

2013

3D Hallway Modeling Using a Single Image

Gregory M. Olmschenk
CUNY City College

[How does access to this work benefit you? Let us know!](#)

Follow this and additional works at: http://academicworks.cuny.edu/cc_etds_theses

 Part of the [Computer Sciences Commons](#)

Recommended Citation

Olmschenk, Gregory M., "3D Hallway Modeling Using a Single Image" (2013). *CUNY Academic Works*.
http://academicworks.cuny.edu/cc_etds_theses/190

This Thesis is brought to you for free and open access by the City College of New York at CUNY Academic Works. It has been accepted for inclusion in Master's Theses by an authorized administrator of CUNY Academic Works. For more information, please contact AcademicWorks@cuny.edu.

3D Hallway Modeling Using a Single Image

Submitted in partial fulfillment of
the requirement for the degree of

Master of Science (Computer Science)

at

The City College of New York

of the

City University of New York

by

Gregory M. Olmschenk

May 2013

Approved:

Professor Zhigang Zhu, Thesis Advisor

Professor Douglas Troeger, Chairman
Department of Computer Science

Abstract

Real-time, low-resource corridor reconstruction using a single consumer grade RGB camera is a powerful tool for allowing a fast, inexpensive solution to indoor mobility of a visually impaired person or a robot. The perspective and known geometry of a corridor is used to extract the important features of the image and create a 3D model from a single image. Multiple 3D models can be combined to increase confidence and provide a global 3D model. This paper presents our results on 3D corridor modeling using single images. First a simple but effective 3D corridor modeling approach is introduced which makes very few assumptions of the camera information. Second, a perspective based Hough transform algorithm is proposed to detect vertical lines in order to determine the edges of the corridor. Finally, issues in real-time implementation on a smartphone are discussed. Experimental results are provided to validate the proposed approach.

Index Terms — indoor modeling, vanishing point, visual impairment, perspective based Hough transform.

Acknowledgments

This work is partially supported by US National Science Foundation Emerging Frontiers in Research and Innovation Program under Award No. EFRI-1137172. The proposed future work of this thesis has been granted a PSC-CUNY enhanced research grant (Award # 66786-00 44).

I would like to thank my advisor Professor Zhigang Zhu for his time and effort in mentoring my research. I would also like to thank Edgardo Molina and the other members of the Visual Computing Laboratory for their support and help during my work.

Table of Contents

Abstract.....	2
Acknowledgments.....	3
Table of Contents.....	4
1. Introduction.....	5
2. Related Work.....	7
3. Overview of Our Approach.....	9
4. Corridor Modeling Algorithms.....	11
4.1. 3D Corridor Modeling.....	11
4.2. Vanishing Point Detection.....	12
4.3. Perspective Based Hough Transform.....	13
5. Experimental Results and Discussions.....	15
5.1. Experimental Results.....	15
5.2. Discussions on Time Performance and Integration.....	20
6. Conclusions and Future Work.....	22
7. References.....	23

1. Introduction

In the world built by humans, hallways are everywhere. As such, there are many cases where the information about the corridor is important. For one, hallway detection is an important component for any indoor navigation system including those used by the visually impaired, and can be used to help visually impaired people avoid the walls as well as find the doors and cross section corridors necessary to take them to their destination. Furthermore, robots are continually becoming more common. A mobile robot needs to be able to navigate its environment. If that environment involves the inside of a building, it's likely the robot will frequently encounter hallways.

The most common way to extract the information on a corridor in computer vision currently is to use a range sensor and perform the computations necessary to find the walls of the corridor. Unfortunately, these methods require relatively expensive equipment, heavy processing, and high energy consumption. This makes the method impractical for a robot that needs to move quickly, a wearable system that need to be built cheaply, or a blind person who doesn't want to carry a laptop on their back processing data all day. The most recent relatively low-cost 3D sensors are the RGB-D sensors, such as Microsoft Kinect and Asus Xtion Pro, but their sensing ranges are quite limited (0.5 to 4 meters) and are not very suitable for corridor detection. We propose a high speed, low resource, inexpensive solution to these needs. Effective portable corridor detection capabilities with nothing more than a single consumer camera such as a webcam or the cameras found on a smartphone. This is done through the detection, localization and 3D reconstruction of corridors from individual images, while multiple images can provide more detailed and confident information.

Furthermore, the methods described here can be used alongside their slower range sensor counterparts. One such example is providing quick results so that a robot can begin to act while the range sensors finish processing a more detailed report. As the RGB camera does not have a range limitation as the range sensors do, another example is using the methods described here combined with data already retrieved from range sensors to make assumptions about what the camera can see beyond the range sensors limits.

We have made the following three contributions for enabling a visually impaired people to navigate in a typical corridor scenario. First a simple but effective 3D corridor modeling approach is introduced which makes very few assumptions of the camera information: we only need to assume the camera height and that it is without an obvious rolling angle, and the 3D information of the corridor and the floor plan is obtained by utilizing the vanishing point of all the horizontal lines. It should be noted that the general approach can still work even with an unknown rolling angle, even though our current implementation assumes none. The assumption of known height can also be relaxed when accurate 3D measurements are not required. Second, a perspective based Hough transform (PBHT) algorithm is proposed to detect feature lines in order to determine the edges of the corridor. The algorithm makes full

use of the 3D information obtained via vanishing point analysis so feature lines at varying distances are all used for determine the edges of the corridor. Finally, issues in real-time implementation on a smartphone are discussed, including model representations, computation costs and communications with the blind user. Experimental results are provided to validate the proposed approach.

