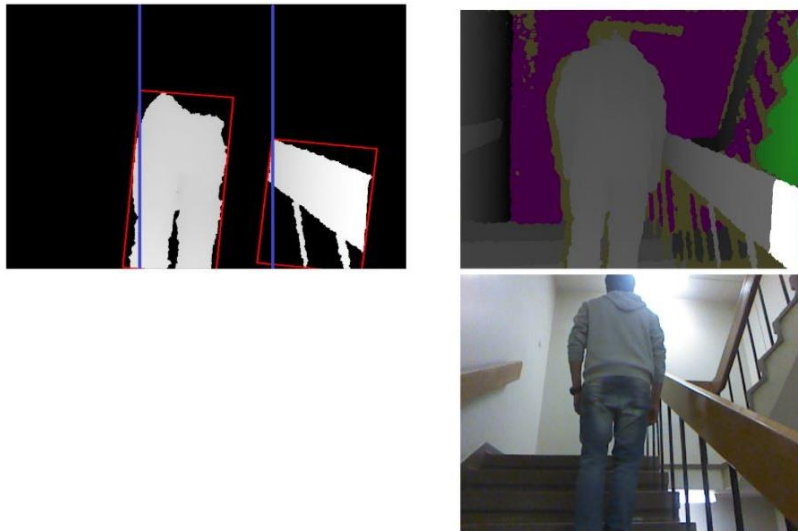


Figure 5.28 Follow mode stairs going up

For the follow mode guidance, the user follows the person using railings on the sides as shown in the figure 5.28 and figure 5.29 for stairs going up and down respectively.



Figure 5.29 Follow mode stairs down

For navigating through the stairs, special safety measures were taken. Now once the user reaches the level one, he is supposed to walk through the corridors towards the destination. Each lab is provided with a unique QR code which are attached outside all the labs on the same hallway. The user will continuously get the information from the environment as soon the sensor scans the code.

In the figure 5.30 below, we can see the user approaching a QR code on the way towards the destination according to the test course which has been discussed in the previous section.



Figure 5.30 Corridor level one

For the follow mode the sighted individual who is blindfolded just continuously follows the person to the destination as shown in the figure 5.31.



Figure 5.31 Navigating in hallway level one

In the above figure 5.31 as well we can notice another person passing by the person being followed and still the sensor is accurately able not to detect the other human and keep on following the same person. Finally, the user reaches the destination, and there is another QR that guides the user by providing the auditory information about the lab or classroom. The figure 5.32 shows the user about to enter the course end point.

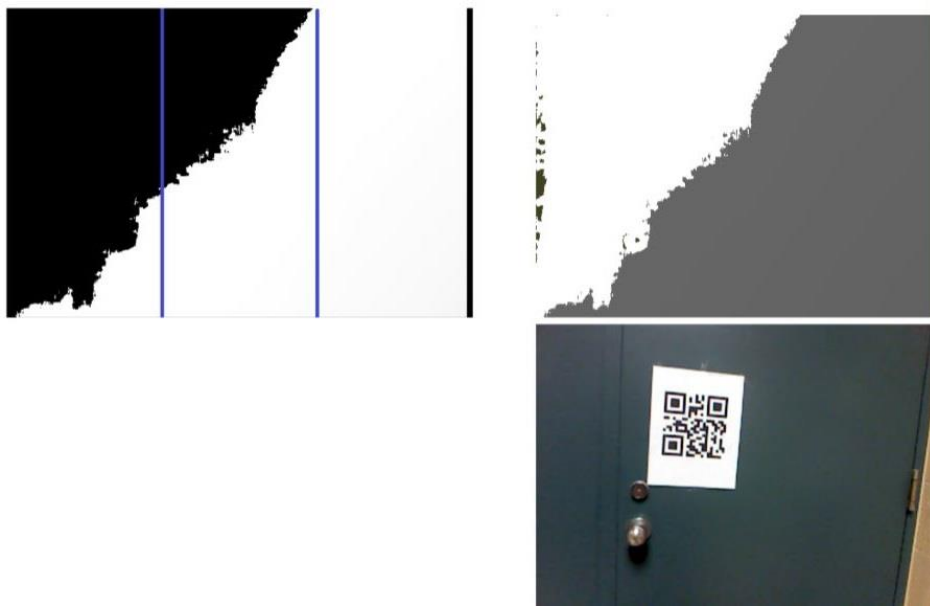


Figure 5.32 Course endpoint

The efficiency and the effectiveness of the device were measured by feedbacks from the blindfolded participants. Significantly the usefulness of the device is measured by the effectiveness to avoid the obstacles. However, there were few instances as shown in the above figure 5.33 where the user was not able to detect chairs, other obstacles, elevators and stairs. This is due to the range of the Kinect sensor and its positioning when equipped by the user, as it might not cover the obstacles outside its range and sometimes miss out on reading a QR code.

