Navigational System Analysis Using Image Processing for Visually Impaired

Thesis submitted in partial fulfilment of the requirement for the degree of

Bachelor of Science
in
Computer Science

Under the Supervision of

Dr. Md. Haider Ali

By

Naireet Chakravorty (15341028),

Nazia Afroz(12101025),

Dipayan Sen Gupta (15241002)

School of Engineering & Computer Science

Department of Computer Science & Engineering

BRAC University

DECLARATION

We, hereby, declare that the work presented in this Thesis Project is the outcome of the investigation performed by us under the supervision of Dr. Md. Haider Ali as Supervisor, Department of Computer Science Engineering, BRAC University Bangladesh. We also declare that no part of this Thesis and thereof has been or is being submitted elsewhere for the award of any degree or Diploma.

....

(Dr. Md. Haider Ali)

Supervisor

Department of CSE

BRAC University

Signature of authors:

______________________	______________________	______________________
Naireet Chakravorty	Nazia Afroz	Dipayan Sen Gupta
15341028	12101025	15241002

Abstract

Blindness is a severe disease that limits a person's ability to navigate their surroundings properly. Usually using a crane has been the only solution for blind people so far, which is a solution not suitable to detect objects. In our thesis we will be analyzing image processing and computer science to come up with the most practical solution to encounter this situation. We will be using various algorithms to determine which one allows the best navigation system with best efficiency at the least expense. The analysis will also provide audio feedback to round off a complete autonomous system that can be easily used by a visually impaired person.

Acknowledgement

We would like to take the opportunity to express our gratitude to respected supervisor Dr. Md. Haider Ali sir for his patience, support and guidance patience assistance. He not only gave us time but also proper guidance and valuable advice whenever we faced with some difficulties. His comments and guidance helped us in preparing our thesis report.

We would also like to thank BRAC University for giving us the opportunity to complete our BSc. Degree in Computer Science.

We would also like to thank our parents and family for their support and encouragement.

Table of contents:

List of figures:

4.4 Line outside safe zone

List of tables:

Abbreviations:

GPS: Global Positioning System.

HRTF: Head Related Transfer Function.

OS: Operating System

CV: Computer Vision

RGB – Red Green Blue (Color Model)

CCS - Cartesian Co-ordinate System

HT – Hough Transform

SHT- Standard Hough Transform

PHT – Probabilistic Hough Transform

Chapter-1

Introduction:

Blindness is a common disease all around the world, almost 285 million [1] are blind all over the world. Even though science has come up with a lot of advancements for blind people to aid in their daily life. There are still not enough optimal solutions for all the problems that the blind people face. There are over 7, 50,000 blind people in Bangladesh. Vision is the most important sense needed to navigate the surroundings hence, navigation and traversal is one of the hardest actions to perform for a blind person. They find going anywhere to be a really difficult task. The traditional solution has always been to use a crane to navigate around the world but not only is that not an intuitive solution, it also doesn't tell if there is an object or obstacle ahead of the person holding the crane.

Many solutions have been proposed already, one of them is GuideCane [2]. This is a device that uses echolocation to detect objects directly in front of the device. But in reality it doesn't really offer significant advantage over the traditional crane model and also disrupts the personal navigation method one might have developed internally.

Another solution that was proposed named Drishti [3], takes the echolocation principle and takes it a step forward using the ultrasound for identification of objects indoors and for outdoor navigation it uses GPS. The drawback is that the device can only identify where potential objects are but not where a visually impaired person can navigate around. GPS has also shown to have an error up to 9.14 meters. Which is insufficient to guide its user away from a potential object. Even though the Drishti system can warn the user of hazards, the limitations has affected it from being adopted by visually impaired people.

Our system addresses this problem from a new perspective, using a smart phone camera as it is really common nowadays. Smart phones have gotten dramatically cheaper as well. So we have chosen to utilize the wide availability to develop a system to tackle this situation.

Proposed System:

In theory, it is very simple to detect a line in an image based on radon transforms. But in most real life cases this does not yield us an accurate result. We are aiming to find a solution that can find this lanes as accurately as possible in the least amount of time. For this we will be using real time image grabbing. As we grab the images we will apply probabilistic Hough line transformation and simple Hough transformation to them. Then we perform a bitwise addition of the resultant images to get a more accurate image.

Using this data the application will send an audio cue to the user asking him to move ahead.

Image Processing:

Image processing is usually referred to manipulation of images using a variety of mathematical operations. In computer science image processing usually refers to digital image processing. Image processing is a huge field in CS that can be used for satellite photos, character recognition and many other fields.

Computer Vision:

Computer vision is the method of acquiring, processing, analyzing and understanding images in high dimensional data from the real world and produce relevant numerical or symbolic information. It is also able to automate and integrate a wide range of processes and representations used in vision perception. In computer science, the subject is concerned with the theory behind the artificial systems that can be extracted from images.

Existing image processing libraries:

For our thesis we will be looking at a few image processing libraries that are available to use. There are a few image processing libraries that can perform a wide range of complex image manipulations. These libraries are:

- **OpenCV:**OpenCV is short for Open source computer vision is an image processing library mainly aimed at real time computer vision. This library is cross platform, so works on a lot of platforms and also free to use under the open source license[1].
- **VXL**: Vision-something-library also known as VXL is a collection of open source C++ library that can tweaked to serve different purposes of computer vision[2].
- **ImageJ:**ImageJ is a public domain Java based image processing library. Users have to write plugins to solve image processing problems using this library[3].

[1] http://opencv.org
[2] http://vxl.sourceforge.net
[3] http://imagej.net/Welcome

Thesis outline:

Our thesis will have three segments, these will be:

- **Literature Review:** This segment will focus on previous works done in the field by other people. We will look at how they approached the situation, what they came up with as the solution and how they solved it. We will also look into what has been done in the fields of lane detection and image segmentation as of now.
- **Technical synopsis:** This part of our thesis will delve into the theoretical aspects of our application. We will talk about everything involving image processing from RGB scale, grayscale images to the more advanced image manipulation theory such as Hough transformation, Sobel edge detection and canny edge detection. This segment will also describe why we are using these particular techniques and how we intend to implement them in our final application.
- **System design:** The last segment will be totally based on how we are developing a system using the aforementioned research that we did. We will talking about the algorithms in use, tools we are using and how we are collecting data. Ultimately leading to implementing the system.

Chapter-2: Literature Review

This chapter reviews existing navigational assistance for the visually disabled people. Firstly mobility of vision disabled people is presented in section 2.1 .Then conventional mobility tools are described in section 2.2. Next associative technologies are reviewed in section 2.3. Finally the content of the Chapter is summarized in section 2.4.

2.1) Mobility of the Visually Impaired:

Mobility is the ability to travel safely from one place to another place and is an important aspect of life. Vision plays a very significant role in mobility as it allows sighted people to collect most of the information required for perceiving the surrounding environments. People with vision loss must depend on some other senses (hearing and touch) to gather information of the surrounding objects and therefore face great difficulties in traveling. The mobility of blind people includes two aspects: perception and orientation as shown in figure 2.1[4].