The rest of thesis is organized as follows: First, Section 2 discusses what's been done in the past, how it relates to what we did in our research, and what we've done new. Next, Section 3 gives an overview of the approach used in this thesis. Section 4 describes the algorithms used to model the corridors in detail. Then experiment results are shown in Section 5. Finally, conclusions as well as discussions about ongoing work are provided at Section 6.

2. Related Work

The majority of existing corridor detection methods use range sensor data to extract corridor information. The most detailed indoor modeling methods use point cloud data gathered by 3D range sensors [1]. This 3D data can often be processed more quickly by reducing the point clouds to surfaces [2]. Other range methods use 2D laser range sensors to provide a faster, less resource intensive approach [3]. These range sensor methods use powerful hardware to extract precise position data without the need to interpret information based on a RGB image. The most recent low-cost RGB-D sensors, such as Microsoft Kinect and Asus Xtion Pro, can obtain depth maps in real-time, but their sensing ranges are quite limited (0.5 to 5 meters) and are not very suitable for corridor detection and 3D measurements of over 10 meters long. Our approach allows for fast, efficient processing without the need for 3D equipment. Furthermore, our methods can be used in conjunction with a range sensor to provide initial data and make decisions about the environment beyond the range sensor's limit, while the range data can be processed slower to give more detailed results.

The use of a single RGB camera to detect doors has been used extensively in existing research [4][5]. The means used to find the frames of the doors in these methods is similar to the extraction of corridors in our approach. In particular, the vanishing point is used to determine a perspective so that important feature lines can be extracted. The major differences between the method given in this thesis and the door detection are the means by which the features are found. While these related works use simple edge detection and a search for a specific shape, our method utilizes our PBHT and the characteristics of a hallway to gather important features (such as corridor boundaries as well as vertical lines) in the image. We then generate 3D information from these features using well established methods [6].

Past research has been done on allowing a robot to quickly move through a corridor [7]. However, these past methods rely on existing 3D maps of the corridors which the robot matches it's currently surroundings with. The method described in this paper expands on this by requiring no previous knowledge. Other fast corridor detection is done with powerful laser range sensors [8][9] leading to cumbersome equipment being required.

Detection of corridor vanishing points to determine the direction of the hallway has been extensively studied [10]. We build on these methods by utilizing and creating a model of the hallway and its features such as doors, turns and corners of the hallways.

Other research has previously used the properties of environments similar to hallways to perform 3D reconstruction [11][12][13][14]. However, our method extends this by using the perspective to improve results at a distance. Furthermore, our method is designed efficiently for the detection and 3D measurements of corridors using a smartphone so that useful results can be used in applications such as the guidance of the visually impaired.

Incorporating perspective into a Hough transform is not a new idea [15]. However, existing approaches attempt a more general solution in which performance suffers. In the case of real-time processing on a smartphone, we need to use a more efficient approach with better performance.

3. Overview of Our Approach

While most uses of 3D reconstruction require multiple viewpoints, a range sensor, or some other special imaging component, our methods can perform the task using a single image. This is due to the nature of a hallway (Figure 1, Figure 2). While there are plenty of differences from one hallway to the next, the property of having (mostly) straight walls makes the single image processing possible. From the camera's point of view, many of the lines found in an image run to a vanishing point. More specifically, the lines which define the boundaries between the floor to the walls and the walls to the ceiling are lines that go to the vanishing point. Furthermore, lines that are perpendicular to the hallway (e.g. doorframes, floor tiles, etc.) will appear as lines parallel to one another that do not go to the vanishing point.

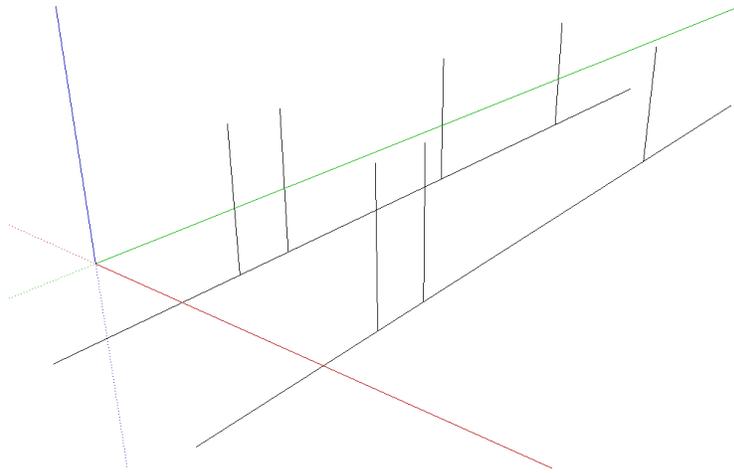


Figure 1: 3D wireframe model of a corridor, with boundaries between the two side walls and the floors, and the frames of doors.

