CALIFORNIA STATE UNIVERSITY, NORTHRIDGE

Wearable Computing for Image-Based Indoor Navigation of the Visually Impaired

A thesis submitted in partial fulfillment of the requirements

For the degree of Master of Science in Computer Science

By

Gladys Mae Ison Garcia

May 2015
The thesis of Gladys Mae Ison Garcia is approved:

___________________________________   _______________________
Dr. George Michael Barnes            Date

___________________________________   _______________________
Dr. Gloria Melara                    Date

___________________________________   _______________________
Dr. Ani Nahapetian, Chair           Date

California State University, Northridge
# Table of Contents

Copyright .............................................................................................................................. ii

Signature ............................................................................................................................. iii

List of Figures ....................................................................................................................... vi

Abstract ............................................................................................................................... viii

1. Introduction ...................................................................................................................... 1

2. Related Work .................................................................................................................. 3

3. System Overview ............................................................................................................. 5

  3.1 Google Glass .............................................................................................................. 5

  3.2 Android Smartphone ................................................................................................. 6

  3.3 OpenCV ...................................................................................................................... 7

  3.4 Mobile Applications ................................................................................................. 7

4. Approach ......................................................................................................................... 9

  4.1 Device Communication .............................................................................................. 9

  4.2 Floor Detection .......................................................................................................... 10

  4.3 Walk Path Analysis .................................................................................................... 12

  4.4 User Feedback ........................................................................................................... 14

5. Results ............................................................................................................................. 15

  5.1 Floor Detection Results ............................................................................................ 16

  5.2 Walk Path Analysis Results ...................................................................................... 21
6. Conclusion........................................................................................................................................... 26

References............................................................................................................................................... 27
List of Figures

Figure 3.1. Google Glass and its specifications .......................................................... 6
Figure 3.2. General overview of the system ............................................................. 8
Figure 4.1. Technical overview of the system ............................................................ 9
Figure 4.2. Sample input images ............................................................................. 10
Figure 4.3. Floor detection step by step output ....................................................... 12
Figure 4.4. Input to walk path analysis phase .......................................................... 13
Figure 5.1. Sample images showing floor space from the user to the wall ............. 15
Figure 5.2. Sample images captured when the user is standing close to the wall .... 16
Figure 5.3. Successful floor detection results ............................................................ 18
Figure 5.4. Unsuccessful floor detection due to other elements present in the image..... 19
Figure 5.5. Unsuccessful floor detection due to low contrast between wall and floor pixels ................................................................................................................. 20
Figure 5.6. Unsuccessful floor detection due to corridor turns .................................. 20
Figure 5.7. Comparison of the height of the floor outline and the actual distance from the wall using dataset 1 ................................................................. 22
Figure 5.8. Comparison of the height of the floor outline and the actual distance from the wall using dataset 3 ................................................................. 23
Figure 5.9. Comparison of the height of the floor outline and the actual distance from the wall using dataset 4 ................................................................. 23
Figure 5.10. Comparison of the height of the floor outline and the actual distance from the wall using dataset 5 ................................................................. 23
Figure 5.11. Comparison of the height of the floor outline and the actual distance from the wall using dataset 6 .......................................................... 24

Figure 5.12. Comparison of the height of the floor outline and the actual distance from the wall using dataset 7 .......................................................... 24

Figure 5.13. Combined graph of all datasets showing the similarity of the trends of each data line ......................................................................................... 25
Abstract

Wearable Computing for Image-Based Indoor Navigation of the Visually Impaired

By

Gladys Mae Ison Garcia

Masters of Science in Computer Science

In this thesis, an image-based non-obtrusive indoor navigation system for the visually impaired is presented. The system makes use of image processing algorithms to extract floor regions from images captured from a wearable eye-mounted heads-up display. A prototype system called VirtualEyes is presented, where floor regions are analyzed to provide the user with voiced guidance for navigation. The floor detection algorithm was tested against over 200 images captured from indoor corridors of various lighting conditions and achieved up to 81.8% accuracy. The system proves to be an effective approach to navigational guidance for the visually impaired.
1. Introduction

There are an estimated 285 million people in the world that are visually impaired according to the data from World Health Organization [1]. 39 million are blind and 246 million are considered as people with low vision. “Low vision” is a common term used for people with moderate to severe visual impairment.