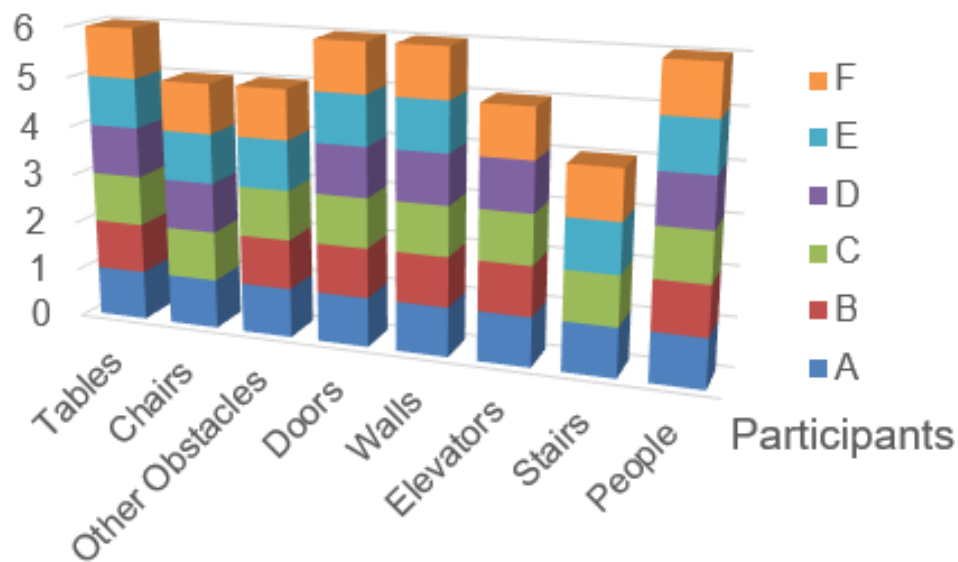


Figure 5.33 Table showing Prototype testing for blindfolded sighted participants

In order to increase the capabilities of the proposed system, we further made some algorithm changes which also changed the way Follow mode works as the part of iteration after the test results.

According to the previously discussed model in which user follows the human skeleton in the follow mode where we assumed that there is not many people around and none of them interferes during the experimentation. Now in this further iteration, the visually impaired person follows a QR code which was found to be more accurate for the follow mode of guidance. Some of the results for this iteration are shown below in figure 5.34.

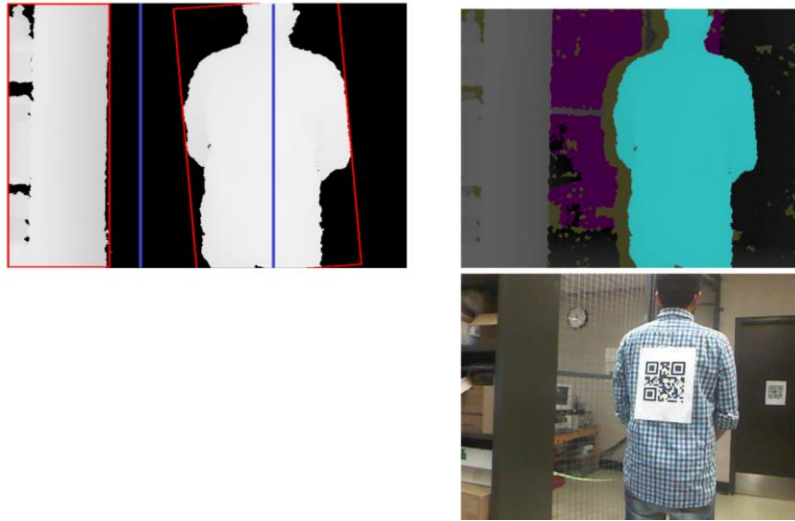


Figure 5.34 Follow mode using QR code in a lab

In the figure 5.35 below we have integrated the skeleton tracking along with the QR code scanning. So, whenever the sensor detects the QR code is saying 'Follow', it will keep on following that person even if any other person interferes or passes by.

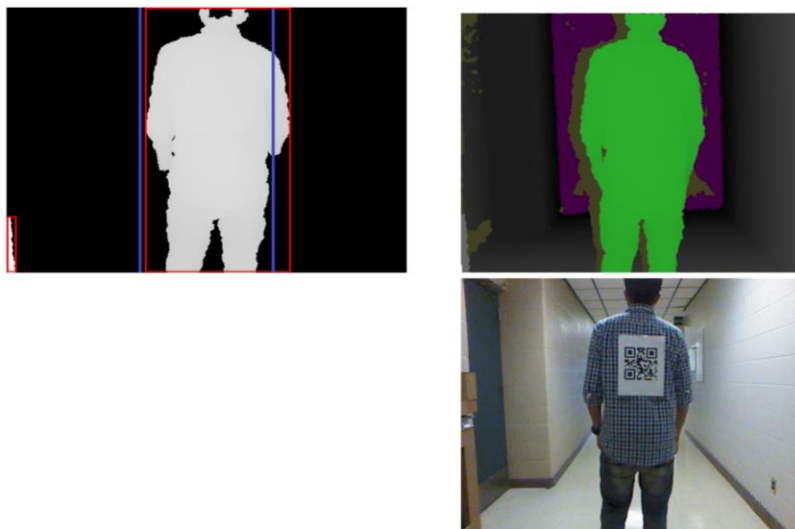


Figure 5.35 Follow mode using QR code in hallway

Further experimentation with people interfering during the testing using follow mode was found to be better with the use of QR code as shown in the figure 5.36. As the results below clearly show that the sensor is only tracking and following the person with the QR code which is highlighted with the red box. And is not tracking the other two humans.