To perform the reconstruction, we begin by performing a canny edge detection [16]. The result is then put through a Hough transform [17] to produce a map of lines. Taking into account the special vanishing point feature of a corridor, we are able to determine that the position of the vanishing point will be where the most lines from the Hough transform converge. From here, we search for surface features, parallel lines which do not run towards the vanishing point. This is done using a perspective based Hough transform (PBHT) which retrieves line segments. This PBHT is specifically built to be based on the perspective of the image. That is, in the hallway image there may be strongly defined lines (such as doorframes) far down the hallway, yet they would be normally excluded from a regular Hough transform because these lines are so much smaller than lines physically closer to the camera. To account for this, the Hough transform adds a weighting biased towards objects closer to the vanishing point. The line segments found serve two purposes. First, the line segment will lie on one of the surfaces, so the end of this segment can be used to estimate which line that passes through the vanishing point defines the

boundary between one surface to the next. Second, the line segment can be used to find features. The most obvious example of this is the line that outlines a door or connecting hallway. From this information, a 3D wireframe model of the hallway can be built.

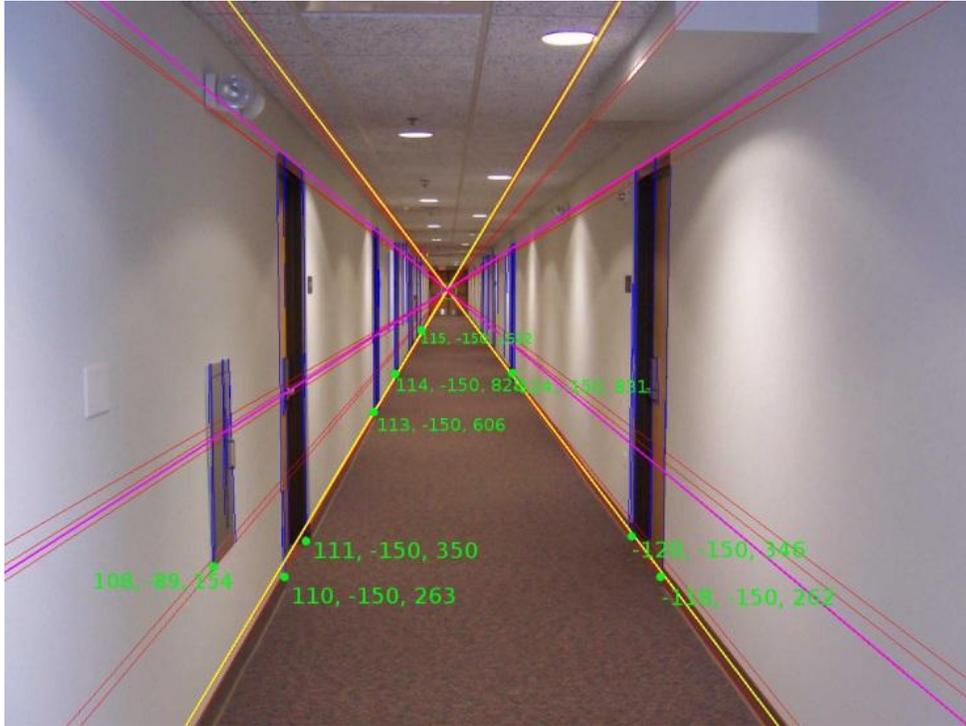


Figure 2: A typical image of a corridor. The direction of the corridor is detected by finding the vanishing point of horizontal lines (in pink), the boundaries of the corridor (the yellow lines below the vanishing point) via the detection of vertical structures (the blue lines). 3D measurements (X,Y,Z) are also labeled for a few vertical lines (in cm).

As a summary, the approach of this research consists of four steps:

- (1) Create reliable single image 3D reconstruction of corridors;
- (2) Use multiple reconstructions to increase confidence and get absolute distance information;
- (3) Case studies to rigorously test and compare the accuracy, speed, and resource consumption of the methods described here in comparison to other common methods;
- (4) The integration of these methods with existing methods and technologies.

The following sections explain in detail the progress so far, how the methods will work, and how the testing and integration will be done. We will mainly focus on the first step.

4. Corridor Modeling Algorithms

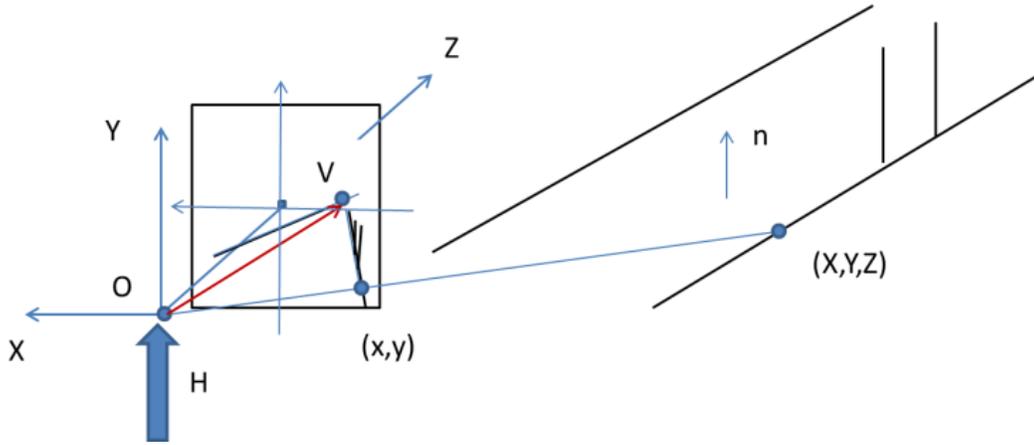


Figure 3: Corridor Detection And Modeling: The Geometry

We will give describe the mathematical model of a 3D corridor as the foundation of our algorithms in detecting vanishing points and boundaries of the corridor. Then the algorithm for vanishing point detection will be presented, followed by the description of our proposed perspective based Hough Transform algorithm.