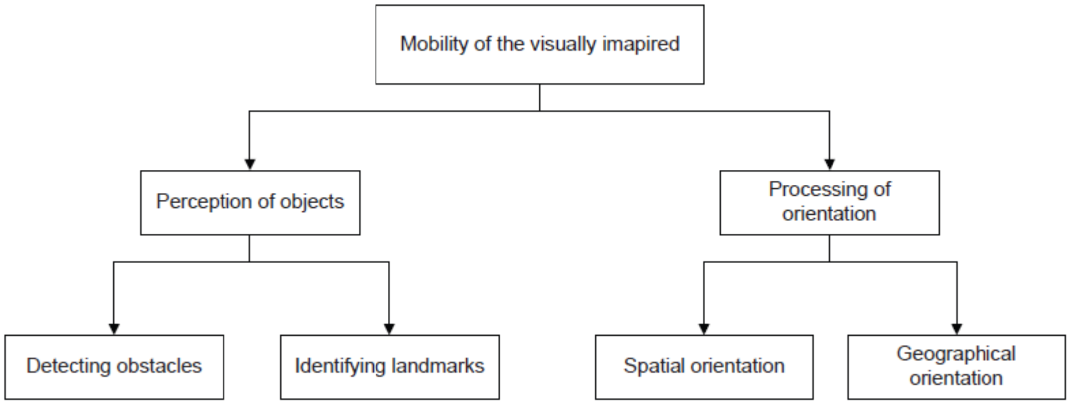


Figure 2.1: Brambing’s model for mobility of blind people.

Perception refers to the sense of detecting obstacles and landmarks without vision. Orientation refers to the knowledge to recognize the position in lieu of surrounding objects and the location in the route of entire journey. To understand the surrounding, vision impaired people use touching and listening as primary modalities. Blind people employ many different types of landmarks to determine their position in the environment. Examples include rises and fall in the walking path, changes in the texture of the walking path, the presence of walls and hedges, traffic sounds and temperature changes.

People with vision impairment has three major barriers in travelling.

1) Relying mainly on touching and hearing, visually impaired people are limited in detecting and avoiding obstacles, finding travel path and identifying early hazard ahead.
2) They have great difficulties in determining routes of a journey, understanding the scene layout and identifying their position in the scene.
3) They cannot obtain visual or textual information such as road signs or bus numbers. These barriers make visually impaired people unable to move safely and independently in surrounding environment.

2.2) Traditional aids:

There are two popular ways of navigation for blind people. One of them is using a white cane and the other is to use a guide dog.

The White Cane:

This is a tool that is generally made out of wood or aluminum. It's a lightweight cylindrical cane that is adjustable according to user's height. In order to use, the user swings the cane directly in action with the ground. User get a feedback from the vibration made from the cane and by that the user get a realization of what is in front of them. Blind people have to go through a training of 100 hours to get their sense right, of the vibrations coming from the cane’s impact on the ground. The disadvantages of using a white cane is that if user uses it for a long time then it causes arm fatigue and muscle weakness. Again if user uses it for a long period of time, bad alignment of spine in using it would cause severe back pain.

The Guide Dog:

A guide dog is a specially trained dog that is specially trained for the mobility of the blind people. The dog generally walks on the left side a couple distance ahead of the user. The dog stops at all kind of curbs at its path. It helps in taking the user to any transport. It helps the user in crossing road and also it remember the familiar routes. The perks of a guide dog is that a full trained guide dog generally costs 40,000 USD. The dog then only survive a time span of 8-10 years only. Due to high cost of training and maintainability only 3 percent of visually impaired population are reported to use guard dog.

2.3) Technology aids:

Many associative technology has been developed in improving the mobility of visually disabled people. These systems can be categorized in 3 parts. Electronic obstacle detectors, computer vision based system and GPS based system.

2.3.1) Electronic Obstacle Detectors:

Obstacle detectors are small electronic devices that can be attached with white canes or glasses. This devices apply their electro location principle of emitted signals for detecting objects. Based on the type of employed signals they are classified into laser detectors and ultra-sound detectors.

Laser Detectors:

These devices are based on Cranberg's principle of optical Triangulation. These devices emit infrared light pulses from the transmitter and this pulses reflects back from the transmitter when they meet obstacles. The receivers measure the angle of the reflected pulses in order to compute the distances to obstacles. Examples of laser detectors are the C-5 laser cane and the talking laser Cane. The C5 laser cane employs three beams of infrared lights to detectObstacles in upward, for ward and downward directions and a proximate range of4 m as shown in Figure. 2.2.

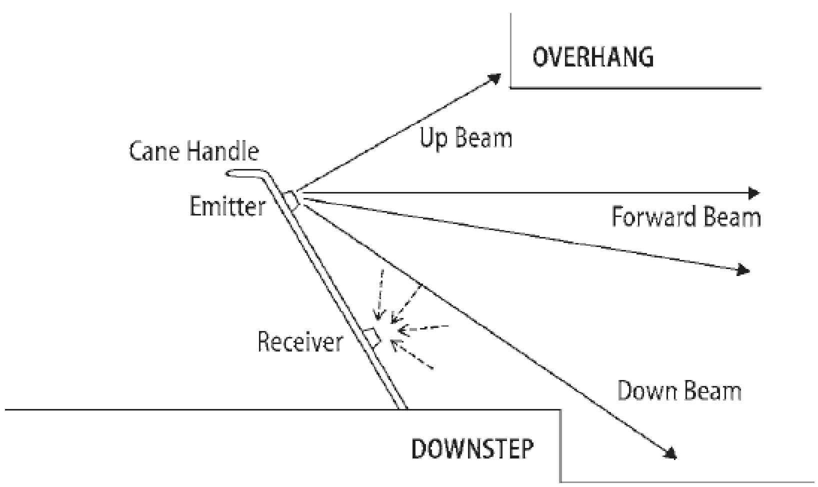


Figure 2.2: Beam Geometry of Laser cane.

Laser objects have several obstacles .First of all they cannot detect object that are transparent or non-reflective. Secondly the lasers employ short laser pulses thus detecting and navigating only on a short proximate distance.

Ultrasound Devices:

These devices are made of ultra-sonic devices. When ultra-sonic beams meet a device it reflects back. The time period between emitting and receiving the ultra sound beam is measured to compute the distance to the obstacle.

(a) (b) (c) (d)

Fig 2.3: Examples of ultrasound devices: (a) Sonicguide [5], (b) Miniguide [6], (c) Ultra-Cane, (d) Nav Belt [7].

Systems	Number of sensors	Maximum distance	Orientation aid
Sonic Guide[5]	1	6	No
Mini Guide[6]	1	8	No
Ultra-Cane	2	4	No
Nev Belt[7]	8	5	Yes
Guide Cane	10-16	5	Yes

Table 2.1: Comparison of ultrasound detectors.

The Sonic guide consists of an ultrasonic transmitter and two ultra sound receivers embedded into eyeglasses as shown in the figure. This image is designed to capture sonic image of the environment. The sonic image represents the distances to the obstacles by sound tone of different frequencies, it also describes the direction of object through the bin aural headphones. The mini guide includes an ultra sound transmitter and Revere. The ultra-cane is a combination of ultrasonic sensors and a cane. The Nav-belt employs eight ultrasound sensors integrated into a belt. The Guide Cane is designed to hold as cane .This Cane is made of 10-16 ultra sound sensors. This machine changes it path upon detecting any object.

