Winter 11-27-2018

Virtual Reality as Navigation Tool: Creating Interactive Environments For Individuals With Visual Impairments

Nick Murphy

Virtual Reality as Navigation Tool: Creating Interactive Environments For Individuals With Visual Impairments

A Thesis Presented to

The Faculty of the Computer Science Department

by

Nick Murphy

In Partial Fulfillment

of Requirements for the Degree

Master of Science in Computer Science

Kennesaw State University

December 2018

Virtual Reality as Navigation Tool: Creating Interactive Environments For Individuals With Visual Impairments

Approved:

In presenting this thesis as a partial fulfillment of the requirements for an advanced degree from Kennesaw State University, I agree that the university library shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish, this thesis may be granted by the professor under whose direction it was written, or, in his absence, by the dean of the appropriate school when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from or publication of, this thesis which involves potential financial gain will not be allowed without written permission.

Nick Murphy

Notice To
Borrowers

Unpublished theses deposited in the Library of Kennesaw State University must be used only in accordance with the stipulations prescribed by the author in the preceding statement.

The author of this thesis is:

Nick Murphy

1100 South Marietta Pkwy
Marietta, GA 30060

The director of this thesis is:

Dr. Chia-Tien Dan Lo

1100 South Marietta Pkwy
Marietta, GA 30060

Users of this thesis not regularly enrolled as students at Kennesaw State University are required to attest acceptance of the preceding stipulations by signing below. Libraries borrowing this thesis for the use of their patrons are required to see that each user records here the information requested.

Virtual Reality as Navigation Tool: Creating Interactive Environments For Individuals With Visual Impairments

An Abstract of

A Thesis Presented to

The Faculty of the Computer Science Department

by

Nick Murphy

Bachelor of Arts, Criminal Justice, 2011

Bachelor of Science, Computer Science, 2012

In Partial Fulfillment

of Requirements for the Degree

Master of Science in Computer Science

Kennesaw State University

December 2018

ABSTRACT

Research into the creation of assistive technologies is increasingly incorporating the use of virtual reality experiments. One area of application is as an orientation and mobility assistance tool for people with visual impairments. Some of the challenges are developing useful knowledge of the user's surroundings and effectively conveying that information to the user. This thesis examines the feasibility of using virtual environments conveyed via auditory feedback as part of an autonomous mobility assistance system. Two separate experiments were conducted to study key aspects of a potential system: navigation assistance and map generation. The results of this research include mesh models that were fitted to the walk pathways of an environment, and collected data that provide insights on the viability of virtual reality based guidance systems.

Virtual Reality as Navigation Tool: Creating Interactive Environments For Individuals With Visual Impairments

A Thesis Presented to

The Faculty of the Computer Science Department

by

Nick Murphy

In Partial Fulfillment

of Requirements for the Degree

Master of Science in Computer Science

Advisor: Dr. Rongkai Guo

Kennesaw State University

December 2018

DEDICATION

PREFACE

We're blind to our blindness. We have very little idea of how little we know. We're not designed to know how little we know.

Kahneman*

*Kahneman, "Thinking, Fast and Slow", New York City, Farrar, Straus and Giroux, 2013.

ACKNOWLEDGMENTS

I would like to thank Dr. Rongkai Guo for his mentorship and support during my graduate studies and this thesis process

I am also indebted to the other members of the thesis committee, Dr. Kai Qian and Dr. Chao Mei, for their willingness to assist and offer guidance

I would like to thank the rest of the members of the CCSE Realities Lab for making me and my work look much better than it is:
Devan Patel, Drew Savas, Karis Kim, and Derek Martin

This thesis was made possible by the support and guidance of my mother, my uncle, and my friends

The research described in this thesis was funded in part by a grant from the National Science Foundation

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Chapter I

Introduction

The virtual, augmented, and mixed reality fields are growing at a fast pace. Thanks in part to developer interest, improvements in technology, and falling prices, immersive virtual reality experiences have become increasingly affordable at the individual level. Such consumer grade devices – from the high end Head Mounted Display headsets like the HTC Vive and Oculus Rift to more humble Google Cardboard – have made virtual environments accessible to a wider range of users.

One potential user group that is often not considered when designing virtual reality applications is people with visual impairments. While this oversight may be partially due to the relative size of that population, another potential reason is the common assumption that virtual reality means *visual* reality only. This bias ignores the variety of ways in which a virtual world can be conveyed to a user beyond that of sight. Much research has been conducted into how audio and haptic feedback can add greater detail to the virtual reality experience; for instance, Romano et al's work on simulating vibration-based textures for virtual objects and Geronazzo et al's study on providing additional distance and depth information with digital sound [20][4]. Improvements in these and other areas produce more immersive virtual environments for all users while increasing accessibility for those who have visual impairments.

In addition to entertainment, virtual environments have the potential to be used in serious applications. One such form is as an Assistive Technology – tools that help individuals with disabilities perform many activities necessary for daily living. One area of particular interest is that of mobility assistance. Navigation between locations is often a challenge for people who lack a sense of sight – many helpful cues for how to orient oneself towards an intended destination are either not available or more difficult to find. Furthermore, the surroundings itself can often change during the navigation process (such as construction, temporarily closed off spaces, or even the amount or speed of fellow travelers in the area) making it more difficult to move through an environment even if it is familiar to the individual.

This thesis presents two experiments that examine two main aspects of a potential autonomous, virtual environment-based mobility assistance system for people with visual impairments. The first study looks into how virtual reality technology can supplement the traditional training methods that aid many individuals with indoor navigation. This idea was tested with an experimental navigation system that provided verbal orientation feedback to participants as they navigated through virtual and mixed reality environments. User confidence was also recorded during and after the study to learn more about the system's viability for daily use or training. The second study explores how a traversable virtual environment could be built from user traffic pattern data within a space. This was accomplished by recording position points using BlueTooth Low Energy beacons, then aggregating, filtering, and clustering the data to extract features and meaning about the environment. The resulting navigation mesh could be used to provide an environmentally aware, automatically updating map for a mobility assistance system.