Blind people, in some cases, require help from another person in order to navigate unfamiliar environments. Others rely on a walking stick or a guide dog to navigate around an environment and avoid obstructions. Researchers have focused on alternative ways to help blind people navigate spaces. Advancements in mobile and wearable technology have the potential to advance research in this area.

In the robotics field, image-based approaches are commonly used for navigational guidance or obstacle detection. Information is extracted from the captured images of the environment with the use of image processing algorithms. Image processing is the use of algorithms on input images and outputs either another transformed image or a set of characteristics related to the image.

In this thesis, a floor detection algorithm that was used for the automatic navigation of a mobile robot was adapted to create a mobile indoor navigation system for the visually impaired. The created system, VirtualEyes, is composed of a paired Google Glass and Android smartphone connected via a secure Bluetooth connection.

The Google Glass is a wearable device that is capable of running Android applications. It has a camera feature that was utilized in this system in order to capture images from the surroundings. This is an ideal device to use for this application as it is
non-obtrusive and is worn in a way that allows for the camera to capture an unobstructed view of the environment.

The following chapters in this thesis will present the various research work related to the system, the hardware and software components of the system, and the technical approach used for navigational guidance using image processing. The algorithms developed for VirtualEyes were tested using over 200 different images and the results are presented along with the conclusion of the research.
2. Related Work

Wearable devices have been widely used for different applications. In this chapter, the different works related to wearable devices and assistive technology are presented.

Najeeb et al. presented a wearable system that uses an off the shelf EEG device that reads brain signals to select letters, compile words, and create sentences meant for people with paralysis [19]. A smart watch was used in [20] to recognize arm gestures for hands-free interaction. Altwajry et al. presented a system that uses Google Glass in [21] that can recognize landmarks by capturing an image of the scene and the GPS information if available.

There has also been research in wearable devices for guiding the visually impaired in unfamiliar environments. The underlying technology varies from a modernized version of the walking stick (a.k.a. white cane) and image-based approaches.

Fernandes et al. presented a system that uses RFID tags attached at the end of the white cane in [2]. A virtual white cane was presented in [3] by using a laser pointer attached to a smartphone. Both approaches require the use of specialized hardware.

In terms of the image-based approach, different systems made use of a smartphone [4], Microsoft Kinect [5], and custom hardware using two cameras mounted on the user’s shoulders [6] as interfaces to gather images of the environment. The use of pre-installed special markers in the environment to identify a safe walking path for the user is also presented in [7].
Such image-based systems commonly adapt the floor detection or obstacle detection implementation in the robotics field. The use of stereo vision is common in this approach as discussed in [6][8]. These systems are able to detect floor regions and obstacles in the environment and calculate the distance of such objects from the user.

The work of Tapu et al. [4] uses monocular vision by utilizing the camera in a smartphone which is a less obtrusive design. Obstacle detection is performed to guide the user when walking in the outdoor environment.

Another common approach in the robotics field is to use image sequences from the video feed as described in [9][10] to track the movement in the scene.

In terms of floor detection, other approaches make use of a single indoor image of the environment to classify floor regions. The implementation presented in [11] makes use of image segmentation to identify floor regions in the image. Authors of [12][13] use horizontal and vertical lines found in the image to detect floor regions.
3. System Overview

In this chapter, the architectural overview of VirtualEyes, a system for the visually impaired navigation guidance, is presented. The system is composed of a paired Google Glass and Android smartphone. These two devices, connected via Bluetooth, work together in gathering image data from the indoor environment using the Google Glass camera, process the data on the smartphone, and provide valuable feedback to the user through the use of the built-in speaker on the Glass.

In the following sections, the overview of the hardware and software components is presented.

3.1 Google Glass

The Google Glass, as shown in Figure 3.1, is a head-mounted, rechargeable battery-operated wearable device, developed by Google and was initially released in 2013, which is capable of running Android applications. This device has features similar to a smartphone ranging from high resolution display, camera, Bluetooth, Wi-Fi, etc. [14]. The camera in the Glass is capable of taking 5 megapixel images. Since the device is worn over the eyes of the user, similar to prescription glasses, the images taken from the camera captures the surroundings in the perspective of the user.
Figure 3.1. Google Glass is a rechargeable battery-operated wearable device with built-in features such as camera, speakers, Bluetooth, Wi-Fi, etc.