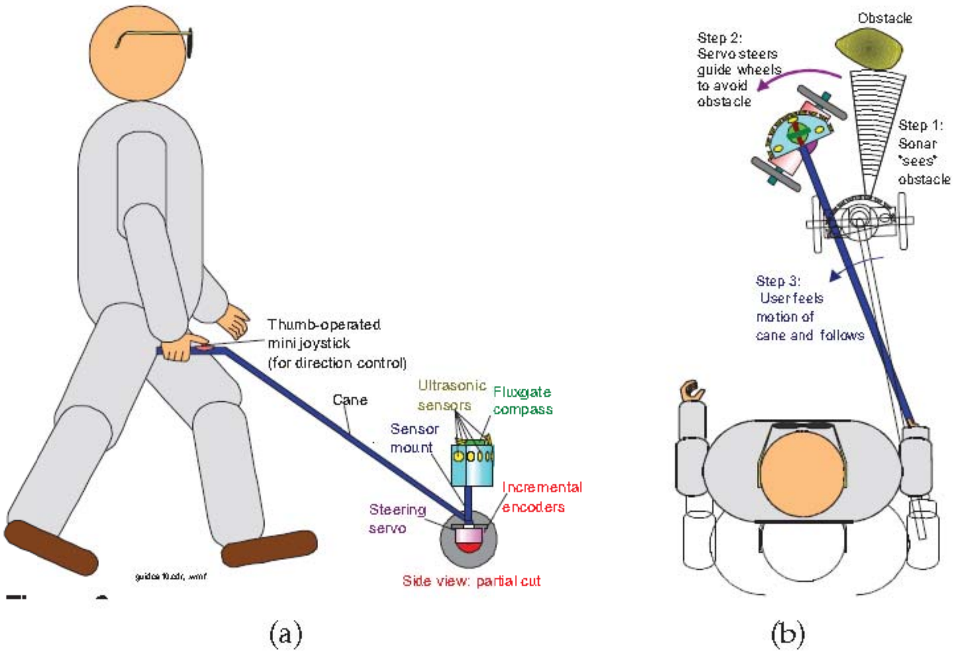


Figure2.4: The Guide Cane Prototype [7] (a) schematic (b) operation.

Ultrasonic devices are effective in use before hard objects. Their efficiency tend to decrease against soft objects where soft objects absorb their ultrasonic waves. These devices also employ a line of sight propagation and therefor for crowded environments, navigation is affected significantly.

2.3.2) GPS Based System:

Many assisting technology s combining with global positioning system and geographic system assist the blind people in determining routes, landmarks, locations during their journey. Examples of GPS based system for visually impaired include personal guide system (PGS) [8], MoBIC system [9], Trekker Brezze [10], BrailleNote GPS [11] and Street Talk [12].

Systems	Interface and components	Advantages	Disadvantages
Personal Guide System (PGS) [8]	A GPS receiver, compass, computer and headphones	Portable; determining routes, landmarks and lo-cations; providing travel instructions	Separated components and not convenient for users; blocking the user's Hearing.
MoBIC sys-tem [9]	A computer, a GPS receiver, a compass nd mobile telecommunications facilities	Allowing users to plan a journey; providing the in-formation of surrounding environments and travel Instructions.	Separated components and not convenient for users; blocking the user's Hearing.
Trekker Breeze [10]	A PDA integrated with an internal GPS receiver and an internal peaker	Small and light-weight package; determining, the surrounding objects, and locations; providing travel instruct ions; recording routes, location s and landmark-s for using later.	Routes limited to the map Of the device.
BrailleNote GPS [11]	A PDA combined with an external GPS receiver; two options for output: Speech and Braille.	Providingtravel Instructions nd informationabout the surroundingenvironments. all wing users to access the maps, Internet and email.	Software is only compatible with specific Windows CE devices.
StreetTalk [12]	A PacMate PDA combined with a external GPS receiver; two options for output: speech And Braille.	Allowing users to plan A journey; providing travel instructions and information bout the surrounding environments.	Software is developed from a GPS program forsi ghted people and notall functions are access ible; not allowing for inpu t of personalized points of interest or personalized routes.

Table 2.2: Examples of GPS based system for visually impaired people.

GPS based Systems have several limitations. They are often expensive due to the cost of the hardware and the maps. Furthermore the GPS signals are often disrupted when the user travel between tall buildings or under dense foliage. Another disadvantage of GPS based system is that this system cannot detect immediate changes of the surroundings.

2.3.3. Computer Vision Based System:

Computer vision systems use images captured from cameras to obtain information about the environment. Voice [13], Brain Port [14], virtual Acoustic Space [15], Electro Neural Vision System [16], Virtual Cane are part of many trending assisting systems.

Systems	Functionality	Interface and components	Advantages	Disadvantages
vOICe [13]	Representing acoustically the environment	A digital camera embedded in eye glasse headphones; Portable computer	Portable.	Require extensive training; blocking the user's hearing; only represent the 2D structure of Scenes.
BrainPort [14]	Representing the environment by gentle electrical stimulation	A camera mounted on sunglasses; a postage-stamp-size electrode A hand-held controller	Portable; does not block user's hearing.	Require extensive training; only represent the 2D structure of scenes.
Virtual Acoustic Space [15]	Representing acoustically the environment	Two cameras embedded in eye-glasses; headphones; portabl computer	Portable; small size; reconstructing the 3D space of the Environment.	Require extensive training; blocking The user's hearing.
Electron-neural vision sys-tem [16]	Detecting obstacles; providinginformatio byelectric stimulatio in both hands; each finger representsa zone in forward field of view.	Two cameras; a compass; a lapto with GPS; gloves With stimulators in each finger.	Real-time performance, d es not block user'shearing.	blocks using hands; does not detect obje s at head and Ground levels.

Table 2.3: Examples of Computer Vision Based System

A handheld camera mounted on a sunglass and a hand held controller for settings and processing in included in Brain port. The captured image is sent to the controller to translate into a stimulation pattern for displaying on the tongue. The bright level of pixels are represented by stimulation levels. The voice system is designed to transform the image into a soundscape. Each pixel is represented in a sinusoidal tone, where each audible frequency corresponds to a vertical position, each time corresponds to horizontal position and amplitude levels denote bright levels. The Voice system consists of digital camera attached to eyeglasses, headphone and a portable where the software is installed.

Figure 2.5: The Voice System

The Voice and Brain port systems are simple lightweight and cheap. However these systems require months of training for users because of complicated patterns representing for the environment. These systems do not provide depth information for the user.

To represent the 3D space of the surroundings environments, many assistive systems apply the stereo vision technique .In this systems, the depth map of the surroundings is obtained from the images, and then is conveyed to the user in different methods. The virtual acoustic space is a depth map. The map is created as a 3D sound environment and is formed using HRTF.

(a) (b)

Figure2.6: The Brain Port Vision Device (a) implementation of the Brain Port (b) illustration of Object detection.

The electronic-neural vision system employs electric simulation on the user’s fingers for representing the detected obstacles and landmarks. Each finger indicates object in a forward zone of view, Simulation level is proportional to the distances.

Figure 2.7: The electron-neural vision system and its components.

These systems do not provide assistance for detecting the travel path, despite the fact that this is a major task for traveling safely and independently in an unknown environment.