The remaining chapters of this thesis serve to support, demonstrate, and analyze the insights gained from these experiments. Chapter 2 discussed relevant related work in the fields of mobility training, assistive technology, signal-based localization, and navigation meshes. The procedures and methodology used for these experiments is stated in Chapter 3. A discussion of the findings and results comprises Chapter 4, while Chapter 5 provides the conclusion of the work. Finally, Chapter 6 contains ideas for future steps that can be taken from the project.

Chapter II

Related Works

Navigation and Orientation

In order to design equipment and training that aids individuals with visual impairments during daily mobility tasks, it is helpful to review the existing body of literature on the subject. Such skills have been taught to individuals using traditional means as well as through the use of technology, including the more recent advent of virtual and augmented reality.

Traditional Training

Long before the existence of modern computer technology, people have devised training methods to assist people with visual impairments. It is worthwhile to study some of these historical teaching techniques in order to develop more useful technological solutions. McDowell conducted a review of such teaching methods at the California School for the Blind over a period between the nineteen forties and fifties [13]. The primary system for instructing the deaf-blind students was the then popular "Tadoma Method", which utilized vocal to haptic feedback [13]. During each session, a student would place their hands on their teacher's neck and face; as the teacher spoke, the resulting vibrations would be transmitted to the pupil [13]. In this manner, sounds, words, and eventually sentences could be taught to the student [13]. Words were then physically related to objects and actions (McDowell describes teachers lifting students up and down to convey the meaning of “jump”) and students were encouraged to repeat the words they learned [13].

More modern examples of teaching efforts for individuals with visual impairments can be found in the work of Habulezi and Phasha [5]. Similar to McDowell, they focused on the curriculum of one school: The Rehabilitation and Development Trust Centre for the Blind in Botswana, the only training facility in the country for individuals with visual impairments [5]. Due to the age of the students, the school’s focus was on teaching science and trades rather than basic language and motor skills like the California school [5]. Course material was convey to students via Braille writing with support from sighted and specialist teachers; additional time was allotted for individuals to complete assignments [5].

However, researchers found that efforts to accommodate students with disabilities extended beyond the manner of teaching. For instance, students with less hearing or visual acuity were specifically seated in the front of the class, while students with sensitive skin due to albinism were placed in shaded areas away from windows [5]. Assistive technology was provided to students with visual impairments so that they could enroll in classes that were traditionally only available to sighted individuals [5]. Additionally, students were encouraged to maintain independent mobility via guided orientation training, which allowed them to memorize pathways to relevant parts of the campus [5].

Habulezi and Phasha, while commending the current methods of the facility, noted areas for improvement in their report [5]. Based on discussions with students with visual impairments, they found that many had issues with light sensitivity and were distracted by changing levels of sunlight due to their proximity to the classroom windows [5]. Due to budget limitations, assistive technology was not available for all subjects offered by the Center, thus preventing students with visual impairments from enrolling in them [5]. Additionally, students reported that faculty were often too helpful in assisting them during daily navigation, which lead to greater reliance on outside help and less personal independence in mobility [5].

It should be noted that while there are legal definitions for blindness, there is no absolute standard for “visual impairment”; members of this population can have varying degrees of visual acuity, light sensitivity, and a host of other differences. Rönnberg and Borg provide further examples of disparities within the classification in their research [21]. They reviewed studies which showed that much like the students in Habulezi and Phasha’s work, individuals with visual impairments can retain high levels of light sensitivity [21]. Such sensitivity was even found to explain the so-called extra-sensory “facial vision” that blind individuals were once believed to possess [21]. The researchers also highlighted the difference causes of visual impairment – from genetic disorders at birth like Usher’s syndrome to later in life accidents – and how the origin of a person’s blindness can impact what senses they retain access to [21].

Rönnberg and Borg believed that an individual's unique capabilities and senses should be taken into account when devising training regimens for them [21]. One example they provided was holding an inflated balloon to assist with feeling useful environmental auditory cues via vibrations [21]. The researchers also discussed different tactile learning acquisition tools such as Braille and Moon-Code [21]. They then examined the Tadoma method previously described by McDowell, and found that both new and experienced students of the system produced speech that was 60 to 70% discernible to a native speaker of the language [21]. They noted that contrary to popular opinion however, studies had yet to find conclusive evidence that an individual's sense of smell was heightened enough to allow precise identification of individuals or locations [21].

The information presented in this section served as an important guide as the experiments described in this thesis were designed and structured. As noted by Rönnberg and Borg, there are many assumptions about the capabilities of individuals with visual impairments that must be reviewed to create effective training tools [21]. In this thesis' study, decisions such as which type of feedback to utilize in the experimental navigation system were made based on findings in the above literature. Further, environmental and individual constraints were put in place in an attempt to standardize the experience for all participants.

Assistive Technology

With an understanding of traditional training methods for people with visual impairments established, it is now possible to examine how computer technology can supplement and improve on existing techniques. Examples of experimental assistive technology range from external communication devices to fully wearable navigation and object detection devices.

In Tulashvili's work, professional tools were combined with voice-activated technology to enable individuals with visual impairments to create digital content [25]. During the study, participants were provided with custom protractor rulers; angles and units of measure were denoted by Braille representations of the numbers [25]. The rulers were placed upon sturdy rubber mats, which the participants used to draft their designs. As they made measures and draw lines, participants verbally called out their instructions; these commands were then processed by speech-to-text software and used to update a map in Auto2D [25]. Participants reported that this method was intuitive and easy for them to use [25].