Since this device is still in its early stages, there are a few limitations to its performance. The battery in the device typically lasts one hour of usage, especially with the use of Bluetooth and the camera. Heating of the device could also cause discomfort to the user while the glass is in operation. Furthermore, there is not much processing power available in the glass to perform the image processing required in this system. To overcome these limitations, the Bluetooth capability of the Glass was utilized to pair it with an Android smartphone and offload processing that would require substantial power.

3.2 Android Smartphone

Mobile smartphones have been a ubiquitous device that is accessible for most people. The higher processing power and better battery life in these smartphones as compared to Google Glass allows for an ideal mobile and lightweight device for performing powerful operations that might prove difficult to run on the Glass.
The use of Android operating system allows the integration of many open source third party libraries that provides an easy to use framework in performing tasks required by the system such as OpenCV.

3.3 OpenCV

Open Source Computer Vision (OpenCV) is a widely used library of image processing algorithms. The library supports different operating systems including Android and has interfaces for a variety of programming languages such as C, C++, and Java [15]. The built-in functions in the OpenCV library were used in this system for most of the image processing tasks.

3.4 Mobile Applications

There are two different applications developed for the system which are installed in the respective devices. Figure 3.2 shows an overview of the functionalities and communication between the applications.

An android application (client app) is installed in the Google Glass that will start up the Glass camera and send captured image frames to the paired Android smartphone. This application also receives text information coming from the Android smartphone and converts this into voice guidance using the Text-To-Speech framework of the Android operating system.
Figure 3.2. Images captured by the Google Glass are sent over to the Android smartphone via Bluetooth for Image Processing and Floor Detection and Analysis. The results of the analysis are sent back to the Google Glass for voiced guidance.

Another android application (server app) is installed in the Android smartphone that is paired with the Google Glass. This application performs various image processing algorithms using OpenCV in order to extract information from the received images. The features from the image are extracted which are then evaluated to analyze the floor region. Once the image analysis has been completed, a feedback is sent to the Google Glass through the Bluetooth connection that has been established.
4. Approach

In this chapter, the implementation of VirtualEyes is discussed which includes the communication between the paired devices, floor extraction and analysis, and user feedback as seen in Figure 4.1.

![Diagram of Device Communication]

**Figure 4.1.** The Google Glass app provides the input and output information from and to the user. The Smartphone app performs Floor Extraction and Analysis by performing a series of image processing algorithms.

4.1 Device Communication

The paired devices transmit data to each other over a Bluetooth connection. The Glass application continuously sends image frames to the Android smartphone for processing.
A 320x240 RGBA image frame captured by the Glass, as shown in Figure 4.2, is about 300 kilobytes. To increase the frame rate of the application, this is compressed to a jpeg format using the built-in compression function in the OpenCV library. This reduces the size of the image to less than 100 kilobytes.

![Sample 320x240 indoor images captured by the Google Glass showing a variety of floor pattern, shape, and lighting](image)

**Figure 4.2. Sample 320x240 indoor images captured by the Google Glass showing a variety of floor pattern, shape, and lighting**

### 4.2 Floor Detection

The images captured by the glass typically contain the walls, floor, ceiling, and other objects within the frame. The floor region is surrounded by walls in all sides. By detecting the wall-floor boundaries from the image, the floor region can be detected
within the image as shown in Figure 4.3 and later analyzed to provide feedback to the user.

The floor detection approach discussed in [12] was adapted in the implementation of this system. This approach is capable of detecting floor regions from a single indoor corridor image.

The first step is to apply the Canny Edge detection [16] algorithm in the image to identify the edges in the image. An edge is a region in the image where there is a sudden change in the pixel intensity. This outputs a black and white image where the white pixels are the identified edges in the image.