2.4) Chapter Summary

Detecting pedestrian path is crucial and challenging task for travels of visually impaired people. Straying outside the walking region is dangerous for the blind travelers. However there has been little work in detecting pedestrian lanes for assisting navigation of visually impaired people. Furthermore existing approach to visual impairment is not either viable or user friendly .Therefore an assistive system for detecting edges of street to help visually impaired people is very much essential.

Proposed Method:

1. A color image is retrieved from the device’s camera.
2. The color image is converted to gray-scale
3. Line is detected in the binary image using the Hough transform.
4. The detected lines are overlaid in the color image.
5. The resulting color image with detected lines is displayed on the screen of the device.
6. There is a safe zone to make sure the drawn line has not deviated horizontally.
7. If the line happens to be inside the safe zone the app sends a continuous audio feedback to the user, otherwise it does not.

Chapter-3: Technical Synopsis

3.1. Converting color image into grayscale image:

Various image processing applications require conversion of color image to grayscale image for different purpose. Grayscale is a range of monochromatic shades from black to white. Therefore, a grayscale image contains only shades of gray and no color. While digital images can be saved as grayscale (or black and white) images, even color images contain grayscale information. This is because each pixel has a luminance value, regardless of its color. Luminance can also be described as brightness or intensity, which can be measured on a scale from black (zero intensity) to white (full intensity). Most image file formats support a minimum of 8-bit grayscale, which provides 2^8 or 256 levels of luminance per pixel. Some formats support 16-bit grayscale, which provides 2^16 or 65,536 levels of luminance [17].

We are converting color images to grayscale images to simplify mathematics, to increase speed of the process and to reduce complexity of the code. Techniques of converting into grayscale remove all color information, leaving only the luminance of each pixel. Since digital images are displayed using a combination of red, green, and blue (RGB) colors, each pixel has three separate luminance values. Therefore, these three values must be combined into a single value when removing color from an image. There are several ways to do this. One option is to average all luminance values for each pixel. This method simply averages the values: (R + G + B) / 3. Another method involves keeping only the luminance values from the red, green, or blue channel.

3.2. Edge Detection:

Due to the importance of image edge detection in image analysis, object recognition and many applications, many edge detection algorithms are used to detect edges of objects in the image. Edges typically occur on the boundary between two different regions in the image. There are a number of algorithms for this, but these may be classified as derivative based where the algorithm takes first or second derivative on each pixel, or gradient based where a gradient of consecutive pixels is taken in x and y direction.

In our project we are using gradient based image edge detection. An image gradient is a directional change in the intensity or color in an image. Image gradient is used to extract information from images. The gradient of a two variable function is a vector which has a magnitude and direction value.

An operation called kernel operation is usually carried out. A kernel is a small matrix centered on a chosen pixel of the image matrix, multiplied the coefficients of the filter with the corresponding pixels of image matrix for the specified pixel located on the center of the matrix, if the calculated value is above a specified threshold, then the middle pixel is classified as an edge, and such calculation is repeated for each pixel of the image, sliding over the image matrix from left to right and from up down [18]. Sobel, Prewitt and Canny are examples of gradient based methods of edge detection.

3.2.1. Sobel edge detector:

The kernel for x and y directions are:

-1	0	1
-2	0	2
-1	0	1

-1	-2	-1
0	0	0
1	2	1

dy dx

The edge detection operation is essentially an operation to detect significant local changes in the intensity level in an image. The change in intensity level is measured by the gradient of the image. Since an image f(x, y) is a two-dimensional function, its gradient is a vector [18]. The magnitude and the direction of the gradient may be computed as given by the formulae 1 and 2 respectively:

$$d = (dx^2 + dy^2)^{\frac{1}{2}} \dots\dots\dots\dots (1)$$

$$\theta = \tan^{-1}\left(\frac{dy}{dx}\right) \dots\dots\dots\dots\dots (2)$$

3.2.2. Canny Edge Detector:

The canny edge detector is the most popular edge detection technique at present. It was developed by John F. Canny in 1986. This algorithm is known to many as the optimal edge detector. The Canny operator works in a multi-stage process. Canny edge detector ensures good noise immunity and at the same time detects true edge points with minimum error. Canny has optimized the edge detection process by-

1. Maximizing the signal-to-noise ratio of the gradient.
2. Edge localization factor, which ensures that the detected edge is localized as accurately as possible. In other way, the detected edges should be as close as possible to the real edges.
3. Minimizing multiple responses to a single edge. One real edge should not result in more than one detected edge.

The signal-to-noise ratio of the gradient can be maximized by detecting true edges and by avoiding false edges. Thus by discarding the false responses when there are multiple number of responses to a single edge, the noise corrupted edges may be removed. In this method the image is first convolved with Gaussian smoothing filter with standard deviation sigma**.** This operation is followed by gradient computation onthe resultant smoothed image [19].

- Non-Maxima Suppression: This is edge thinning process. The Canny edge detector produces thick edges wider than a pixel. The operation of non-maxima the one with lesser edge magnitude is discarded. This will give a thin line in the output image. Suppression thins down the broad ridges of gradient magnitude. There are several techniques for such a thinning operation. In one technique, the edge magnitudes of two neighboring edge pixels, perpendicular to the edge direction are considered.

- Double Thresholding: The image may still contain many false edge points after applying non-maxima suppression. An appropriate threshold need to be selected to remove false edge points. All the edge points having magnitude greater than the threshold may be preserved as true edge points, while others are removed as false edge points. Now the problem arises when the threshold is small, then a number of false edge points may be detected as true edge points, otherwise some true edge points may be missed. To solve this problem, we choose two thresholds T1 and T2 to create two different edge images E1and E2, where T2~1.5T1. E1will contain some false edge points, whereas E2will contain very few false edge points and miss a few true edge points but will have gaps in the contours. The algorithm then combines the results from E1 and E2 in such a manner that: it links the edges of e2 into contours until it reaches a gap, then it links the edge from E2 with edge pixels from a E1 contour until a E2 edge is found again [19].

3.3Hough Transform:

The Hough transform is a technique that is used to isolate features of a particular shape within an image. To be specific, it has been used to extract lines, circles and ellipses. Hough transform can also detect other structures if their parametric equation is known. This method was introduced in 1962 by Paul Hough. The main advantage of the Hough transform is that it is tolerant of gaps in the edges of shapes to some extent and is relatively unaffected by noise in an image and uneven illumination.

- Hough Space: A straight line can be described in 2D coordinate system in various ways. For example:

  In a Cartesian coordinate system (CCS) with parameters (m, b):

  $$y = m * x + b \dots\dots\dots\dots\dots\dots..(1)$$

  Where m is the slope and b is the y-axis intercept

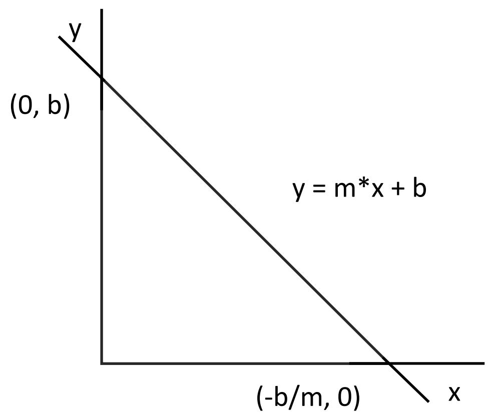


Figure 3.1: Line in Cartesian coordinate system

Equation 1 is not convenient for us to use because we cannot describe vertical lines with it - parameter m becomes infinite [20].