The methodology developed by Tulashvili required a relatively low level of integration with the user. Fernandes, Costa, Filipe and Hadjileontiadis experimented with technology on the other end of that spectrum, a stereo-vision-based navigation assistance system [3]. The proposed *SmartVision* device consisted for five modules – mounted to the participant's chest, back, and walking cane – that gathered information and user input, processed and interpreted the data, and provided output for decision making [3]. Images captured from the environment are combined with GPS, WiFi, and nearby RFID signals to help form a picture of the immediate surroundings; this knowledge is then imparted to the user via text-to-speech audio output [3]. Users can push button on the *SmartVision* controller to adjust their destination or inquiry about features in their current location [3].

Unlike the highly specialized technology exemplified by projects like *SmartVision*, Kacorri, Kitani, Bigham, and Asakawa attempted to utilize the ubiquitous mobile phone to provide an object detection and identification service for people with visual impairments. They developed a machine-learning supported app that allowed participants to create personalized lists of objects, which they could then have their camera phone identify for them [8]. One of the key areas of investigation in Kacorri et al’s study was the potential differences between lists created by sighted individuals versus those of people with visual impairments [8]. The preliminary observations showed that participants trained their systems to record more specific details such as brand, color, and flavor compared to the information in prebuilt profiles [8].

The studies described in this section demonstrate the potential for technology to have a positive impact on the daily lives of people with visual impairments. While there is existing research on technology-based orientation and mobility aids, these devices often require obtrusive or specialized equipment to be worn in order to function effectively. The experiments described in this thesis approach the challenge of providing navigation assistance using common place equipment in a similar manner as the Kacorri et al study.

Virtual Reality Environments

Virtual and Augmented Reality technology is one of the newest fields to be applied to providing aid to individuals with visual impairments. As with other assistive technologies previously mentioned, there are different levels of device and software complexity depending on the intended application.

Lahav, Schloerb, Kumar, and Srinivasan devised a relatively low impact virtual landmark system called *BlindAid* [11]. The device itself was similar to a standard walking cane; this tool allowed the researchers to convey a virtual environment via sound and haptic feedback [11]. Participant were asked to walk through a series of virtual rooms, "feeling" the texture of the virtual objects and "hearing" the sound as the cane came into contact with them [11]. As they traveled through the space, participants could place virtual landmarks throughout the room, which they could navigate back towards by following the sound of the virtual beacon that was attached [11]. The researchers discovered that despite their efforts to make the virtual world more detailed and textured, many participants preferred the simpler audio and haptic cues [11]. When interviewed, several participants stated that the more complex textures that the *BlindAid* system were confusing, and that the high level difference between simple smooth and rough textures was sufficient for navigation purposes [11].

The population that Silva de Souza, Cardoso, and Lamounier worked with required more than just orientation and navigation assistance to reach their intended destination. They conducted a case study on virtual environment navigation, focusing on an individual with visual impairments who also required the use of a wheelchair [23]. The researchers developed an external brain computer interface which could interpret basic signals from the outer layer of the cerebrum [23]. This device allowed the participant to navigate through a 3D audio virtual environment without having to physically move their wheelchair [23]. In addition to performance, researchers were also concerned about the health of the participant during the trials [23]. As the individual navigated through the virtual course, researchers monitored their heart rate, temperature, and blood pressure [23]. Silva de Souza et al found that while all three metrics were elevated after the participant had finished a session, the vital signs were still in a safe and healthy range [23].

Following in the footsteps of Lahav et al's *BlindAid* system, Zhao, Bennett, Benko, Cutrell, Holz, Morris, and Sinclair experimented with methods to convey virtual environments to individuals with visual impairments. Tests were conducted using the *Canetroller* device, which delivered haptic and auditory sensations when the walking cane extension came into contact with objects in the virtual world [29]. Additional feedback was provided via a braking mechanism worn across the chest; this module simulated resistance when the individual collided with a virtual wall or other object [29]. Using this device, participants were asked to explore two purely virtual spaces and meet certain objectives such as crossing a virtual street [29]. During the study, the researchers noted that participants moved their cane in several different patterns to explore the room, such as trailing, two point contact, and shore-lining [29]. According to participants, they switched between different cane usage styles depending on the feedback they received from from the *Canetroller* modules [29].

The above research provides a brief summary of current work with virtual reality as an assistive technology. Experiments described in this thesis sought to extend the work detailed by Lahav et al and Zhao et al into a mixed reality space; this extension would potentially allow users to navigate physical environments with the assistance of virtual world guides and feedback. Silva de Souza et al's emphasis on user health inspired this thesis's interest in examining user confidence and comfort during virtual reality navigation.

Mapping

Having reviewed the method of conveying a virtual map to a user with visual impairments, the creation of the map itself must now be discussed. Prior research has been conducted into many areas related to this task, including environmental localization, feature encoding, and data filtering and correction.

Localization

Much research has been conducted into developing autonomous awareness of surroundings, especially when that environment is difficult to access. One such example is the ocean and submarine navigation, which was the subject of studies by Wang, Chen, Gu, and Hu. At the time of their research Autonomous Underwater Vehicles required expensive and specialized equipment to orient correctly during travel; however, Wang et al felt that localization errors still occurred at an unacceptable rate [26]. They devised a cooperative algorithm where nearby submersibles could collectively locate themselves using each other and a single mobile beacon on the surface [26]. The researchers discussed how their method utilized filtering (such as extended Kalman) in combination with techniques like Moving Horizon Estimation to correct many of the unbounded localization errors they had observed in other systems [26]. During simulation testing, Wang et al found that their algorithm converged towards the ground truth much closer than either filtering or estimation methods alone did [26].