From the black and white edge image, we try to find vertical and horizontal lines in the image using Hough Line Transform [17]. Vertical lines are defined as lines that are within 10 degrees from the vertical direction. Horizontal lines, on the other hand, can go from 40-70 degrees from the horizontal direction. Due to the noisy conditions in the scene (i.e. posters on the walls, shadows from lighting, etc.), there could be vertical and horizontal lines that are detected which are not part of the wall-floor boundary. In order to minimize the incorrect line extraction, lines that match any of the below conditions are removed:

1. Lines that are shorter than 30 pixels
2. Vertical lines that exist entirely on the upper half of the image
3. Horizontal lines that appear above the vanishing point

All the remaining vertical and horizontal lines are assumed to be part of the wall-floor boundaries. The convex hull for all the endpoints of the lines is computed which
gives the rough estimate shape of the floor region in the image. The convex hull implementation in OpenCV [18] was used for the prototype.

![Image](image-url)

**Figure 4.3.** Results of each image processing step. From top-bottom, left-right: (a) input image, (b) vertical lines in red, (c) horizontal lines in green, (d) cyan dots as the intersections of every pair of horizontal lines (vanishing point), (e) yellow line as the average y-axis value of all vanishing points, (f) convex hull of detected horizontal lines

### 4.3 Walk Path Analysis

The output from the previous floor detection step is a polygon indicating the detected floor. In the walk path analysis step, the outline of the floor is used to determine how much floor space is ahead of the user.

When walking along a corridor, the perspective of the user shows the walls on each side of the corridor, floor, and ceiling as seen in Figure 4.3 (a). The vanishing point of the perspective line in the image is roughly located at the center of the image depending on the height and viewing angle of the user. The floor region in such viewing angle is roughly shaped like a trapezoid where the base is wider than the top. The height of the
floor region would indicate the proximity of the user to the end of the corridor. The 
height decreases as the user approaches the end of the corridor as seen in Figure 4.4. By 
using the height of the estimated floor outline, the system can make an analysis on 
whether the user is safe to proceed walking or should stop to avoid hitting a wall.

The floor detection phase returns a list of points that forms the outline of the detected 
floor region. The height of this floor region is computed by taking the difference between 
the lowest and highest point in the outline. By testing the system using 320x240 pixel 
images from multiple environments, it was found that a good threshold for the floor 
outline height is 30 pixels. An image where the height of the floor region is less than the 
threshold value indicates that the user is standing close to a wall. On the other hand, a 
floor region height that is greater than the threshold indicates that the user has enough 
walking space from the wall.

![Figure 4.4. Consecutive image frames showing the decreasing height of the detected floor region as the user approaches the end of the corridor](image-url)
4.4 User Feedback

The Google Glass has a built-in speaker that uses bone conduction technology. The speaker is utilized to give guidance to the user while navigating in an indoor environment.

The walk path analysis phase determines whether it is safe for the user to continue walking forward or should the user stop. This information is delivered to the user using the built-in speaker. By using the Text-To-Speech library in Android, the user can hear alerts from the system. VirtualEyes will tell the user to “Stop” or “Walk” every few seconds.
5. Results

The floor detection and walk path analysis algorithms discussed in Chapter 3 are tested using test images taken from various locations in California State University, Northridge (CSUN) campus. The client application was installed in a Google Glass Explorer Edition version 2 with firmware version XE22. This device runs on a Texas Instruments OMAP 4430 SoC 1.2Ghz Dual (ARMv7) processor with 2GB of RAM. The server application is installed in a Samsung Galaxy S4 running Android version 4.4. This smartphone has a Qualcomm MDM9215 + APQ8064T 1.9GHz Quad-core with 2GB of RAM.

Different datasets were collected from various corridors in the CSUN campus specifically in Jacaranda Building, Bayramian Hall, and Sierra Hall. The images were captured while walking in a constant pace along the corridor towards a wall. For each of the 7 datasets, the first image was taken with a distance from the user to the wall that ranges from 30 to 60 feet. As the user approaches the wall in a constant pace, this distance becomes smaller as seen in Figure 5.1. The last few images in the dataset were about 2 to 5 feet from the wall where the floor is no longer visible which is shown in Figure 5.2.