Instead, we shall represent a line in a Polar coordinate system with parameters (r, θ):

$$y = \left(-\frac{Sin\theta}{Cos\theta}\right) * x + (r/Sin\theta)$$

Where r is the length of a line perpendicular to this line, starting from the origin and θ is the orientation angle of r with respect to the x-axis.

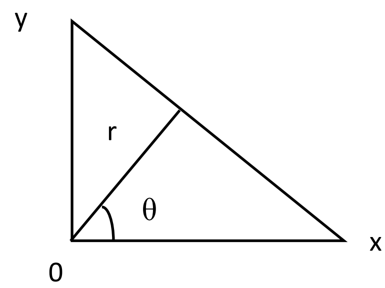
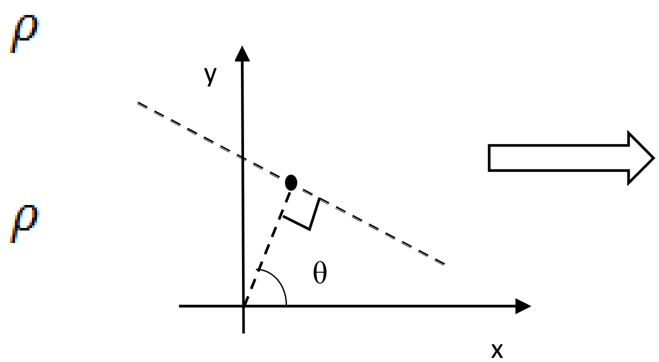


Figure 3.2: Line in Polar coordinate system

We express the parameter r: $r = x * Cos\theta + y * Sin\theta$------------ (2)

Each point [x0, y0] from x, y - plane (figure 3) which satisfies equation 2 gives a sinusoid (figure 4) in a so called Hough space (r, θ - plane) [20].



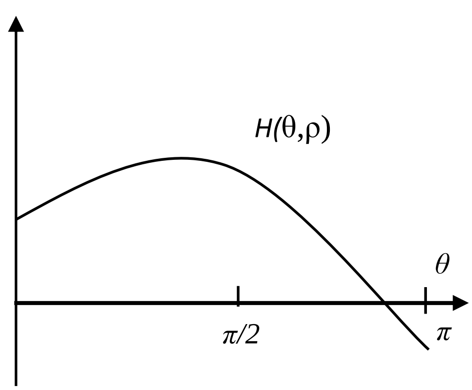


Figure 3.3: Point in CCS   Figure 3.4: Point from CCS projected to HoughSpace

- Detecting Straight Line: When we transform points which lie on the same line in Cartesian Coordinate System we can see that their corresponding sinusoids in Hough space (HS) all intersect at one point of intersection

Hough transform is so effective. Because, interrupted, dashed, even partially damaged Lines in input image will still get detected, because their undamaged segments will form the intersection points which indicate presence of lines. So when we search for local maxima in HS, extract these points of intersection, map them back to Cartesian space and overlay this image on the original image - we get the detected lines.

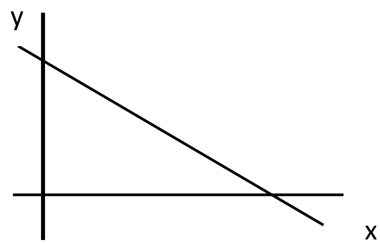


Figure 3.5: Detected Lines

- Algorithm for Hough Transform: The steps are as follows-

1. Using any suitable edge detection scheme find all the edge points in the image.
2. Quantize the (m, c) space into a two-dimensional matrix H with appropriate quantization levels.
3. Initialize the matrix H to zero.
4. Each element of H matrix, H (mi, ci), which is found to correspond to an edge point is incremented by 1. The result is a histogram or a vote matrix showing the frequency of edge points corresponding to certain (m, c) values (For example, points lying on a line).

5. Only the large valued elements are taken by thresholding the histogram these elements correspond to lines in the original image [18].

3.3.1. Probabilistic Hough Transform:

Background: We use standard Hough transform to detect shapes which have parametric equations, for example lines, circles, curves etc. It takes binary image as input and part of an edge feature is represented by each active pixel. These pixels are mapped as many points in Hough space by SHT. All possible lines that could pass through that image point is represented by a single edge is pixel that is mapped to a sinusoid in 2D parameter space (θ ,p) . This is called the voting stage. Sinusoids in parameter space will cross if multiple points in the image are collinear. Now parameters for the line in input images are detected by finding the dissection point in Hough space where most sinusoids cross. This is referred to the search stage [20].

Though the Hough transform is considered as a powerful tool in shape analysis which gives good results even in the presence of noise and occlusion, it has some drawbacks. Main shortcomings of the technique are excessive storage requirements and computational complexity. It needs a lot of computation power to iterate over all the points and to add vote. In order to reduce this computation, researchers have come up with some new probabilistic techniques which would increase the computing speed without losing much accuracy [22]. Different approaches have been introduced concerning Hough Transform to solve these problems. Probabilistic Hough transform is one of the proposed techniques to increase the efficiency of HT.

Probabilistic Hough Transform:

PHT is commonly known as a mathematically corrected version of Hough Transform. The definition of PHT is as follows:

"The Probabilistic Hough Transform H(y) is defined as the log of the probability density function of the output parameters, given all available input features."

Now we say, there is an input image with a setof feature measurements {X1, X2….Xn} which is Xnand a specific point in parameter space y. Then the probability density function in Hough space is p (y| Xn) and the PHT can be written as:

$$H(y) = ln[p(y|x_n)]\ which, by\ Bayes'rule, is$$

$$H(y) = \sum_{i=1}^{n} ln[p(x_i|y)] + ln[p(y)] + c$$

Here C is an arbitrary constant and p(y) istheprior probability distribution (assumed to be uniform).

The Probabilistic Hough Transform introduced by Kiryati*et al* only uses randomly selected small subset of the edge points of the input image and so it reduce the complexity of the voting stage and make the algorithm faster [23].

An instance of computation of PHT is the familiar case of finding straight lines from oriented edges.Different experiments have shown that the normal methods of handling high dimensional Hough spaces suffer a degradation of robustness when the dimensionality exceeds four or five.Although the PHT is not a feasible alternative to conventional methods in problems with few unknown parameters, but the PHT does have some definite advantages when the dimensionality increases.Contrary to conventional HT approaches, the (PHT) is independent of the size, shape, and arrangement of the accumulator array. When the cells have a large coverage in dynamic Hough Transforms, these factors have a strong influence on the number of votes accumulated. In conventional HT, there is a sharp cut-off between the cells receiving a vote or not, which create something a bit like aliasing distortion. The PHT has a lower bandwidth and a larger immunity against noise. Since probabilistic approaches to HT use the idea of random sampling, drastic computational savings are obtained by applying PHT [22].

3.4 Thresholding:

Thresholding is the simplest method of image segmentation. The pixels are divided into regions based on their intensity: the pixels with intensity higher than a predefined threshold value belong to a different region than the pixels with intensity lower than the threshold. Different methods are used to threshold an image. Here we are giving brief description of Global and Local thresholding.