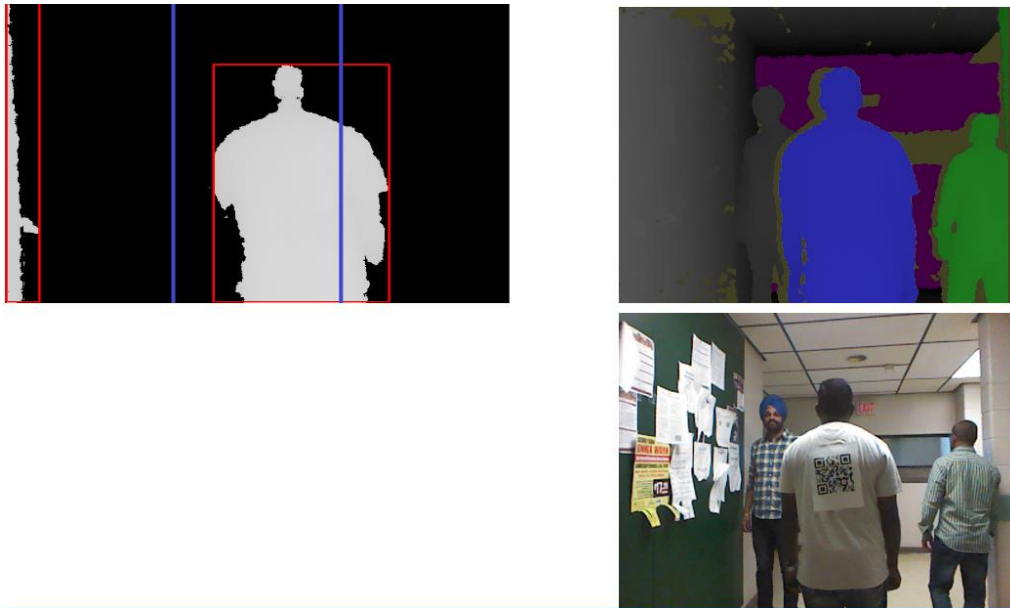


Figure 5.36 Follow mode with interference

In the other scenario shown in figure 5.37, where there is another person in between the visually impaired person and the person being followed. In this case, the sensor highlights both the humans in the range for the normal follow mode but with the iteration which was done with the use of QR for the following mode, now the sensor will guide the user to follow the person with the QR code at the back.

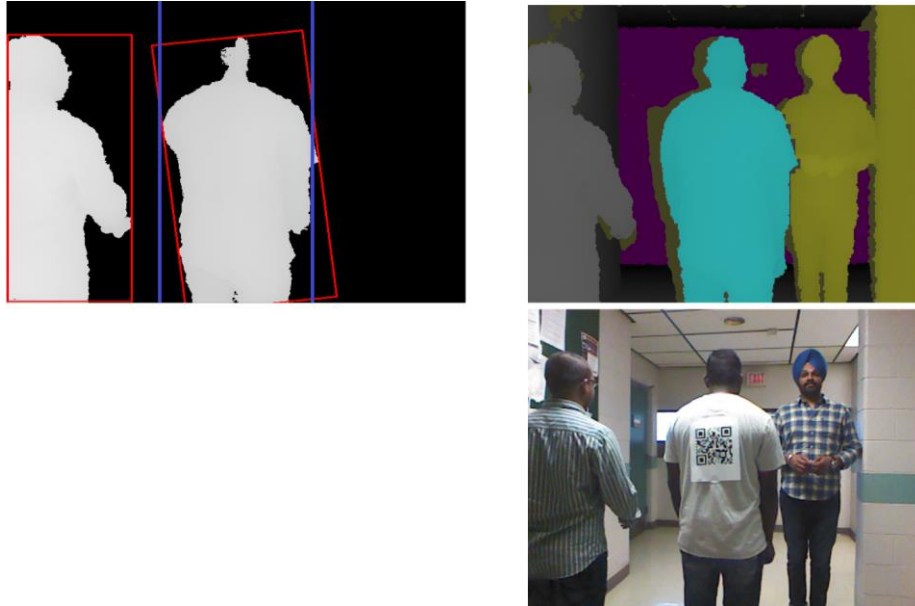


Figure 5.37 Follow mode with another person in between user and the person being followed

The accuracy of test object through the test course helps to measure the effectiveness of our device. The ability of the device in tracking all kinds of obstacles and providing auditory outputs was found to be significant as compared to the previous studies discussed in the literature review. The success of the system is how it does the obstacle detection, avoidance, scan QR codes and track humans providing audio information for the navigation purposes. The Kinect sensor helped to put up a novel approach for navigation as compared to the previously available devices in which there were more components used for single use. However, by using Kinect all major issues were covered in this study.

Chapter 6

Discussion

6.1 Future Work

Having created a successful prototype that can assist blind or visually impaired, there is further scope for improvement. The next steps are to build a more stable mount for the camera, for better viewing angles and calibration. The new version of Kinect can also be brought into use with all new technology and better quality camera. These would increase the scanning range of obstacles. There is also a possibility of using multiple Kinect sensors that might help to provide more independence while navigating and there might be few changes in the algorithm to make it more robust. The future iterations to this will keep on augmenting the use of this device. RFID technology can be another solution which might help visually impaired user to navigate through public places using RFID tag grids, canes or other various devices. This technology can be further integrated with our system for better results in outdoor areas.