Once a 3D wireframe model has been generated, the data can be analyzed. Features on the surfaces of the walls, floor, and ceiling can now be compared with relative distances. Particularly, important features such as doorways/openings may be extracted. Furthermore, a set of wireframe data easily combines with another set of wireframe data. Two such sets quickly provide absolute distances and increase the confidence of the data.

4.1. 3D Corridor Modeling

We assume the height of the camera is known (H), and the focal length is F . We also assume that the camera is leveled in the horizontal direction (along the X axis), so that we only have a heading angle (α) and a tilt angle (β) relative to the corridor (Figure 3). If the vanishing point of the corridor is $V = (x_0, y_0, F)$, representing in the camera coordinate system, then the heading angle of the camera is:

$$\alpha = \text{atan}\left(\frac{x_0}{2F}\right) \quad (1)$$

And the tilt angle is:

$$\beta = \text{atan} \left(\frac{y_0}{2F} \right) \quad (2)$$

The normal of the ground plane in the camera coordinate system is $\mathbf{n} = (0, F, -y_0)$. Define the magnitude of the normal vector as $n = |\mathbf{n}|$. Since the camera is leveled, then the plane equation can be written as:

$$FY - y_0Z = nH \quad (3)$$

Hence, given a floor image point (x, y) , the 3D coordinates of the corresponding point can be calculated as:

$$Z = \frac{n}{y - y_0} H \quad (4)$$

$$X = \frac{nx}{F(y - y_0)} H \quad (5)$$

$$Y = \frac{ny}{F(y - y_0)} H \quad (6)$$

when $y \neq y_0$. Otherwise the ground plane equation would be $Y = H$, and the corresponding 3D point would be simply $(H \frac{x}{y}, H, H \frac{F}{y})$.

From this, the distance of Z each image point on the floor can be estimated, which will be used in the PBHT algorithm in Section 4.3. The location and the width of the corridor, and the location/width of each door can be calculated. The width of a door can be calculated as the distance between two endpoints of the door frames on the floor. The distance of each edge of the corridor can be calculated as the perpendicular distance of the camera center to the 3D line of the edge once the 3D coordinates of two points on the lines are obtained.

4.2. Vanishing Point Detection

First we need to determine the vanishing point of the corridor image. This vanishing point and the characteristics of a corridor shape are essential to generate the 3D model from a single image.

The canny edge detection [18] in our implementation begins with a noise reduction using a 5x5 Gaussian filter. Next, a pair of convolution masks are applied in both the x and y directions using the 3x3 Sobel operators. From this a gradient intensity map of the image is generated. A non-maximum suppression is applied which sets all points to zero which are not part of the local maximum. This results in only thin lines (edges) remaining. The canny edge detection terminates after filtering pixels based on thresholds. The detection results in an edge map. Such an edge map is generated for each of red, green, and blue versions of the image individually, and then the maps are combined into a single edge map.

The edge map then is then subjected to a standard Hough transform. Every edge point is used to add to an accumulator array specified by polar coordinates. Then, each bin of the array which exceeds a given threshold specifies an existing line. The result of the transform is a list of lines given in polar coordinates.

The pixels of the image are then surveyed to find which pixel has the most lines passing through or near it within a given buffer distance. Based on the characteristics of a corridor, this location is determined to be the vanishing point of the image. The vanishing point of the corridor is noted as (x_0, y_0) .

4.3. Perspective Based Hough Transform

Next our method extracts the surface feature lines from the image. To do this we propose a perspective based Hough transform (PBHT).

The PBHT is a modified method of the progressive probabilistic Hough transform (PPHT) [11]. The use of PPHT has several advantages over a standard Hough transform. It is generally faster as it only needs to add a portion of the total points to the accumulator, it results in a finite line segment rather than an infinite line, and it can be stopped early and still present information on the most prominent lines.

The PBHT weights voting values based on the proposed line's normal distance from the vanishing point. The algorithm works by first randomly selecting a point to vote. After a vote has been cast, we check the bins to see if the current values would be achieved by random noise. More specifically, from the total value N allocated to all bins, does any bin exceed a threshold of s . The threshold s changes as votes are cast. Once a line is found, the values of the points on that line are removed from the bins and other points supporting that line are removed from the map of points which have not yet voted.