3.4.1 Global (Static) thresholding:

Global thresholding, use an appropriate threshold T. Here the threshold is equal for all pixels in the image (independent of their position in the image).This thresholding operation can be expressed as:

$$dst(x,y) = \begin{cases} 255 & if\ src(x,y) > thresh(T) \\ 0 & otherwise \end{cases}$$

Where dst (x, y) is the intensity value of the pixel at position (x, y) in the destination image, src (x, y) is the intensity of the pixel at (x, y) in the source image and thresh is the numeric threshold value.

Peaks and valleys of the image histogram can help in choosing the appropriate value for the threshold. To choose global threshold T we have to examine the image histogram. If we find two dominant peaks, we will set the threshold between these peaks. Now, if there are multiple peaks then we can adopt some methods to fix the threshold T. We can use iterative method for finding T. The steps for this method are:

1. Estimate value of T (start with mean)
2. Divide histogram into two regions, R1 and R2 using T
3. Calculate the mean intensity values $\mu 1$ and $\mu 2$ in regions R1 and R2
4. Select a new threshold $T = (\mu 1 + \mu 2)/2$
5. Repeat 2-4 until the mean values $\mu 1$ and $\mu 2$ do not change in successive iterations

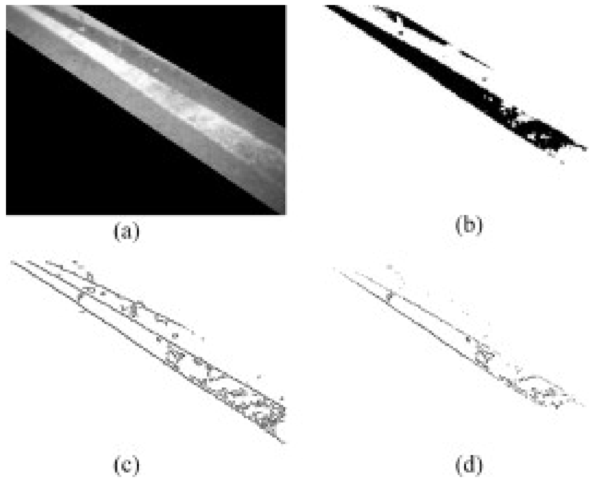


Figure 3.6: Global thresholding

A simple algorithm for global thresholding:

- Initial estimate of T
- Segmentation using T:

  G1, pixels brighter than T;

  G2, pixels darker than (or equal to) T.

- Computation of the average intensities m1 and m2 of G1 and G2.
- Computation of the average intensities m1 and m2 of G1 and G2.

$$T_{new} = (m1+m2)/2$$

- If $|T - T_{new}| > \Delta T$, back to step 2, otherwise stop.

Global thresholding has some drawbacks. It can't handle changing illumination, Can give poor results for certain types of images. We can see, this technique gives us good results when the threshold is carefully adjusted for a particular scenery. The main drawbacks of the global thresholding method are that it gives us bad results when the camera is tilted upwards or downwards just by a fraction of a centimeter, when the scenery changes and when there is a varying illumination in the image. Another problem of this method is that it returns too many white pixels which would later slow down the shape detection algorithm.

3.4.2 Local (adaptive) thresholding:

Adaptive thresholding changes the threshold dynamically over the image. The threshold value depends on the position of the pixel in the image. New threshold is calculated for every pixel from its neighboring pixels. Local thresholding operation can be expressed as:

$$dst(x,y) = \begin{cases} 255 & if\ src(x,y) > T(x,y) \\ 0 & otherwise \end{cases}$$

Where T (x, y) is a mean of the intensity values of all pixels in the n × n neighborhood of (x, y) minus constant C. Usually grayscale or color images are the input images for adaptive thresholding and in simplest implementation output images are the binary imagesrepresenting the segmentation[24]. A threshold is calculated for each pixel in the image. If the pixel value is below the threshold it is set to the background value, otherwise it assumes the foreground value.

To find the threshold we can go for two approaches: (1) The Chow and Kaneko approach and (2) Local thresholding. The assumption behind both approaches is that smaller image regions are more likely to have approximately uniform illumination, thus being more suitable for thresholding. Chow and Kaneko divide an image into an array of overlapping sub-images and then find the optimum threshold for each sub-image by examining its histogram. The threshold for each single pixel is found by interpolating the results of the sub-images. There are some drawbacks of this method, for example it is computational expensive and is inefficient for real-time applications.

Another approach for finding the local threshold is to statistically examine the intensity values of the local neighborhood of each pixel. The statistic which is most appropriate depends largely on the input image. Simple and fast functions include the mean of the local intensity distribution,

$$T = mean$$

$$T = \frac{(max + min)}{2}$$

We should cover sufficient foreground and background pixels and for this the size of the neighborhood has to be large enough, or it will be a poor threshold. On the other hand, choosing regions which are too large can violate the assumption of approximately uniform illumination. This method produces good results for some applications and is less computationally intensive. Adaptive thresholding is used to separate desirable foreground image objects from the background based on the difference in pixel intensities of each region. Local adaptive thresholding selects an individual threshold for each pixel based on the range of intensity values in its local neighborhood. In which images global intensity histogram doesn't contain distinctive peaks, this thresholding is useful for this kind of images.

On the margin of the images or any part where the range of intensity values within a local neighborhood is very small and their mean is close to the value of the center pixel the mean of the local area is not suitable as a threshold. The situation can be improved if the threshold employed is not the mean but (mean-C), where C is a constant. This technique allow us to set all pixels which exist in a uniform neighborhood (*e.g.* along the margins) to background.

Figure 3.7: Adaptive thresholding

Chapter-4: System Design

4.1) Algorithms Used:

A Hough Transform is considered probabilistic if it uses random sampling of the edge points. These algorithms can be divided based on how they map image space to parameter space.

One of the easiest probabilistic method is to choose m edge points from the set M edge points. The complexity of the *voting stage* reduces from O (M.NΘ) to O (m.NΘ). This works because a random subset of M will fairly represent the edge points and the surrounding noise and distortion.

Smaller value of m will result in fast computation but lower accuracy. So the value of m should be appropriately chosen with respect to M.

Pseudo Code for Probabilistic Hough Transform:

1. Quantize the Hough transform space: identify maximum and minimum values of r and theta and the total number r and (θ) values.
2. Generate an accumulator array A(r,θ); set all values to 0.
3. Random set edge points from all edge points (x, y)in the image.

   Do

   Compute the normal direction (θ) (gradient direction or orientation-90 degree)

   Compute r from $xi\ cos$ (θ) +$yi\ sin$ (θ) = r

   Increment A (r, θ)

4.For all cells in the accumulator array

[Accumulator is a 2D array for m and c representation. It has mxn number of bins each representing a particular (r,θ)]

Do

Search for maximum values

The coordinates r and (θ) give the equation of the corresponding line in the image.

4.2) Tools:

Android

The key elements to consider when choosing the appropriate library for a CV application that is supposed to solve road lane detection problems and run on a mobile platform are:

- Speed-we want to detect road lanes in real time, which means that the library has to be highly optimized because we need to process multiple frames, make our own computations and draw the desired result back to the user every second.
- Low complexity - we want the used algorithms to be as efficient as possible and have a low complexity, because of the limited hardware in smart phones.
- Low battery consumption-since smart phones are powered by batteries with limited capacity and we cannot assume with absolute certainty that the smart phone running the application will be connected to a power source at all times.
- Android ecosystem-the library should be written in Java or at least provide a port for the Android operating system in JNI5, so that it could be easily integrated to the application since Android applications are developed mainly in Java.