6.2 Conclusion

This thesis presents an indoor navigation system using a Kinect sensor, for free and follow mode guidance. The system designed is durable, lightweight, and cost-effective, and anyone can use it with ease without any training because of its simplistic working.

The real-time image processing done with the help of computer vision algorithms helped to detect all kinds of obstacles such as tables, chairs, humans, walls, doors, and stairs. In our study, we also used QR codes for augmenting the power of our proposed system. The audio information is provided to the blind person whenever there is some obstacle or QR is scanned by the camera and even for the follow mode for directional information. The system was evaluated by six blindfolded users in both types of navigation modes. The conclusion obtained from the results clearly show the effectiveness and efficiency of our system in helping visually impaired users in the indoor environment which was based on the accuracy of the users through the test course.

References

- [1] “Fast Facts about Vision Loss. <http://www.cnib.ca/en/about/media/vision-loss/Pages/default.aspx>”
- [2] “Visual impairment and blindness.” 2014, from <http://www.who.int/mediacentre/factsheets/fs282/en/>
- [3] “University of Nebraska Medical Center.” College of Medicine Department of Ophthalmology & Visual Sciences, from <https://www.unmc.edu/eye/international/mission.html>
- [4] “Vision Loss in Canada.” 2011, from http://www.cos-sco.ca/wp-content/uploads/2012/09/VisionLossinCanada_e.pdf
- [5] T. Morton and M. Yousuf, “Technological innovations in transportation for people with disabilities workshop summary report,” 2011.
- [6] “Kinect. (n.d.). In Wikipedia. Retrieved August 10, 2004, from <https://en.wikipedia.org/wiki/Kinect>.” .
- [7] J. R. Marston and R. G. Golledge, “The hidden demand for participation in activities and travel by persons who are visually impaired,” *J. Vis. Impair. Blind.*, vol. 97, no. 8, pp. 475–488, 2003.
- [8] R. G. Golledge, J. R. Marston, and C. M. Costanzo, “Attitudes of visually impaired persons toward the use of public transportation,” *Univ. Calif. Transp. Cent.*, 1997.
- [9] J. Brabyn, “Developments in electronic aids for the blind and visually impaired,” *IEEE Eng. Med. Biol. Mag.*, vol. 4, no. 4, pp. 33–37, 1985.

- [10] J. M. Loomis, R. G. Golledge, and R. L. Klatzky, "Navigation system for the blind: Auditory display modes and guidance," *Presence Teleoperators Virtual Environ.*, vol. 7, no. 2, pp. 193–203, 1998.
- [11] J. L. González-Mora, A. Rodríguez-Hernandez, L. Rodríguez-Ramos, L. Díaz-Saco, and N. Sosa, "Development of a new space perception system for blind people, based on the creation of a virtual acoustic space," presented at the International Work-Conference on Artificial Neural Networks, 1999, pp. 321–330.
- [12] J. L. González-Mora, A. Rodríguez-Hernandez, E. Burunat, F. Martin, and M. A. Castellano, "Seeing the world by hearing: Virtual Acoustic Space (VAS) a new space perception system for blind people.," presented at the Information and Communication Technologies, 2006. ICTTA'06. 2nd, 2006, vol. 1, pp. 837–842.
- [13] A. Awada, Y. B. Issa, C. Ghannam, J. Tekli, and R. Chbeir, "Towards Digital Image Accessibility for Blind Users Via Vibrating Touch Screen: A Feasibility Test Protocol," presented at the Signal Image Technology and Internet Based Systems (SITIS), 2012 Eighth International Conference on, 2012, pp. 547–554.
- [14] S. Ram and J. Sharf, "The people sensor: a mobility aid for the visually impaired," presented at the Wearable Computers, 1998. Digest of Papers. Second International Symposium on, 1998, pp. 166–167.
- [15] A. Belov *et al.*, "A Kinect Based Indoor Navigation System for the Blind," 2015.
- [16] J. Borenstein and I. Ulrich, "The guidecane-a computerized travel aid for the active guidance of blind pedestrians," presented at the Robotics and Automation, 1997. Proceedings., 1997 IEEE International Conference on, 1997, vol. 2, pp. 1283–1288.

- [17] A. G. Dodds, "The Sonic Pathfinder: An Evaluation.," *J. Vis. Impair. Blind.*, vol. 78, no. 5, pp. 203–6, 1984.
- [18] J. Benjamin, N. Ali, and A. Schepis, "A laser cane for the blind," presented at the Proceedings of the San Diego Biomedical Symposium, 1973, vol. 12.
- [19] D. Yuan and R. Manduchi, "Dynamic environment exploration using a virtual white cane," presented at the Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, 2005, vol. 1, pp. 243–249.
- [20] S. T. Brassai, L. Bako, and L. Losonczi, "Assistive technologies for visually impaired people," *Acta Univ. Sapientiae Electr. Mech. Eng.*, vol. 3, pp. 39–50, 2011.
- [21] S. Alghamdi, R. van Schyndel, and I. Khalil, "Safe trajectory estimation at a pedestrian crossing to assist visually impaired people," presented at the Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE, 2012, pp. 5114–5117.
- [22] V. Pradeep, G. Medioni, and J. Weiland, "Robot vision for the visually impaired," presented at the Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on, 2010, pp. 15–22.
- [23] H. Jabnoun, F. Benzarti, and H. Amiri, "Visual substitution system for blind people based on SIFT description," presented at the Soft Computing and Pattern Recognition (SoCPaR), 2014 6th International Conference of, 2014, pp. 300–305.
- [24] N. Mahmud, R. Saha, R. Zafar, M. Bhuian, and S. Sarwar, "Vibration and voice operated navigation system for visually impaired person," presented at the Informatics, Electronics & Vision (ICIEV), 2014 International Conference on, 2014, pp. 1–5.