Every pixel which votes does so into exactly one bin for each θ value. This allows us to only require analysis along the ρ dimension of the accumulator. The value allocated to each bin is based on the shortest distance of the line defined by that bin to the vanishing point. The value attributed to the bin based on a line distance is proportional to:

$$v = \frac{1}{d + c} \quad (7)$$

where d is the shortest distance of the line from the vanishing point and c is a constant based on the size of the image. Since we are searching for vertical lines, this is equivalent to using the distance information of these vertical lines: the closer a line to the vanishing point, the larger the distance of the line to the camera. This constant prevents extremely small numbers of pixels close to the vanishing point from being considered lines.

The equation of the line is found by converting the polar point (ρ, θ) to a Cartesian point (x, y) . Then a line normal to the line runs from the origin to (x, y) , so the slope of the line is $\frac{-x}{y}$ and it passes through (x, y) . Since the vanishing point is (x_0, y_0) it is simple to find the shortest distance from here.

The PBHT algorithm works in detail as follows:

PBHT algorithm

- 1) If the input image is empty, end the algorithm.
- 2) Update the accumulator with random pixel from the input image.
- 3) Choose a random pixel from the input image.
- 4) For each line that passes through the pixel, get the distance of that line to the vanishing point. Update the accumulator with values inversely proportional to the distance.
- 5) Remove the pixel from the input image.
- 6) Check if the bin with the highest value of the bins modified by the new point exceeds the threshold. If not, go to step 1.
- 7) Search the given line for the longest number of continuous pixels which do not have a gap larger than a given threshold.
- 8) Remove the pixels in the line segment from the input image and subtract the vote values for each bin of those which have already voted.
- 9) If the line segment total vote value exceeds a given threshold value, add it to the output line list.
- 10) Go to step 1.

Once the surface line segments have been found, the transition from floor to the walls is found based on the end points of the line segments. A simple linear regression of the end points is taken and the line from the original Hough transform which most closely matches the result is determined to be the transition line.

5. Experimental Results and Discussions

5.1. Experimental Results

In our experiments, the center of the image and the focal length are determined simply by using the metadata of the image. The center is simply calculated as the middle of the image and the focal length is converted into pixels using the size of the sensor target.

In our first example (Figure 2), a perfect hallway was captured. In this simple case, we see that we can accurately find the vanishing point, the transitions from the walls to the floor and ceiling, and the edges of the first few door frames. The door frames farther down the hall only return a single feature line, but based on this we can assume that there will likely be a door at these locations. In this example, the image center and the focal length are: $C_{image} = (610px, 457px)$ and $F = 343px$. Based on these parameters we find the vanishing point $V = (53px, 92px)$. Therefore, the heading and tilt angles of the camera can be calculated, using Eq. (1) and Eq. (2), as $\alpha = 4.41^\circ$ and $\beta = 7.64^\circ$. This means the heading angle is almost straight forward and there is only a slight tilt. From here, the normal of the ground is calculated as $\mathbf{n} = (0, 343px, -92px)$. The calculated 3D position for several of the door features are given in Table 1, some of them are also labeled in Figure 2.

Table 1. 3D measurements of a simple example (Figure 2)

Left Side			Right Side		
X (cm)	Y (cm)	Z (cm)	X (cm)	Y (cm)	Z (cm)
108	-89	154	-118	-150	262
110	-150	263	-120	-150	346
111	-150	350	-124	-150	831
114	-150	606			
114	-150	828			
115	-150	1622			

In our second example (Figure 4), we have a more complex scene than in the first. Lockers provide many additional features but pillars of unconventional shape affect the data. In this example we see that boundaries of the hallway are still accurately determined. However, due to the complexity of the bank of lockers, some false positives may be considered doorways. While these may be corrected by combining data from multiple images, it does show the single image variation has some limitations.

The image center and the focal length are: $C_{image} = (246px, 185px)$ and $F = 181px$. Based on these parameters we find the vanishing point $V = (-15px, 47px)$. Therefore, the heading and tilt angles of the camera can be calculated using, Eq. (1) and Eq. (2), as $\alpha = 2.37^\circ$ and $\beta = 7.40^\circ$. This means the heading angle is almost straight forward and there is only a slight tilt. From here, the normal of the ground is calculated as $\mathbf{n} = (0, 181px, -47px)$. These calculations also verified the correctness of the detection. The calculated 3D position for several of the door features are given in Table 2.

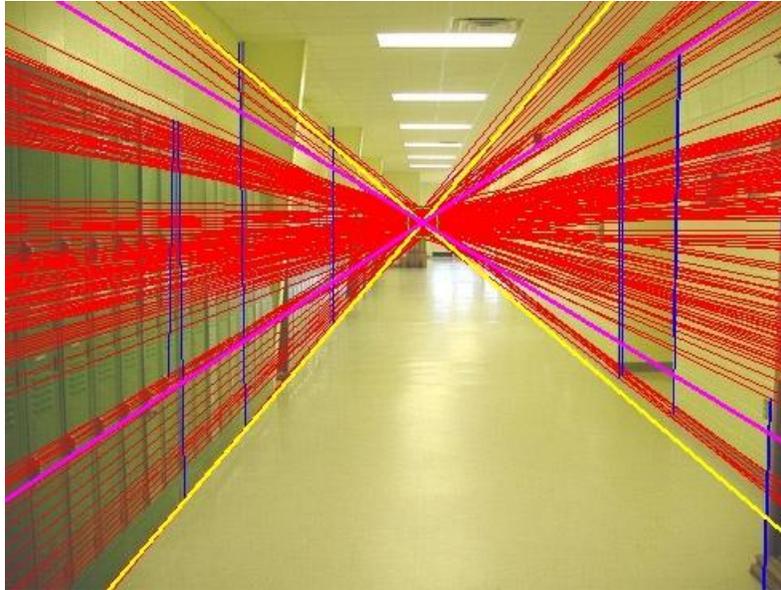


Figure 4: Example 2: A corridor with a bank of lockers. The red lines are all lines which voted for the accepted vanishing point. The blue are the detected vertical line features. The yellow are the transitions from the floor to the walls. The purple are the transitions from the walls to the ceiling.