![Sample images with enough distance from the user to the wall that shows the floor region](image1)

Figure 5.1. Sample images with enough distance from the user to the wall that shows the floor region
5.1 Floor Detection Results

The test data were taken from different locations with a variety of color and texture of the floor and walls. Dataset 1 contains images of corridors with good floor and wall color contrast. The images have varying lighting conditions due to the windows that are present on the right side of the corridor. Dataset 4 contains images where the floor and walls have different colors. These images contain reflective floor surfaces as opposed to dataset 1 and have bulletin boards on the wall. The rest of the datasets are composed of images where the floor and walls have a poor contrast. However, there is a darker colored baseboard that separates the wall and floor in images in dataset 3, 5, 6, and 7. Sample results from each dataset are shown in Figure 5.3.
Figure 5.3. The floor detection algorithm is able to estimate the floor region from captured images of corridors with different color, texture, and lighting. The images above shows the results of the floor detection phase on datasets 1 to 7 (top to bottom).

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Total Images</th>
<th>Correctly Identified</th>
<th>% of Correctly Identified</th>
<th>Incorrectly Identified</th>
<th>% of Incorrectly Identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>48</td>
<td>36</td>
<td>75%</td>
<td>12</td>
<td>25%</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>2</td>
<td>10%</td>
<td>18</td>
<td>90%</td>
</tr>
<tr>
<td>3</td>
<td>22</td>
<td>18</td>
<td>81.82%</td>
<td>4</td>
<td>18.18%</td>
</tr>
<tr>
<td>4</td>
<td>27</td>
<td>18</td>
<td>66.67%</td>
<td>9</td>
<td>33.33%</td>
</tr>
<tr>
<td>5</td>
<td>48</td>
<td>37</td>
<td>77.08%</td>
<td>11</td>
<td>22.92</td>
</tr>
<tr>
<td>6</td>
<td>24</td>
<td>17</td>
<td>70.83%</td>
<td>7</td>
<td>29.17%</td>
</tr>
<tr>
<td>7</td>
<td>94</td>
<td>62</td>
<td>65.96%</td>
<td>32</td>
<td>34.04%</td>
</tr>
</tbody>
</table>

Table 5.1. Table shows the number of correctly and incorrectly identified floor regions in the different sets of test images
The floor detection algorithm heavily relies on edges found in the images. If there is a good contrast between the floor and the wall pixels in the image, the system will more accurately detect the floor region in the image. Datasets 3 and 7 have the highest accuracy out of all the datasets that were tested with about 81.8% and 77.1% respectively. Although the floor and walls have a similar color, there is a darker colored baseboard on the wall that clearly separates floor pixels from the wall pixels. The algorithm, however, failed in situations where the user is turning into another corridor.

Images from datasets 1 and 4 have very distinct floor and wall boundaries but some images were affected by other conditions, shown in Figure 5.4. Images from dataset 1 contain a window on the right side of the image. Objects outside the window contain edges that were also detected by the edge detection algorithm which negatively affected the floor detection. For dataset 4, bulletin boards that are attached on the wall caused stray edges to be detected which cause the floor detection to incorrectly identify the floor region.

![Figure 5.4. Other elements present in the image that could confuse the edge detection algorithm such as windows and bulletin boards cause the floor detection algorithm to fail.](image)

Of all the sets of data for testing, dataset 2 has the lowest accuracy rate with just10%. These are images of corridors where the floor and wall color are very similar as shown in
Figure 5.5. The edge detection step failed to detect the wall-floor boundary which caused the floor detection to fail. In this kind of input images, it might help to have an image pre-processing step that would enhance the edges in the image without making the image noisy.

![Image of floor detection failure](image1.png)

**Figure 5.5. The floor detection fails on images where the floor and wall pixels have low contrast**

Furthermore, the floor detection does not perform well on images captured when turning in corridors as shown in Figure 5.6. The algorithm relies on finding the wall-floor boundaries on both sides of the floor. When turning in corridors, there is only one side where the wall is visible.