Paull, Saeedi, Seto, and Li took a different approach to the problem of Autonomous Underwater Vehicle localization. They noted that land-based localization methods like GPS and RFID were of limited use in underwater situations, and that even the use of beacons per Wang et al was too restrictive for daily operations [18]. Based on their research, Paull et al recommended the use of simultaneous localization and mapping technology to overcome these limitations [18]. As the name suggests, this technology allows AUV agents to both model its surroundings and then find its own location within the map [18]. In their report, the researchers explored several variations of the localization method; possible implementations included feature-based (which extracted environmental features and stored them for later reference) and view-based (which recorded metrics like the current trajectory and angle and tracked them over time). While Paull et al stated that the technology was in its infancy, they believed it was a significant step forward in the quest for effective autonomous underwater navigation systems [18].

Localization in environments above sea level has also been the subject of much study. Shao, Zhang, and Wang experimented with a method for individual nodes in a Wireless Sensor Network to determine their current location [22]. While reviewing existing systems, the researchers noted that each node was required to know its exact location at all times in order to allow other nodes to update themselves based off their proximity [22]. To over this restriction, Shao et al designed an algorithm that utilized auxiliary variables to estimate the current node's location and orientation, requiring only partial information from Angle of Arrival received signal measurements [22]. The researchers found that the efficiency of this method was susceptible to variance due to environmental factors; however, if the signal to noise ratio was calculated and accounted for, the algorithm produced more accurate orientation and location data than the standard triangulation methods [22].

The above research provides a preliminary understanding of localization methods utilized in different scenarios. While these studies were not focused on generating data for mapping virtual environments, many of the same principles can be applied to both tasks. Additionally, several of the concepts mentioned in these studies – such as filtering data to remove errant values – inspired the design of experiments described in this thesis.

Bluetooth/RSSI

The previous section described several types of signal-based localization methods. After review, it was determined that Bluetooth signals could potentially provide that service for the experiments described in this thesis. Once this was decided upon, it was necessary to study how other researchers had utilized Bluetooth beacons for localization purposes.

Palumbo, Barsocchi, Chessa, and Augusto took a stigmergic approach to indoor localization with Bluetooth Low Energy beacons. They contrasted their approach to the standard fingerprinting method, which compares current measures to predefined values to determine where the user is located at a given time [17]. To overcome the inherent limitations of that system, Palumbo et al developed a position marking process that produced a shareable online probability map of user positions [17]. After an initial signal strength test, guide marks were dispersed from beacons within individual result areas calculated from a MinMax localization algorithm [17]. The "strength" of a mark lessened as it moved further from its beacon's origin; this enabled the researcher's algorithm to develop bounding boxes to track the user's position inside the map [17]. Using this system, Palumbo et al were able to build a grid-based heat map of furniture and objects in a room [17].

Huang, Chang, and Chen utilized historical BlueTooth readings for their localization method. Their work took inspiration from principles of computer graphics – specifically, the concept of a "Hit Box" that determines the boundaries of sprites and models [6]. Huang et al proposed a *HitBall* algorithm that would provide a low energy and communication cost system for range-free localization within a mobile sensor network [6]. As part of the process, individual beacons would retain a limited list of prior readings from nearby mobile and stationary transmitters, in addition to the most recent signals [6]. Readings were “faded” and assigned a reduced weight in the localization calculation according to their age, which increased estimated position stability while allowing the beacon to move freely [6]. Huang et al simulated an obstacle filled course to test their algorithm; they found that their sensors were able to locate themselves effectively with only one-hop broadcasting communication required [6].

Rather than exclusively relying on BlueTooth signals for localization, Kriz, Maly, and Kozel devised an experiment to compare its effectiveness with that of WiFi. Their goal was to understand if user positioning in indoor layouts (which often have an existing WiFi coverage) could be enhanced or replaced with BlueTooth beacon technology [10]. Similar to Palumbo et al, the researchers reviewed fingerprinting methods for localization as well as the comparatively simpler (and often more variable) triangulation technique; Kriz et al ultimately determined that a modified fingerprinting approach was the best option for their study [10]. During trials, the researchers found that pure WiFi localization outperformed pure BlueTooth and BlueTooth enhanced WiFi tracking in the first few seconds of the test [10]. However, once the BlueTooth devices had “warmed up”, the BlueTooth enhanced WiFi solution had a reduced variance rate and an improved accuracy rate of 23% compared to the other methods [10].

Many studies have utilized BlueTooth signals for finding users and sensors in unknown or changing environments. The experiments described in this thesis seek to expand on this body of work, and develop navigation mesh maps from information gathered during BlueTooth-based localization (similar to the “simultaneous localization and mapping" technologies for submersibles as described by Paull et al). The above literature has also reinforced a concern previously mentioned in the general Localization section: the risk of map distortions due to intermittent signal interference.

Data Correction and Analysis

As seen repeatedly in the prior literature review, signal-based localization methods can often suffer from errant values due to natural signal fluctuation or environmental factors. Given this potential issue, it is worthwhile to conduct further examination of methods to compensate for errors and correct the resulting maps.

The Kalman filter was utilized in several other studies previously reviewed (such as Wang et al and Paull et al), so it is worthy of further explanation. Welch and Bishop gave a comprehensive overview of the filter in their 2001 SIGGRAPH lecture [27]. They described the filtering technique as an "optimal estimator", intended to reduce potential error covariance based on a set of predefined guidelines [27]. The original implementation was referred to as the *Discrete Kalman Filter*; it used prior samples in the data set to develop a probabilistic "noise profile" which it then used to clean up future received data [27]. Prior studies reviewed in this chapter utilized versions of the *Extended Kalman Filter,* which was designed to work with non-linear data processes [24]. This filter generates partial derivatives of the current state to develop a linear estimate that is easier to process using Kalman filter principles [27].