OpenCV:

Taking all these aspects into account I have chosen the OpenCV library for the following reasons:

- It is the most advanced open source CV library with the widest community support.
- It is designed to be high performance .The algorithms are written in C++ and are compiled to a highly optimized native code, which suits our needs since Android runs on

a Linux kernel.

- Even though the algorithms are written in C++, OpenCV maintains direct port for the Android platform
- It includes all the algorithms that can be found in a desktop version in the form of native Android libraries and at the same time provides Java interfaces for an access to those algorithms.

4.3) Collecting data:

To collect data for all the testing and experimenting we took photos of footpaths found around the BRAC University campus. Collecting good data was really difficult as most of the footpaths do not have visibly clear lines that we can work on. So we had to take a lot of photos to compile our result.

4.4) Methodology:

1. Taking photos with the android smart phone camera:

Using the smart phone's native camera we will be taking pictures from the road ahead. The roads ahead need to be clear to work properly. The lines should be visible. After this step we will be turning this image into a grayscale image. Gray scaling image opens up a lot of new opportunities to manipulate the image we receive from the camera. We have already discussed how grayscale works. Shown in figure 4.1.

Figure 4.1: Grayscale image

2. Running the image through edge detection filters:

To detect the lines we are going to need the edges of the lane that we are going to follow. To do that we have to find the edges first. There are two edge detectors that we are going to apply to see which one works better. First we will be applying the canny edge operator.

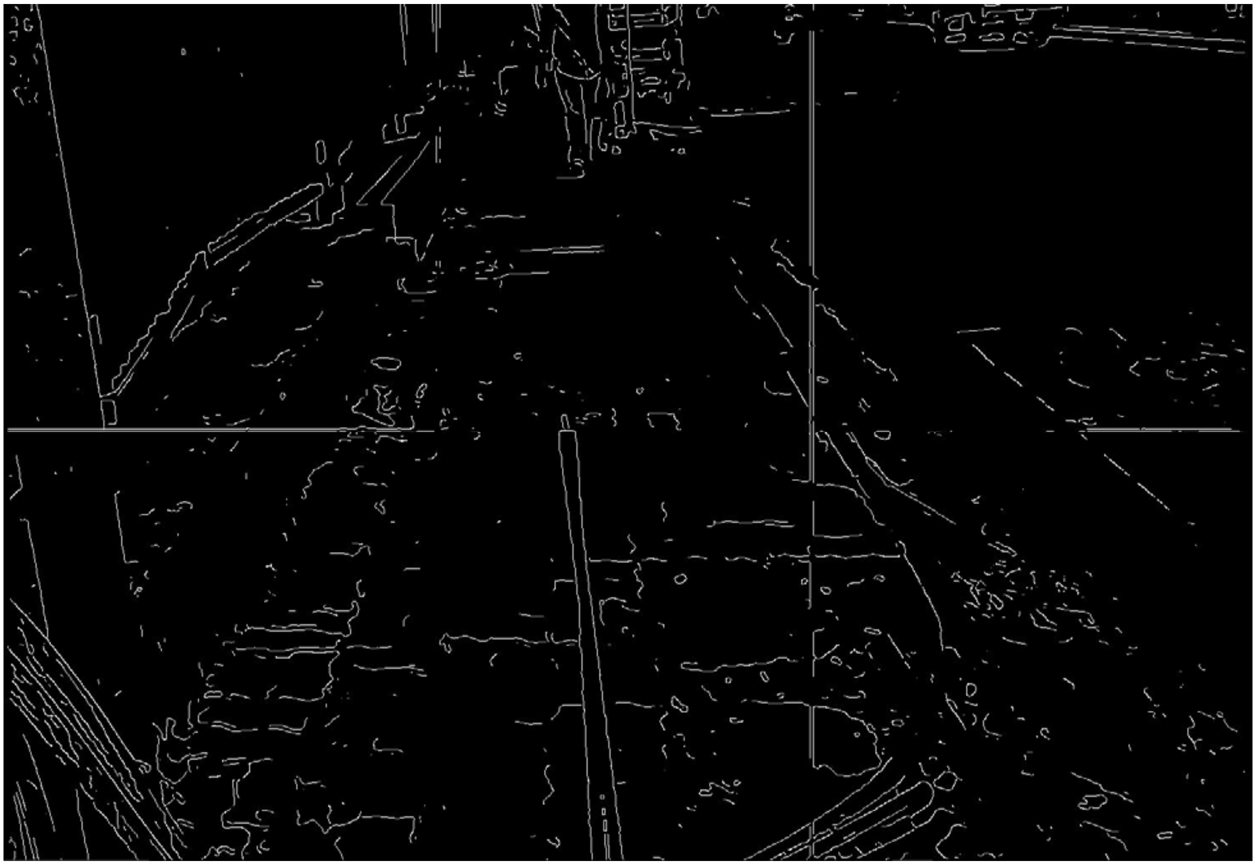


Figure 4.2: Canny edge

We will be using canny for our purpose.

3. Probabilistic Hough man transformation:

This is the most vital step of the whole process. Using probabilistic Hough transform we will be able to find out the lane we are going to follow. Overlaying the line on the image gives us the following result. Figure 4.3 is a screen cap that demonstrates when the line is within the safe zone. On the other hand figure 4.4 shows an example of the tracking line outside the safe zone. When the line is in the unsafe zone it will stop sending audio cues to the user to move forward.

Figure 4.3: Line in safe zone

Figure 4.4: Line in unsafe zone

4.Region of interest:

A Region of Interest is an area of an image, which is graphically selected from a window displaying that image. User will focus on this part of the image for further analysis. In region of interest a selected subset within a dataset is selected for a particular purpose. Many intelligence, surveillance, and reconnaissance applications need to detect potential targets which is region of interest.The dataset could be any of the following:

- Waveform or 1D dataset: The ROI is a time or frequency interval on the waveform (a graph of some quantity plotted against time).
- Image or 2D dataset: The ROI is defined by given boundaries on an image of an object or on a drawing.
- Volume or 3D dataset: The ROI is the contours or the surfaces defining a physical object.
- Time-Volume or 4D dataset: Concerning the changing 3D dataset of an object changing in shape with time.

We need region of interest to sort out the significant part of the image. In blind navigational system, it is ideal to keep the line perfectly underneath the user or it will not be safe. We are calling this region of interest the safe zone.The Android application detects the significant line from the image captured from the rear camera and in response gives an audio feedback to the user. The main reason of ROI implementation is that we would want to keep the user inclined with the detected line. If ROI isn't used the user may get audio response upon detecting the significant safe line but he won't be directly on it. The deviation from that safe line wouldn't be ideal. The definition of safe line here is the type of line which is bold enough to stand out and there is not much obstruction on it.

5. Feedback:

In physical science we perceive sound as waves. Waves are then perceived by our ear drum and our brain and we hear things. It is commonly assumed that the improvement in the remaining senses is a result of learned behavior; in the absence of vision, blind people pay attention to auditory cues and learn how to use them more efficiently. But there is mounting evidence that people missing one sense don't just learn to use the others better. The brain adapts to the loss by giving itself a makeover. If one sense is lost, the areas of the brain normally devoted to handling that sensory information do not go unused — they get rewired and put to work processing other senses. Most of the research on cross-modal neuroplasticity has focused on blind individuals, who often have enhanced auditory abilities. Brain imaging studies show the visual cortex of the blind people is overtaken by other senses, such as hearing or touch, and contributes to language processing.