- [25] B. Sujith and V. Safeeda, "Computer Vision-Based Aid for the Visually Impaired Persons- A Survey And Proposing New Framework," *Int. J. Innov. Res. Comput. Commun. Eng.*, Oct. 2014.
- [26] A. Zeb, S. Ullah, and I. Rabbi, "Indoor vision-based auditory assistance for blind people in semi controlled environments," presented at the Image Processing Theory, Tools and Applications (IPTA), 2014 4th International Conference on, 2014, pp. 1–6.
- [27] C. K. Lakde and P. S. Prasad, "Review paper on navigation system for visually impaired people," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 4, no. 1, pp. 2278–1021, 2015.
- [28] C. K. Lakde and P. S. Prasad, "Navigation system for visually impaired people," presented at the Computation of Power, Energy Information and Commuincation (ICCPEIC), 2015 International Conference on, 2015, pp. 0093–0098.
- [29] Ş. Aymaz and T. Çavdar, "Ultrasonic Assistive Headset for visually impaired people," presented at the Telecommunications and Signal Processing (TSP), 2016 39th International Conference on, 2016, pp. 388–391.
- [30] A. Joshi, H. Agrawal, and P. Agrawal, "Simultaneous Localization and Mapping for Visually Impaired People for Outdoor Environment," presented at the Proceedings of the Second International Conference on Computer and Communication Technologies, 2016, pp. 107–115.
- [31] T. Schwarze, M. Lauer, M. Schwaab, M. Romanovas, S. Böhm, and T. Jürgensohn, "A camera-based mobility aid for visually impaired people," *KI-Künstl. Intell.*, vol. 30, no. 1, pp. 29–36, 2016.

- [32] N. Molton, S. Se, J. Brady, D. Lee, and P. Probert, "A stereo vision-based aid for the visually impaired," *Image Vis. Comput.*, vol. 16, no. 4, pp. 251–263, 1998.
- [33] A. Marsh, M. May, and M. Saarelainen, "Pharos: coupling GSM and GPS-TALK technologies to provide orientation, navigation and location-based services for the blind," presented at the Information Technology Applications in Biomedicine, 2000. Proceedings. 2000 IEEE EMBS International Conference on, 2000, pp. 38–43.
- [34] Y. Tian, X. Yang, and A. Arditi, "Computer vision-based door detection for accessibility of unfamiliar environments to blind persons," presented at the International Conference on Computers for Handicapped Persons, 2010, pp. 263–270.
- [35] F. Wong, R. Nagarajan, and S. Yaacob, "Application of stereovision in a navigation aid for blind people," presented at the Information, Communications and Signal Processing, 2003 and Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint Conference of the Fourth International Conference on, 2003, vol. 2, pp. 734–737.
- [36] S. Mann *et al.*, "Blind Navigation with a Wearable Range Camera and Vibrotactile Helmet," in *Proceedings of the 19th ACM International Conference on Multimedia*, New York, NY, USA, 2011, pp. 1325–1328.
- [37] F. Van der Heijden and P. Regtien, "Wearable navigation assistance-a tool for the blind," 2005.
- [38] G. Balakrishnan, G. Sainarayanan, R. Nagarajan, and S. Yaacob, "A stereo image processing system for visually impaired," *Int. J. Inf. Commun. Eng.*, vol. 2, no. 3, pp. 136–145, 2006.

- [39] M. Zöllner, S. Huber, H.-C. Jetter, and H. Reiterer, "NAVI—a proof-of-concept of a mobile navigational aid for visually impaired based on the microsoft kinect," presented at the IFIP Conference on Human-Computer Interaction, 2011, pp. 584–587.
- [40] B. Schauerte, M. Martinez, A. Constantinescu, and R. Stiefelhagen, "An assistive vision system for the blind that helps find lost things," *Comput. Help. People Spec. Needs*, pp. 566–572, 2012.
- [41] P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox, "RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments," *Int. J. Robot. Res.*, vol. 31, no. 5, pp. 647–663, 2012.
- [42] J. P. Bigham, C. Jayant, A. Miller, B. White, and T. Yeh, "VizWiz:: Locatelt-enabling blind people to locate objects in their environment," presented at the Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on, 2010, pp. 65–72.
- [43] Y. Tian, X. Yang, and A. Ardit, "Computer vision-based door detection for accessibility of unfamiliar environments to blind persons," *Comput. Help. People Spec. Needs*, pp. 263–270, 2010.
- [44] A. Khan, F. Moideen, J. Lopez, W. L. Khoo, and Z. Zhu, "KinDectect: Kinect Detecting Objects," in *Computers Helping People with Special Needs: 13th International Conference, ICCHP 2012, Linz, Austria, July 11-13, 2012, Proceedings, Part II*, K. Miesenberger, A. Karshmer, P. Penaz, and W. Zagler, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 588–595.
- [45] H. Takizawa, S. Yamaguchi, M. Aoyagi, N. Ezaki, and S. Mizuno, "Kinect cane: An assistive system for the visually impaired based on three-dimensional object

recognition,” presented at the System Integration (SII), 2012 IEEE/SICE International Symposium on, 2012, pp. 740–745.