Another common filter seen in localization literature (such as that of Shao et al) is the Gaussian filter. Deng and Cahill provided information about its standard form as well as their own proposed improvement in their paper [2]. Whereas Kalman filtering attempts to adjust values based on a noise profile, the Gaussian approach focuses on removing errant values and "smoothing out" the data set [2]. The standard implementation utilizes a set window of values to develop a Gaussian curve; values outside a standard deviation threshold are considered to be errant and removed [2]. Deng and Cahill developed an extension based on these principles, the Adaptive Gaussian filter, to reduce issues with excessive blurriness in key areas of a signal [2]. Their algorithm changed its filtering behavior based on the size of the standard deviation in each window – less data was removed when the standard deviation was higher, while more data was excluded with lower standard deviation [2]. During processing, the researchers found that their altered algorithm produced sharper signal edges and a lower constant mean square error rate compared to the standard version of Gaussian filtering; however, they ultimately concluded that the algorithm was too computational intensive for practical usage [2].

While filtering data to remove or reduce errant values is helpful, it alone does not provide the necessary information to build a navigation mesh model of an area. Other studies in this literature review suggest the use of machine learning techniques to group and derive meaning from the aggregated data sets. As an example, the previously referenced study by Kriz et al used a weighted form of *k*-Nearest Neighbor clustering in their modified fingerprinting system for localization [10]. Their algorithm determined groups based on Euclidean distance from previously measured centroids; *k* number of positions were then retained and assigned a weight based on their distance from their neighbors [10]. After a set amount of iterations, the weights were used to determine the heat map reading of each remaining cluster group [10].

Altintas and Serif utilized a different variation of *k*-Nearest Neighbor clustering for their work with autonomous indoor positioning. Their fingerprinting system was trained using measured signal strength readings from defined access points [1]. Once this data was established, *k* number of centroids were deployed to build clusters around the training-defined reference points over a set number of iterations [1]. After the groupings have been solidified, their distance to the user based on signal strength is calculated and the closest clusters are then used to develop an estimate of what space the user is currently occupying [1]. Altintas and Serif observed that, as with many applications of *k*-based clustering, the size chosen for *k* had an impact on the accuracy of the results [1]. However, once their experiments had determined an appropriate value for *k* compared to the number of access points, their clustered fingerprinting system provided greater accuracy in user location than the conventional non-clustered variants [1].

K-Means clustering is an unsupervised learning algorithm which is also commonly used for cluster creation. Na, Xumin, and Yong provide an overview of this process and their suggestions for improving upon it. Unlike the previously described *k*-Nearest Neighbor algorithms, *k*-Means clusters are not concerned with an individual point's proximity to its neighbors [15]. Rather, points are grouped by their distance to the nearest randomly assigned central point; over a set number of iterations, these centroid points are repeatedly reassigned with the goal of creating the best groupings [15]. Na et al found that k-Means clustering enabled researchers to discover hidden trends in data, even when there were significant amounts of noise within the values [15]. However, the researchers felt that standard implementations were too slow and inefficient for many practical uses [15]. After studying the structure of the algorithm, they determined that a number of steps – such as recalculating distances between all centroids and data point each iteration – were unnecessary and could be minimized [15]. As a result of their changes, Na et al reported improvements in both computation time and clustering accuracy once the data set was over a certain size [15].

With signal-based localization systems, it is important to cull the errant values that are inherent to the transmission medium. That said, many methods of data clustering and visualization can overcome certain levels of variance in order to extract features and meaning from data. The previously described research provide ample examples of both data correction and analysis; these insights were essential to designing the methodology of the experiments described in this thesis.

Navigation Mesh

Having discussed the methods of collecting environmental data, and techniques for cleaning and analyzing it, all that remains is to discuss the generation of virtual maps. The SIGGRAPH 2014 lecture by Kallmann and Kapadia offers an in-depth description these maps, commonly referred to as navigation mesh models. In broad terms, the goal of a navigation mesh is to model an area and allow an agent to traverse through it [9]. It is a cell-based data structure, constructed to simplify a complex environment and define the walkable pathways within it [9]. Kallmann and Kapadia noted that design concerns for navigation meshes included how to best allow for deletions, updates, and path reconstruction due to changes in the modeled environment [9].

The subject of research by Toll, Cook, and Geraerts was how to efficiently update these navigation meshes. Toll et al found that many environmental models were kept static due to the computational cost of updating them [24]. The researchers speculated that this cost could be avoid if changes were only made in localized areas rather than repairing the entire map whenever changes occurred [24]. By treating the convex points of the mesh as partitions in a Voronoi diagram, a medial axis skeleton could be formed [24]. Toll et al used this substructure to insert and delete points, lines, and layers of 2D and 2.5D navigation mesh maps [24]. When experimenting with an Explicit Corridor Map, the researchers were able to make changes in under one millisecond per operation [24].

Olivia and Pelechano noted that many navigation meshes are manually built, and often degrade or become partially inaccessible when updated during program operation [16]. To solve this problem, the researchers devised a process for automatically subdividing existing polygons in the model into smaller convex cells [16]. These subdivided cells could then be used by local path-finding algorithms such as A* Search to avoid invalid and inaccessible areas caused by navigation mesh alteration [16]. Olivia and Pelechano that their algorithm produces a suboptimal number of cells for localized traversal; however, they were able to apply convex relaxation techniques to reduce the number of excessive polygons and points [16]. With this addition, the automated process results in maps that only required 0.67 cells per geometry notch, while running in $O(r \cdot n)$ linear time [16].

The above research presents an overview of the current state of navigation mesh models and the associated creation and updating issues. The experiments conducted for this thesis focused on automating the process of environmental mapping, albeit by relying on crowd-sourced data rather than altering existing structures like the work of Olivia and Pelechano. While some of the challenges researched by Toll et al were outside the scope of this thesis, they were noted to be addressed in potential future work for this project.