The application is focused on creating an audio response upon detecting the line in the ROI of the image. In the audio response function of the application. The function works as below:

- There is this static region of interest on the display of the mobile.
- The audio is created on detecting the lane in the region of interest.
- Detection of line in the ROI triggers a default mp3 file.
- The mp3 file is default on purpose so the feedback is not complex.
- If the line is on the display of the mobile but not in the ROI, for obvious safety reason the audio is not triggered.

Chapter-5: Experimentation analysis:

The results gathered from the experiment was widely varied. It was only successful in a select few scenarios. What we could gather from our limited experimentations are:

- The system only works well in well-lit areas. It needs ample lighting to properly find the lines that it can follow.
- Android devices can range from low range to upper high range fidelities. The performance of the system depends a lot on how good the camera on the device is. The processing power needed is marginal, so the efficiency will most likely be same on most of the devices.
- It needs clear lines to follow, which comes as no surprise as we are traversing using the footpath lines. So any footpaths where the lines are unclear or cannot properly be read by the camera can result in inaccurate results.

Limitations and Future Work:

Following the investigations presented in this thesis, improvements to the proposed approaches that could be made in the future includes:

1. The application doesn't account for obstacles and other human beings. This could be further worked on in future.

2. Many of the streets in cities don’t have significant footpath or significant edge to follow .This deficiency of edge will lead to halt in movement for the user.

3. The user need to be positioned on to the lane at first. With the first lane detected the user can proceed on following.

Chapter-6: Conclusion

This thesis presents a vision based system designed to detect robustly pedestrian lanes in different environments for assistive navigation of visually impaired people. The system aims to locate the walking lane in front of the traveler in each scenario. The major tasks of the system includes, identifying the lane type in the image captured from the camera in real time. The experimental results have shown effectiveness and robustness of the proposed approaches in comparison with existing algorithms.

References:

1. WHO, "Visual Impairment and Blindness". N.p., 2014. Web. 13 Dec. 2015.
2. Dorweiler, Jason. "Revisiting Lane Detection Using Opencv". *Transistor.io*. N.p., 2015. Web. 12 Dec. 2015.
3. Ran, L. (n.d.). Drishti: An Integrated Indoor /Outdoor Blind Navigation Systemand Service. Retrieved from http://www.cise.ufl.edu/~helal/projects/publications/percom_04_helal_sumi_camera_ready.pdf
4. M. Brambring, “Mobility and orientation processes of the blind,” in Electronic spatial Sensing for blind, D. Warren and E. Strelow, Eds., Dordrecht, the Netherlands, Nijho , 1985, pp. 493–508.
5. L. Kay, “A sonar aid to enhance spatial perception of the blind: engineering Design and evaluation,” Radio and Electronic Engineer, vol. 44, no. 11, pp. 605–627, 1974.
6. GDP Research, “The miniguide ultrasonic mobility aid,” Tech. Rep., 2011. [Online]. Available: http://www.gdp-research.com.au
7. J. Borenstein and I. Ulrich, “The guidecane - A computerized travel aid for the active guidance of blind pedestrians,” in IEEE International Conference on Robotics and Automation, 1997, pp. 1283–1288.

8. J.M.Loomis, R.G.Golledge, R.L.Klatzy, J.M.Speigel and J.Tietz,"Personal guidance system for virtually impaired," in ACM conference on Assistive technologies,1994,pp.85-91
9. H.Petrie, V.Johnson, T.Strothotte, A.Raab, R.Michel, K.Reichert and A. Schalt," MoBIC: An aid to increase the independent mobility of blind travelers, "British Journal of Visual Impairment, vol.15, no 2, pp.63-66, 1997.
10. Human ware ,"Trekker Breeze," Tech. Rep., 2014[Online].Available http://store.humanware.com/hus/trekker-breeze-plus-handheld-talking-gps.html
11. SenderoGroup," Braillenote GPS", Tech.Rep.,2014 http://store.humanware.com/hus/braillenote-gps-software-and-receiver-package.html
12. RLA.COM RL & Associates - Assistive Software - StreetTalk GPS Solution In-text: (Rla.com) Bibliography: Rla.com, "RL & Associates - Assistive Software - Streettalk GPS Solution". N.p., 2015. Web. 18 Dec. 2015.
13. P.B.L. Meijer," An experimental system for auditory image representations," IEEE Transactions on Biomedical Engineering, vol.39, no.2, pp.112-121, 1992.
14. A.arnoldussen and D.C. Fletcher, "Visual perception for the blind: the brain port vision device," vol.9, 2012. [Online].Available:http://www.retinalphysician.com/articleviewer.aspx?articleid=106585
15. H.Petrie, V.Johnson, T.Strothotte, A.Raab, R.Michel, K.Reichert and A.Schalt," MoBIC: An aid to increase the independent mobility of blind travelers, "British Journal of Visual Impairment, vol.15, no 2, pp.63-66, 1997.
16. 15.J.L, Gonzalez-Mora, A.Rodriguez-Hernndez, L.FRodrguez-Ramos, L.Daz-Saco and N.Sosa, "Development of a new space perception system for blind people, based on the

creation of a virtual acoustic space," Engineering Applications of Bio-Inspired Artificial Neural Networks,vol.1607,pp.321-330,1999

17. Annadurai, S., &Shanmugalakshmi, R. (2007). Fundamentals of Digital Image Processing. New Delhi, ND: Dorling Kindersley (India) Pvt. Ltd.
18. Acharya, T., & Ray, A. K. (2005). Image Processing Principles and Applications. New Jersey, NJ: A John Wiley & Sons, MC. Publication.
19. Nosrati, M., Karimi, R., Hariri, M. &Malekian, K. (2013). Edge Detection Techniques in Processing Digital Images: Investigation of Canny Algorithm and Gabor Method. *World Applied Programming, Vol (3), Issue (3).* Retrieved from http://waprogramming.com/papers/517cbaaeca0d93.40790585.pdf
20. Hamarne, G., Althoff, K. & Abu-Gharbieh, R. (1999). Automatic Line Detection. Project Report for the Computer Vision Course. Retrieved from https://www.cs.sfu.ca/~hamarneh/ecopy/compvis1999_hough.pdf
21. Stephens, R. S. (1990). A Probabilistic Approach to the Hough Transform. Retrieved from http://www.bmva.org/bmvc/1990/bmvc-90-012.pdf
22. Kalviainen, H., Hirvonen, P., Xu, L., &Oja, E. (1995). Probabilistic and non-probabilistic Hough transforms: overview and comparisons. Image and Vision Computing, Volume 13, Issue 4. Retrieved from http://www.cse.cuhk.edu.hk/~lxu/papers/journal/KHXU95rht.PDF
23. Kiryati, N., Eldar, Y., & Buckstein, A. M. (1991). A probabilistic Hough transform. Pattern Recognition, Volume 24, Issue 4. Retrieved from http://dl.acm.org/citation.cfm?id=104026

24. Ferrari, S. (2012). Image segmentation. Retrieved from https://homes.di.unimi.it/ferrari/ElabImm2011_12/EI2011_12_16_segmentation_double.pdf