- [46] S. K. Rao, A. R. Shetty, R. Bhakthavathsalam, and R. Hegde, “Stereo acoustic perception based on real time video acquisition for navigational assistance,” *ArXiv Prepr. ArXiv12081880*, 2012.
- [47] B. Prudhvi and R. Bagani, “Silicon eyes: GPS-GSM based navigation assistant for visually impaired using capacitive touch braille keypad and smart SMS facility,” presented at the Computer and Information Technology (WCCIT), 2013 World Congress on, 2013, pp. 1–3.
- [48] S. L. Hicks, I. Wilson, L. Muhammed, J. Worsfold, S. M. Downes, and C. Kennard, “A depth-based head-mounted visual display to aid navigation in partially sighted individuals,” *PloS One*, vol. 8, no. 7, p. e67695, 2013.
- [49] A. Kumar, R. Patra, M. Mahadevappa, J. Mukhopadhyay, and A. Majumdar, “An embedded system for aiding navigation of visually impaired persons,” *Curr. Sci.*, vol. 104, no. 3, pp. 302–306, 2013.
- [50] D.-M. Tsai, H. Hsu, and W.-Y. Chiu, “3-D vision-assist guidance for robots or the visually impaired,” *Ind. Robot Int. J.*, vol. 41, no. 4, pp. 351–364, 2014.
- [51] A. Helal, S. E. Moore, and B. Ramachandran, “Drishti: An integrated navigation system for visually impaired and disabled,” presented at the Wearable Computers, 2001. Proceedings. Fifth International Symposium on, 2001, pp. 149–156.
- [52] A. Dix, *Human-computer interaction*. Springer, 2009.
- [53] “Wikipedia contributors. ‘Kinect.’ Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 13 May. 2017. Web. 2 Jun. 2017.” .

- [54] "Kinect for Windows Sensor Components and Specifications." .
- [55] "Skeletal Tracking." <https://msdn.microsoft.com/en-us/library/hh973074.aspx>.
- [56] R. Lun and W. Zhao, "A survey of applications and human motion recognition with microsoft kinect," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 29, no. 05, p. 1555008, 2015.
- [57] K. Khoshelham, "Automated localization of a laser scanner in indoor environments using planar objects," presented at the Indoor Positioning and Indoor Navigation (IPIN), 2010 International Conference on, 2010, pp. 1–7.
- [58] S. Rusinkiewicz and M. Levoy, "Efficient variants of the ICP algorithm," presented at the 3-D Digital Imaging and Modeling, 2001. Proceedings. Third International Conference on, 2001, pp. 145–152.
- [59] K. Sharma, "Kinect Sensor based Object Feature Estimation in Depth Images," *Int. J. Signal Process. Image Process. Pattern Recognit.*, vol. 8, no. 12, pp. 237–246, 2015.
- [60] E. Stone and M. Skubic, "Evaluation of an inexpensive depth camera for in-home gait assessment," *J. Ambient Intell. Smart Environ.*, vol. 3, no. 4, pp. 349–361, 2011.
- [61] K. Khoshelham, "Accuracy analysis of kinect depth data," presented at the ISPRS workshop laser scanning, 2011, vol. 38, p. W12.
- [62] A. Khongma, M. Ruchanurucks, T. Koanantakool, T. Phatrapornnant, Y. Koike, and P. Rakprayoon, "Kinect Quality Enhancement for Triangular Mesh Reconstruction with a Medical Image Application," in *Soft Computing Techniques in Engineering Applications*, Springer, 2014, pp. 15–32.

- [63] "Wikipedia contributors. 'Microsoft Visual Studio.' Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 1 Jun. 2017. Web. 2 Jun. 2017." .
- [64] A. Jana, *Kinect for Windows SDK Programming Guide*. 2012.
- [65] "<https://msdn.microsoft.com/en-us/library/hh973074.aspx>." .
- [66] "Emgu CV." http://www.emgu.com/wiki/index.php/Main_Page

Appendix

I. Namespace

Using keyword is used to import various namespace, i.e. collection of classes. Namespaces are used to organise objects. For this study, various namespaces are referred and imported in the WPF to create the algorithm. All the used namespace are given below;

```
using System;  
using System.Windows;  
using System.Windows.Controls;
```