Summation

This thesis is comprised of two experiments, virtual reality navigation and dynamic map generation. While each is a distinct study, they both serve to examine the feasibility of an autonomous mobility assistance system for people with visual impairments. The virtual reality navigation experiment’s connection to the main topic is more readily apparent – it attempts to provide insight into how virtual and augmented reality environments can supplement existing training methods for daily use and route recall. In addition, the study looks into user confidence while using the system, in order to learn more about the practicality of its real world usage. The dynamic map generation study serves as the environmental information source for the mobility assistance software. This study examines the possibility of building 2D navigation meshes based on aggregated traffic collected within a space. It also seeks to learn if different preprocessing filter methods can provide more useful descriptions of the environment, by controlling for noise inherent in the Bluetooth signal-based position recording.

Chapter III

Procedure

Virtual Environment Navigation Assistance Experiment

This section is taken from the paper "Exploring Virtual and Mixed Reality Environments as Mobility Assistance for People with Visual Impairments" which has been submitted for publication as of the writing of this thesis.

Participant Description

Fifteen participants took part in this research study. All participants were adults over the age of twenty one. One participant was unable to complete the study, and an additional four participant results were discarded due to issues with their recorded data. All participants reported some level of visual impairment which ranged from full blindness to below average visual acuity; to account for the difference in ability, blindfolds were secured over the participant's eyes during the training and testing sessions.

System Description

The testing equipment consisted of custom software loaded onto an Android-based mobile device. The application supported three scenarios: training, testing (virtual environment), and testing (physical environment). During each session, the mobile device was stored in a plastic belt and secured around the participant's waist using a runner's belt. Ear-enclosing headphones were connected to the mobile device to relay the program’s auditory output. The software was controlled by a game controller connected to the mobile device via Bluetooth.

In all configurations, the user’s orientation was measured using the mobile device’s built-in gyroscope. The user’s current direction was overlaid on a map of the current environment, along with markers for one of the fifteen total checkpoints. The application continuously calculated the difference between the user’s orientation and the direction of the existing checkpoint, then verbally informed the user of what adjustment were needed. If the user needed to turn more than thirty degrees, the program said “Left” or “Right”; if the difference in direction was between thirty and fifteen degrees, the program said “Slight Left” or “Slight Right”. When the user faced within fifteen degrees of the correct orientation, the program produced a single digital tone. All prompts were repeated with a one second delay until the orientation changed. When a checkpoint was reached, the next one in the path was activated and the difference in orientation was reassessed. This behavior continued until all checkpoints were reached.

The map used for both testing scenarios corresponded to a real world hallway, while the training scenario map displayed a series of directions with no connection to the physical environment. During the training and virtual environment testing scenarios, the user’s marker was automatically advanced towards the checkpoint on the virtual map as long as they are facing within fifteen degrees of the correct direction. In the physical environment testing scenario, the user’s marker was manually moved forward by a set percentage of the total distance when the controller button is pressed. Three to five intermediate checkpoints were placed in-between each major checkpoint depending on the path length to standardize the intervals that the user walked.

The application collected and recorded the following data during both types of testing scenarios: X and Y coordinates, timestamp, current user orientation, and current checkpoint reached. This information was stored in a CSV file on the mobile device for later analysis.

Methodology

The study was conducted on the second floor of the Center for the Visually Impaired in Atlanta, GA. At the beginning of the study, participants were seated in the conference room in the testing area. Researchers read the two page consent form to the participant, outlining the study procedure, risks, and benefits.

Participants were then outfitted with the testing equipment described in the Systems Description section. A blindfold was then secured over the participant's eyes to control for different levels of visual impairment. Once the participant confirmed that all equipment was comfortably and securely fastened, the training portion of the study was started.

During training, the mobile device's software verbally prompted the participant to turn in a series of fifteen directions. At the end of each set, the participant was asked if they felt comfortable with the software and how it worked. If the participant indicated they were not yet confident, they would restart the training session for another set of fifteen checkpoints.

After the training session was successfully completed, individual participants were tested based on one of two conditions. Conditions were preassigned based on block randomization.

Condition 1: Participants completed three test sessions in the virtual environment. At the beginning of each session, participants were guided to a clear section in the testing room. The virtual environment testing software was loaded onto the mobile device. Each session was then conducted in the same manner as the training session.

Condition 2: Participants completed three test sessions in the mixed reality environment. At the beginning of each session, participants were guided to a starting point in the hallway. The mixed reality environment testing software was loaded onto the mobile device. During this session, a researcher followed behind the participant and manually updated the participant's position on the software's map using the controller. When the participant had successfully navigated the course, they were guided back to the starting position before the next session.

In both conditions, the testing equipment provided verbal instructions as described in the Systems Description section. Additionally, the participant's movements were recorded from behind by a researcher during each session. After the end of each session, participants were asked if they needed to take a short break and rest before the next session.

Upon completion of the testing session, participants were asked six questions about their physical abilities and experiences during the study. They were then compensated for their time and invited to participate in the second part of the study.

Participants took part in a follow-up session a minimum of one week after their initial session. They were guided to the starting point in the hallway used in Condition 2 and asked to walk the same path as they had in the first part of the study, based on what they could recall. As during the first testing session, participants were recorded from behind by a researcher. The testing session was concluded when the participant either completed the path, made a wrong turn, or stated that they could not remember the next movement to make.

After the second testing session ended, participants were asked three questions about their experience during the second day of the study. They were then compensated and debriefed about the purpose of the study and what their test condition was measuring.

Navigation Mesh Generation Experiment

This section was previously published as part of the paper “Recreating Virtual Environments from User Traffic Pattern” in the proceedings of the 10th International Conference on Virtual Worlds and Games for Serious Applications [30].

Data Collection

The data gathering phase of the experiment was conducted in three locations within the same building at Kennesaw State University. Area 1 consisted of two rooms inside the research lab. Its dimensions were 8.99 meters long by 9.46 meters wide. Area 2 covered the central portion and stairwell of the third floor, and measured 36.44 meters long by 19.93 meters wide. Area 3 was the hallways that make up the left side of the same floor; it had a length of 36.46 meters and a width of 13.47 meters. These locations were chosen to measure the effectiveness of the proposed method under layouts common in the target environments. A more detailed layout description of the three locations can be observed from Figures 1 – 3 below. White lines represent non-walkable area boundaries or objects, green rectangles represent doors, and blue dots signify the beacon locations.