![Image of floor detection failure in corridors](image2.png)

**Figure 5.6. The floor detection algorithm fails on correctly estimating the floor outline when turning in corridors**
5.2 Walk Path Analysis Results

The floor outline result of the floor detection phase is used as input in the walk path analysis. To analyze the results of the walk path analysis phase, the height of the floor outline was compared against the actual distance of the user from the wall when the image was captured. Since the walk path analysis phase is highly dependent on the accuracy of the results from the floor detection phase, images that did not have successful floor detection results were removed from the dataset for this testing. Furthermore, since dataset 2 had a very low overall accuracy in floor detection, it was not included for this testing.

The results of the comparison are shown as graphs in Figure 5.7 to Figure 5.12. The vertical axis indicates the height of the detected floor region in pixels. The horizontal axis indicates the distance of the user from the wall when the image was captured.

It can be seen from all of the figures that the overall trend of the graph indicates a decreasing height of the floor outline. This reflects the decreasing distance as the user walk closer to the wall. At about 10 feet or less, the height in pixels of the floor outline begins to decrease sharply. And at about 5 feet is where the floor outline height drops to 0. This indicates that the floor detection phase no longer detects any floor in the image which is accurate as seen in sample images in Figure 5.2.

The graph also shows that height of the detected floor region cannot be used to accurately determine the actual distance of the user from the wall. The floor outline height does not linearly decrease along with the decreasing actual distance. There are random spikes in the graph which indicates that the floor region detected increased in
height. This is mainly caused by the user movement while walking. The use of a head mounted camera is sensitive to changes in orientation. There will be a slight change to the height of the camera as the user makes a step forward. Furthermore, camera orientation will also be affected by movements of the head of the user.

If the floor detection phase returns an inaccurate result, this affects the result of the walk path analysis phase. To overcome this, VirtualEyes was designed to keep a running average of the floor outline height as the user walks forward. The average height of the resulting floor outline of the past 10 images is computed. This value is used in determining the appropriate feedback sent to the user. With this approach, if only one image in a continuous image sequence fails in the floor detection step, this will not greatly affect the results of the walk path analysis phase.

![Graph showing comparison of height of floor outline and actual distance from wall using dataset 1](image)

**Figure 5.7.** Comparison of the height of the floor outline and the actual distance from the wall using dataset 1
Figure 5.8. Comparison of the height of the floor outline and the actual distance from the wall using dataset 3

Figure 5.9. Comparison of the height of the floor outline and the actual distance from the wall using dataset 4

Figure 5.10. Comparison of the height of the floor outline and the actual distance from the wall using dataset 5
Figure 5.11. Comparison of the height of the floor outline and the actual distance from the wall using dataset 6

Figure 5.12. Comparison of the height of the floor outline and the actual distance from the wall using dataset 7

Figure 5.13 shows the combined graphs of the results for the walk path analysis phase for all datasets. The chart shows that the images that were captured at the exact same distance from the wall do not have the exact floor region height result using presented approach. However, there is a consistent characteristic from all data lines and they all show a similar trend which begins with a slight decrease that sharply drops to 0 when the user gets to a distance of 10 feet or less from the wall. Therefore, the floor region height can be used as a parameter in estimating the proximity of the user to the wall.
Figure 5.13. Combined graph of all datasets showing the similarity of the trends of each data line.
6. Conclusion

This thesis has shown the effectiveness of using mobile devices for a navigational guidance system for the visually impaired. By using non-obtrusive devices, users can navigate around indoor environments and blend in with the rest of the crowd.

The system uses floor detection in user indoor guidance, instead of the related work’s approach of obstacle detection. The effectiveness of this approach was demonstrated with the VirtualEyes prototype. The system achieved up to 81.8% accurate detection of the floor on a set of over 200 distinct images. The floor detection algorithm implemented in the system works well in corridors where the wall on both sides are visible and have a distinctive color contrast between the floor and the walls. Detection of floors on images with minimal color contrast could be improved with the use of some image pre-processing algorithms.

Although the system is not able to accurately determine the exact distance of the user from the wall, it can still effectively alert the user when the floor outline height reaches a low value which indicates that there is no more walking space ahead of the user.
References

[1] WHO Visual impairment and blindness


[16] Canny Edge Detector - OpenCV 2.4.9.0 Documentation

[17] Hough Line Transform - OpenCV 2.4.9.0 Documentation

[18] Convex Hull – OpenCV 2.4.11.0 Documentation