Table 2. 3D measurements of a more complex scene with a bank of lockers (Figure 4)

Left Side			Right Side		
X (cm)	Y (cm)	Z (cm)	X (cm)	Y (cm)	Z (cm)
151	-130	104	-159	-130	244
150	-130	250	-160	-130	361
148	-130	504	-156	-130	783
147	-130	578			

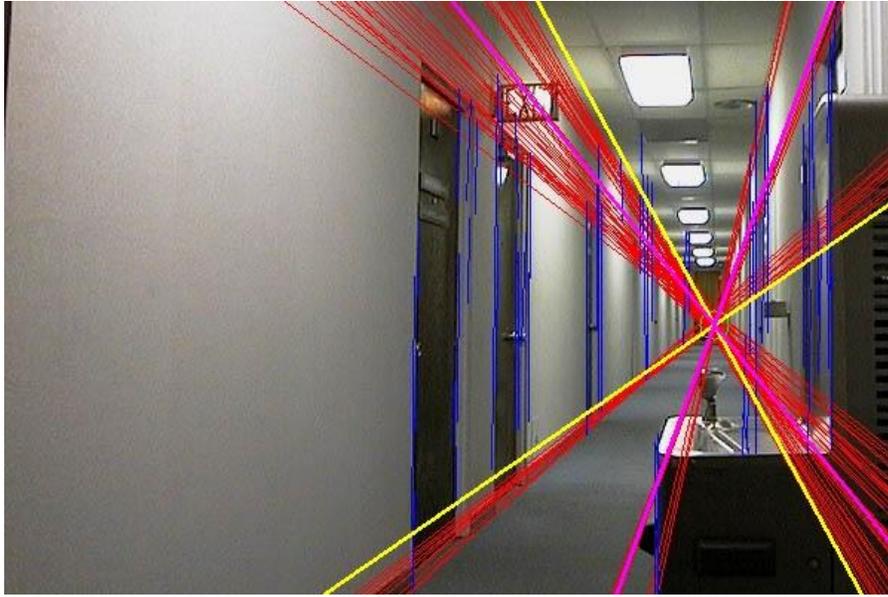


Figure 5: Example 3: Heading to one side of a corridor with a large obstacle. The color of the lines has the same meaning as in Figure 4.

In our third example (Figure 5), we have much of the hallway obscured by a water fountain. The results here show that even when a large portion of the image is obscured, the vanishing point and boundaries of the hallway can still be accurately determined.

Here is a situation where multiple images/wireframe models combined would help determine the correct information. The image center and the focal length are: $C_{image} = (300px, 200px)$ and $F = 269px$. Based on these parameters we find the vanishing point $V = (-178px, -18px)$. Therefore, the heading and tilt angles of the camera can be calculated using, Eq. (1) and Eq. (2), as $\alpha = 18.31^\circ$ and $\beta = 1.92^\circ$. This means the heading angle is largely off center, but there is almost no tilt. From here, the normal of the ground is calculated as $\mathbf{n} = (0, 269px, -18px)$. The calculated 3D position for several of the door features are given in Table 3.

Table 3. 3D measurements of a hallway with a large obstacle (Figure 5)

Left Side			Right Side		
X (cm)	Y (cm)	Z (cm)	X (cm)	Y (cm)	Z (cm)
85	-110	150	-64	-110	157
81	-110	192	-67	-110	578
76	-110	241	-156	-130	783
73	-110	283			
69	-110	523			

In our fourth example (Figure 6), we have a hallway with many students congregated. Here we find that the information on the layout of the corridor is still relatively accurate despite the congregation of students. While several feature lines may be detected that do not match to a door, the majority can immediately be disregarded due to the distance between the two edges of the false “doors” not matching what is the expected range of sizes of a doorway.

The image center and the focal length are: $C_{image} = (512px, 342px)$ and $F = 311px$. Based on these parameters we find the vanishing point $V = (-56px, 4px)$. Therefore, the heading and tilt angles of the camera can be calculated using, Eq. (1) and Eq. (2), as $\alpha = 5.14^\circ$ and $\beta = 0.37^\circ$. The calculated 3D position for several of the door features are given in Table 4.



Figure 6: Example 7: A hallway with many students standing.