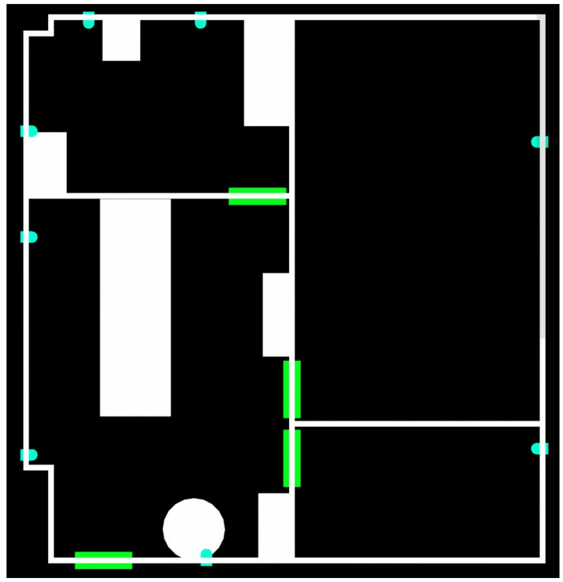

Figure 1. Area 1 Layout

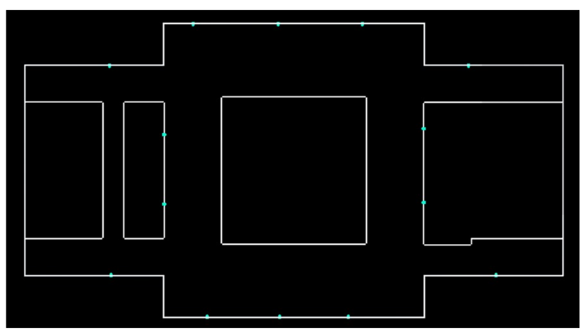

Figure 2. Area 2 Layout

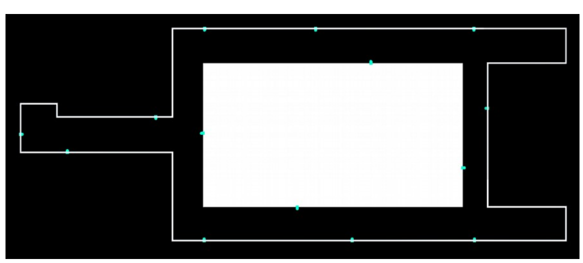

Figure 3. Area 3 Layout

Eight BLE beacons were fastened to the walls of Area 1 during the testing sessions, while fourteen BLE beacons were installed around Areas 2 and 3 at eye-level height. All beacons were set to advertise themselves every 100 ms, and each beacon had its Transmit Power set to the maximum value of 10 dBm. The placement, number, and configurations of the beacons in each area were initially chosen based on the beacon manufacturer's recommendations. Configuration tests were then conducted to assure the recommended values worked well for the test environments. When the number of beacons in an area was increased, it caused a markedly slower response time from the Estimote sensor API, while a smaller beacon count or a change in their distribution in the room caused the user's recorded position to incorrectly "pull" towards areas with higher beacon density. When the Transmit Power was lowered, the range of the beacons was decreased, which led to similar issues. Based on these prior tests, it was concluded the recommended beacon settings were the preferred ones for the sites.

Each participant was given an Android mobile device with Wi-Fi and Bluetooth capabilities. They were then moved to a location within Area 1, Area 2 or Area 3. Once there, they turned on the mobile device and the data collection application installed on it. The application asked the user to confirm which area they are in, and then began to listen for data from the Estimote API. After that, the participant walked freely throughout the area for the length of the session, which lasted on average between thirty minutes and an hour.

The following data was collected every time the Estimote API sent the mobile device an update from the beacons: current time stamp, user's x position, and their y position. The user's application then processed this raw data to generate two additional copies with Gaussian and Kalman filters. These filters were chosen as two different approaches to reduce the variation inherent to RSSI-base positioning: Kalman for its noise correction, and Gaussian for its ability to winnow the data set. The Kalman filter was initialized with a process noise value of 0.1 and a measurement noise value of 0.3, while the Gaussian filter was set to evaluate 10 readings at a time and to reduce weight on the 3 readings on either side of its curve.

Upon completion of a data collection session, the application was closed and the total recorded data was saved to the mobile device's internal storage with a time-stamped label. The data files were then pulled from the device and combined with the other recordings for that particular area. The aggregate data was then utilized for the next step in the research: visualization of the data.

Data Visualization

The collected position information was loaded into a custom-made viewing program built with the Unity3D IDE. When the data was loaded, positions were normalized on a zero to one scale against the testing area's length and width dimensions. The points were then drawn on the screen, with an overlay of the area's boundaries also displayed for initial visual comparison.

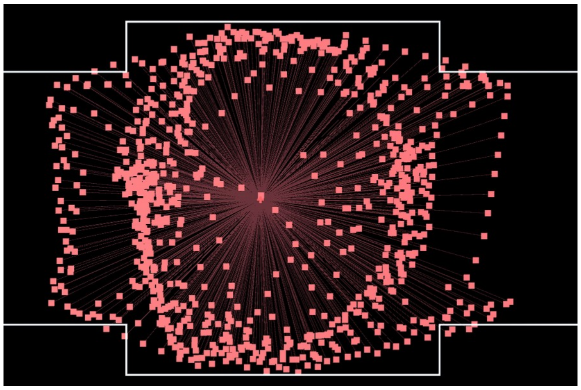

Figure 4. View of Area 2 positions with cluster size k = 1

The next step was to detect trends from the data in regards to the shape and size of the room. To do this, points were assigned into distinct groups via K-means clustering. A detailed discussion of this algorithm and its functions can be found in the previously cited work by Na et al [15]. It was determined the K-means clustering algorithm's use of Euclidean distance as a metric to group points would work well with this study's map-based data set. The time required to render usable clusters and the ease of parameter changes to test for different settings were also factors in the decision to use it.