All the fundamental classes and base classes that define commonly used value and reference data types, events and event handlers, interfaces, attributes, and processing exceptions are part of **System** Namespace. **System.Windows** provides several important Windows Presentation Foundation (WPF) base element classes, various classes that support the WPF property system and event logic, and other types that are more broadly consumed by the WPF core and framework. **System.Windows.Controls** provides classes to create elements, known as controls, that enable a user to interact with an application. The control classes are at the core of the user's experience with any application because they allow a user to view, select, or enter data or other information.

```
using System.Windows.Media;  
using System.Windows.Media.Imaging;  
using System.Drawing;  
using Microsoft.Kinect;  
using Emgu.CV;  
using Emgu.CV.Structure;  
using System.Speech.Synthesis;
```

System.Windows.Media provides types that enable integration of rich media, including drawings, text, and audio/video content in Windows Presentation Foundation (WPF) applications. **System.Windows.Media.Imaging** provides types that are used to encode and decode bitmap images. **System.Drawing** The System.Drawing namespace provides access to GDI + basic graphics functionality. More advanced functionality is provided in the System.Drawing.Drawing2D, System.Drawing.Imaging, and System.Drawing.Text namespaces. **Microsoft.Kinect** The Kinect Services support the following features: Depth image including Player Index, RGB image, Tilt (Get and Set), Microphone Array (not in simulation), Skeleton Tracking (not in simulation)

Emgu.CV Emgu CV is a cross platform .Net wrapper to the OpenCV image processing library [66]. Allowing OpenCV functions to be called from .NET compatible languages such as C#, VB, VC++, IronPython etc. **Emgu.CV.Structure** Defines colors, circle, 2D, Dimensions, Ellipse, Hue, Saturation, and every other parameter regarding the structure, etc. The **System.Speech.Synthesis** namespace contains classes for initializing and configuring a speech synthesis engine, for creating prompts, for generating speech, for responding to events, and for modifying voice characteristics.

II. Initializing Kinect Streams and Bitmap

Firstly, the Kinect sensor is started by running the algorithm in Visual studio. If the sensor is not connected, the program will return error.

```
byte[] colorPixels;  
  
int blobCount = 0;  
  
public MainWindow()  
{  
    InitializeComponent();
```

```

        this.Loaded += MainWindow_Loaded;
        this.Closing += MainWindow_Closing;
        speechSynthesizerObj.SpeakCompleted +=
SpeechSynthesizerObj_SpeakCompleted;
    }
    bool isspeakcompleted = true;
    private void SpeechSynthesizerObj_SpeakCompleted(object sender,
SpeakCompletedEventArgs e)
    {
        isspeakcompleted = true;
    }

    void MainWindow_Loaded(object sender, RoutedEventArgs e)
    {
        kinectSensorChooser1.KinectSensorChanged += new
DependencyPropertyChangedEventHandler(kinectSensorChooser1_KinectSensorChan
ged);
    }

    void StopKinect(KinectSensor sensor)
    {
        if (sensor != null)
        {
            sensor.Stop();
            sensor.AudioSource.Stop();
        }
    }

    Void kinectSensorChooser1_KinectSensorChanged(object sender,
DependencyPropertyChangedEventArgs e)

```

III. Enabling Depth Streams and Bitmap

The following code initializes the Depth and Colour streams for the Kinect sensor.

The depth stream runs the resolution 640×480 at Fps 30.

```

KinectSensor oldSensor = (KinectSensor)e.OldValue;
    StopKinect(oldSensor);

KinectSensor newSensor = (KinectSensor)e.NewValue;

```

```

newSensor.SkeletonStream.Enable();

newSensor.DepthStream.Enable(DepthImageFormat.Resolution640x480Fps30);

newSensor.ColorStream.Enable(ColorImageFormat.RgbResolution640x480Fps30);
    this.colorPixels = new byte[newSensor.ColorStream.FramePixelFormatDataLength];
    this.depthPixels = new
DepthImagePixelFormat[newSensor.DepthStream.FramePixelFormatDataLength];
    this.colorBitmap = new WriteableBitmap(newSensor.ColorStream.FrameWidth,
newSensor.ColorStream.FrameHeight, 96.0, 96.0, PixelFormats.Bgr32, null);
    this.depthBitmap = new WriteableBitmap(newSensor.DepthStream.FrameWidth,
newSensor.DepthStream.FrameHeight, 96.0, 96.0, PixelFormats.Bgr32, null);
    this.colorImg.Source = this.colorBitmap;
    newSensor.AllFramesReady += this.sensor_AllFramesReady;

    try
    { newSensor.Start(); }
    catch (System.IO.IOException)
    { kinectSensorChooser1.AppConflictOccurred(); }
}

```

IV. Finding QR code

The following code is used to identify the QR codes from the colour streams and further an auditory output is sent to the user which says the text which is stored in the code. QR ZXing library is used for decoding the codes within colour images.

```

public string QRText = String.Empty;
String objectPosttion = "";
    SpeechSynthesizer speechSynthesizerObj = new SpeechSynthesizer();
private void sensor_AllFramesReady(object sender, AllFramesReadyEventArgs e)
{
    BitmapSource depthBmp = null;
    blobCount = 0;
    using (ColorImageFrame colorFrame = e.OpenColorImageFrame())
    {
        QRText = String.Empty;

        if (colorFrame != null)