Table 4. 3D measurements of a hallway with many students congregated (Figure 6)

Left Side			Right Side		
X (cm)	Y (cm)	Z (cm)	X (cm)	Y (cm)	Z (cm)
342	-150	204	-250	-150	394
341	-150	302	-250	-150	484
336	-150	612	-249	-150	496
335	-150	709	-249	-150	583

In our fifth example (Figure 7), we have a few people in a hallway with wood paneling on the walls obscuring part of the door frames on the side towards the camera. Similar to the last example, the people obscuring the hallway have little impact on information found of the general hallway structure. On the other hand, the wood paneling obscures the door features and the line across the hinges of the first door on the right could easily be incorrectly matched to the leading edge of the second door as this distance is appropriate for a doorway. With this we again see some of the limitations of the single image method. As with the third example (Figure 5), false positives can arise in such conditions, but these can be corrected with multiple image data.

The image center and the focal length are: $C_{image} = (241px, 213px)$ and $F = 152px$. Based on these parameters we find the vanishing point $V = (68px, 59px)$. Therefore, the heading and tilt angles of the camera can be calculated using, Eq. (1) and Eq. (2), as $\alpha = 12.61^\circ$ and $\beta = 10.98^\circ$. The calculated 3D position for several of the door features are given in Table 5.

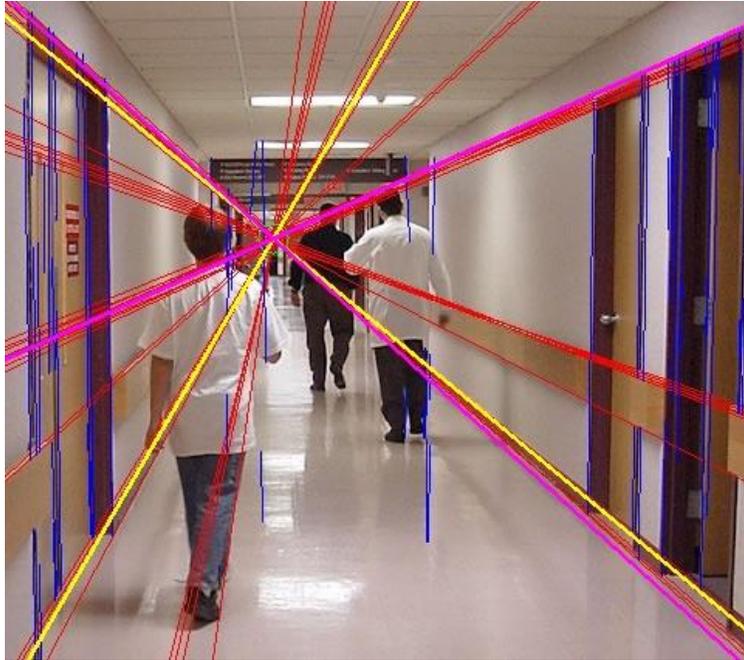


Figure 7: Example 5: A hallway with a few people walking and strange wall features.

Table 5. 3D measurements of a hallway with people and features on wall (Figure 7)

Left Side			Right Side		
X (cm)	Y (cm)	Z (cm)	X (cm)	Y (cm)	Z (cm)
109	-130	210	-211	-130	208
109	-130	315	-210	-130	256
			-209	-130	308
			-207	-130	414

5.2. Discussions on Time Performance and Integration

We have generated 3D wireframe models of the corridor from a single image. The wireframe models will be matched based on the relative positions of the prominent feature lines such as the edges of doorframes. Again, because the geometric characteristics of a corridor are so simple, the matching of these 3D models requires very little processing. In our current implementation on a laptop using C++, the average time for processing an image of a size of 1200x900 is 0.462 seconds. The processing produce important information for doorway navigation and matching, including the axis of the corridor having lines which are parallel that represent the transition from the walls to the floor, and feature lines, such as the door frames, intersecting these lines but are perpendicular to them.

Thus, the matching problem is reduced to matching the distances between the feature lines. Furthermore, full 3D maps of the corridor can be created and referenced. Since from each single image, the 3D measurements of the corridor (the floors and doors, turns, etc) have been obtained, we will use the location measurement data (such as accelerometer and magnetic components) from a smartphone to roughly align the local models and then by matching the local floor and door models, we refine the localization results and generate a global 3D wireframe model. Information about the device/user can also be interpreted from the data. The differences in capture times and change in frame of the images allows for a determination of speed. Since the position of the camera is easily found relative to the 3D model of the corridor, our method can even be used to provide simple real-time 3D SLAM (simultaneous localization and mapping) in a corridor using only the hardware of a smartphone.

Due to the low processing requirements and power consumption, our method makes 3D reconstruction of a corridor in real-time possible on a basic smartphone. With the increasing ubiquity of such devices, the techniques proposed here open up many useful applications. A visually impaired person would not need any special equipment to find their way through a building using this method; all they need to do is take out their phone to get real-time information on their heading. Such a low resource technique would allow such navigation techniques to leave the labs where they're being researched and actually be used in real world applications. Also, even a beginner in amateur robotics would have access to a smartphone and this allows even the most inexpensive of mobile robots a method by which to navigate corridors.

Development of the method on an iPhone is currently under way. Preliminary results show that single image's processing can be completed in less than a second. Furthermore, this is without the method being optimized to complete as quickly as possible. This suggests that the method can run in real-time on the iPhone. Beyond this, work has also been done to allow iPhone streaming of video data to a server. This has been done with the intention of both providing server based processing, as well as allowing human based annotations to be combined with the algorithms results. This method is currently under development.