```

The Kinect colour frame data is copied to bytes and further this data is copied into writable bitmap.

```

    {
        colorFrame.CopyPixelDataTo(this.colorPixels);
        this.colorBitmap.WritePixels(
new Int32Rect(0, 0,this.colorBitmap.PixelWidth,this.colorBitmap.PixelHeight),
        this.colorPixels,
        this.colorBitmap.PixelWidth * sizeof(int),0);
        using (Bitmap bm = colorBitmap.ToBitmap())
        {
            bm.RotateFlip(RotateFlipType.RotateNoneFlipX);
            QRText = ImageHelpers.DecodeQRZxing(bm);

            lblQR.Content = QRText;
        }
    }
}
using (DepthImageFrame depthFrame = e.OpenDepthImageFrame())
{
    if (depthFrame != null & QRText == string.Empty)

```

According to the code below, we can find objects and obstacles from particular depth ranges

```

    {
        blobCount = 0;
        depthBmp =
depthFrame.SliceDepthImage((int)sliderMin.Value, (int)sliderMax.Value);

        Image<Bgr, Byte> openCVImg = new Image<Bgr,
byte>(depthBmp.ToBitmap());
        Image<Gray, byte> gray_image = openCVImg.Convert<Gray, byte>();
        objectPosttion = "";
        using (MemStorage stor = new MemStorage())

```

Emgu.CV cross platform helps in image processing the data streams by using OpenCV image processing library. This is used to find different contours, skipping objects by different width ranges, able to count people in the sensor range and take all the necessary decisions which help the user navigate.

```

{
Contour<System.Drawing.Point> contours = gray_image.FindContours(
Emgu.CV.CvEnum.CHAIN_APPROX_METHOD.CV_CHAIN_APPROX_SIMPLE,
Emgu.CV.CvEnum.RETR_TYPE.CV_RETR_EXTERNAL,stor);

    for (int i = 0; contours != null; contours = contours.HNext)
    {
        i++;
if ((contours.Area > Math.Pow(sliderMinSize.Value, 2)) && (contours.Area <
Math.Pow(sliderMaxSize.Value, 2)))
    {
        int pe = 0;
        if (kinectSkeletonViewer1 != null)
            pe = kinectSkeletonViewer1.getPeoples();
        if (pe == 0 && !isFollowMode)
        {
            MCvBox2D box = contours.GetMinAreaRect();
            int x = contours.BoundingRectangle.X;
            int size = contours.BoundingRectangle.Width;
            if (size > 426)
            {
                objectPosition = "Obstacle Detected Turn Around";
            }

            else if (size > 213 & x > 213 & x < 426)
            {
                objectPosition = "Obstacle Detected Move Right";
            }

            else if (size > 213 & x < 213 & x < 426)
            {
                objectPosition = "Obstacle Detected Move Left";
            }
            else

```

The above code helps the user to get directional feedback. For this the pixel data of Kinect sensor is divided into three different ranges, which say feedback such as, obstacle detected turn around, move right, move left go straight.

```

    {
        }
        openCVImg.Draw(box, new Bgr(System.Drawing.Color.Red), 2);
        blobCount++;
    }

```

For the follow mode guidance, the user will follow the nearest person and neglect all other people out there in the surroundings. And for this study it is assumed that

```

    else if (isFollowMode && pe > 0)
    {
        MCvBox2D box = contours.GetMinAreaRect();
        int x = contours.BoundingRectangle.X;

```

These lines of the code help to find the position of the person being followed.

```

        int size = contours.BoundingRectangle.Width;

        if (size < 213)
        {
            objectPosttion = "Go Straight";
        }
        else if (size > 213 & x > 213 & x < 426)
        {
            objectPosition = "Move Left";
        }
        else if (size > 213 & x < 213 & x < 426)
        {
            objectPosition = "Move Right";
        }
        else
        {
        }
        openCVImg.Draw(box, new Bgr(System.Drawing.Color.Red), 2);
        blobCount++;
    }

```

These lines of the code help differentiate between person and object.

```

        else if (pe > 0)
        {
            blobCount = pe;
            objectPosttion = " person in front of you ";

```

```

    }
  }
}

```

```

this.outImg.Source = ImageHelpers.ToBitmapSource(openCVImg);
txtBlobCount.Text = blobCount.ToString() + "" + objectPosttion;

```

These lines of the codes help by sending all the text of speech Synthesizer which sends the audio information to the user through earphones. And it also asynchronously speaks the contents present in the object position.

```

if (objectPosttion != "" && isspeekcompleted)
{
    isspeekcompleted = false;

    speechSynthesizerObj.SpeakAsync(objectPosttion);
}
else
{
    if (isspeekcompleted)
    {
        isspeekcompleted = false;
    }
}

```

It asynchronously speaks the contents stored in individual QR code.

```

        speechSynthesizerObj.SpeakAsync(QRText);
    }
}

}
}
void MainWindow_Closing(object sender, System.ComponentModel.CancelEventArgs)
{
    if (null != this.sensor)
    {

```

```
        this.sensor.Stop();
    }
}

private void CloseBtnClick(object sender, RoutedEventArgs e)
{
    this.Close();
}
public bool isFollowMode = false;
private void CheckBox_Checked(object sender, RoutedEventArgs e)
{
```

This assigns the values for the follow mode of guidance.

```
    isFollowMode = ((CheckBox)sender).IsChecked.Value;
}
}
}
```