While we have discussed the advantages of our proposed method in cases where there are low resources, the power of the proposed techniques in high resource applications should not be overlooked. Even with more advanced robotics equipment, combining real-time 3D modeling along with high speed mobility is a difficult task. On high power equipment, our technique should produce more than enough 3D models despite any high speeds. Furthermore, 3D range sensors always have distance limitations where our proposed methods have virtually no distance limit. Therefore, our method can also be used, with very little additional computation, to produce preliminary results for information beyond the range of the 3D sensors.

6. Conclusions and Future Work

In this thesis, we have presented a real-time, low resource 3D corridor reconstruction method from a single camera. We have also proposed the perspective based Hough transform algorithm which allows for the fast reconstruction found in this paper. We have shown that by exploiting the characteristics of a corridor, accurate 3D models can be generated quickly with low resources.

Combining this technique with existing 3D modeling techniques allows for both a fast result which would be ready before most other methods as well as a way to extract data beyond the conventional 3D range finders.

Our ongoing work consists of a full implementation on the iPhone. Initial speed tests of the processing show that real-time implementation can be accomplished. This work is then planned to be integrated with human annotation by providing the camera data on a server where users can give feedback. The proposal of this work has been granted a research award by the PSC-CUNY. This work will not only provide visually impaired users with guidance, but, through recording the interactions, will be able to gather information both on what needs the visually impaired have and how the person helping on the server responds. This information is useful in understanding how to advance more efficiently in future research.

7. References

- [1] Hao Du, Peter Henry, Xiaofeng Ren, Marvin Cheng, Dan B. Goldman, Steven M. Seitz, and Dieter Fox. 2011. Interactive 3D modeling of indoor environments with a consumer depth camera. ACM, UbiComp 2011: 75-84
- [2] Mariano Martín Nevado, Jaime Gómez García-bermejo, Eduardo Zalama Casanova. 2004. Obtaining 3D models of indoor environments with a mobile robot by estimating local surface directions. *Robotics and Autonomous Systems* 48: 131-143
- [3] Larsson. J., Mathias Broxvall Aless, Ro Saffiotti. 2008. Laser Based Corridor Detection for Reactive Navigation. *Industrial Robot* 35: 69-70
- [4] Jung-Suk Lee, Nakju Lett Doh, Wan Kyun Chung. 2004. Door Detection Algorithm of Mobile Robot in Hallway Using PC-Camera. ISARC 2004
- [5] Yingli Tian, Xiaodong Yang, Aries Arditi. 2010. Computer vision-based door detection for accessibility of unfamiliar environments to blind persons. ICCHP 2010: 263-270.
- [6] A. Criminisi, I. Reid and A. Zisserman. 2000. Single View Metrology. *International Journal of Computer Vision: Volume 40 Issue 2*: 123-148
- [7] Akio Kosaka, Juiyao Pan. 1995. Purdue Experiments in Model-Based Vision for Hallway Navigation. IROS. In *Proceedings of the Workshop on Vision for Robots*: 87-96
- [8] Larsson, Johan; Broxvall, Mathias. 2005. Fast Laser Based Feature Recognition. *Proceedings of the 3rd Swedish Workshop on Autonomous Robotics*
- [9] J. Forsberg, U. Larsson, A. Wernersson. 1995. Mobile robot navigation using the range-weighted Hough transform. *IEEE Robotics & Automation Magazine*: 18-26
- [10] R. Ebrahimpour. 2012. Vanishing point detection in corridors: using Hough transform and K-means clustering. *IET* 6: 40-51
- [11] Min Sun et al. 2009. 3D tunnel model reconstruction from monocular image. *Proc. SPIE 7496, MIPPR 2009: Pattern Recognition and Computer Vision, 74961A*
- [12] Olga Barinova et al. 2008. Fast Automatic Single-View 3-d Reconstruction of Urban Scenes. *ECCV 2008, 10th European Conference on Computer Vision: Part II*: 100-113
- [13] Erick Delage et al. 2007. Automatic Single-Image 3d Reconstructions of Indoor Manhattan World Scenes. *Springer Tracts in Advanced Robotics Volume 28*: 305-321

- [14] Erick Delage et al. 2006. A Dynamic Bayesian Network Model for Autonomous 3D Reconstruction from a Single Indoor Image. *Computer Vision and Pattern Recognition*, 2006 IEEE Vol 2: 2418-2428
- [15] Rong-Chin Lo, Wen-Hsiang Tsai. 1997. Perspective-transformation-invariant generalized Hough transform for perspective planar shape detection and matching. *Pattern Recognition*, Volume 30, Issue 3: 383–396
- [16] J. Canny. 1986. A Computational Approach to Edge Detection, *IEEE Trans PAMI*, 8(6): 679
- [17] Duda, R. O. and P. E. Hart. 1972. Use of the Hough Transformation to Detect Lines and Curves in Pictures. *Comm. ACM*, Vol. 15: 11-15
- [18] Matas, J., Kittler, J. and Galambos, C. 1999. Progressive Probabilistic Hough Transform for line detection. *Computer Vision and Pattern Recognition*, Volume 1: 554-560