The number of clusters was set by the user beforehand, and then the algorithm was run until the cluster centroid variation between iterations is less than a predefined threshold. In the end, it was determined that a 'k' value between 8 and 11 provided a reasonable segmentation of each area.

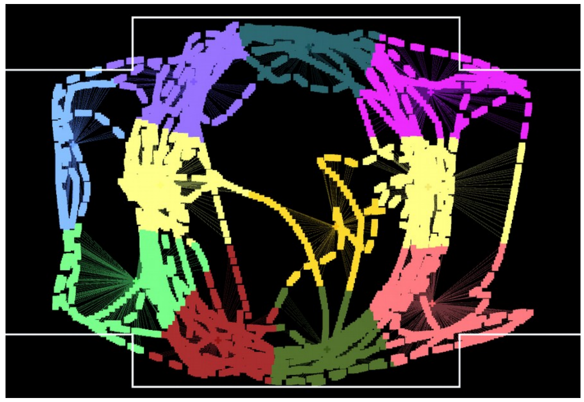

Figure 5. View of Area 2 positions with cluster size k = 11

Navigation Mesh Generation

The final process to build a walkable 2D model of the area began after the position data had been grouped into appropriate clusters. To prepare for this, a grid was defined for the area map, with grid sizes of 10 units by 10 units. This format was chosen as many navigation algorithms, such as A* Search, are built to work with grid-based navigation meshes. Additionally, the test areas are also predominantly square and rectangle shaped – and the round portions can be approximated with rough curves due to the chosen grid size.

Once the grid was set up, each cluster defined by the K-means algorithm was broken into further internal partitions with the addition of random number of additional centroids inside each of them. The same K-means implementation was used for the creation of the sub-clusters. Once the cluster movement had reached equilibrium, the secondary filtering process could begin.

After the clusters were divided into subsections, the next step was to further filter the data set to remove more of the potentially errant position points caused by signal fluctuation. This was done via a comparison of each point's Euclidean distance to the nearest second-level centroid within their cluster group. If the nearest distance was above the threshold, it was considered to be an errant position caused by signal fluctuation and was removed from the data-set.

When the initial pass had been completed, a second "box" filter was performed. A random number of grid cell sized squares were added to the map. They were placed randomly in locations where no data points had been recorded. Once created, these boxes expand outward from their center. When one of the sides of a box collided with a point, it checked the distance between that point and its sub-cluster centroid point. If the distance was above the cut-off value, the point was assumed to be from signal fluctuation and was removed from the map; the box side then continued forward. If the distance was low enough, that side of the box stopped moving. This continued until all boxes had either stopped expanding or reached the end of the map.

At this point, the navigation mesh was ready to be created. Every grid cell on the map was filled with a cell object. Each cell then calculated how many of the remaining position points were inside its boundaries and set its cell density to that number. Cells with a density of zero were then removed from the map, and only the walkable tiles remained. For visualization and comparison purposes, the cell density was used as a metric to display a heat map of the user's path. The green value was determined by the current density of a cell, divided by the maximum cell density value. This maximum density was determined by the ratio of all data points on the map with the surface area of the walkable spaces.

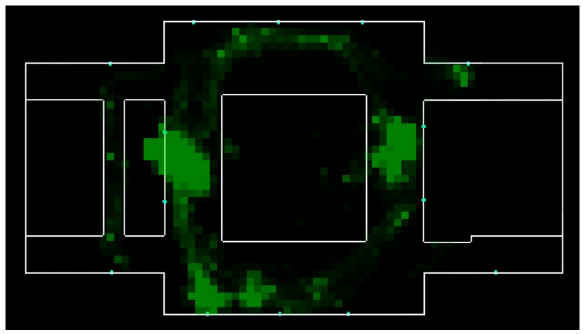

Figure 6. View of Navigation Mesh for Area 2

Chapter IV

Findings and Discussion

Virtual Environment Navigation Assistance Findings

This section is taken from the paper "Exploring Virtual and Mixed Reality Environments as Mobility Assistance for People with Visual Impairments" which has been submitted for publication as of the writing of this thesis.

Quantitative Results

Quantitative data was collected by recording participants' rotation, movement, and time during each of the three rounds in the testing sessions on day one of the study. Additionally, numerical data was collected from certain responses to the post-study questionnaires, when participants rated their level of confidence during the sessions. Testing session movement data and confidence ratings were recorded for all participants; seven participants provided confidence ratings for the second day questionnaire. One participant was unable to complete the third round of the testing session on the first day. The Kruskal–Wallis H test was chosen to analysis statistical significance of conditions results, due to the difference in condition sample sizes and the overall small sample size.

According to participant responses, only one individual had a physical disability that affected mobility. The mode value for the number of years a participant had been a navigation cane user was 1, although the range of values extended from 0 to 43. Only one participant reported using a navigation cane for either more than 1 year or less than 40 years.

Condition	Sample Size (n)	Average Completion Time	SD
1	21	5 min 9 sec	1.434
2	9	13 min 33 sec	3.384

Table I. Average Completion Time over Three Testing Rounds

Metric	Test Statistic	df	Asymptotic Significance
Average Completion Time	5.727	1	0.017

Table II. Statistical Analysis of Average Completion Time

Seven participants were tested under Condition 1, the virtual environment, while three participants were tested under Condition 2, the mixed reality environment. Each participant's total completion time was calculated by averaging their three testing rounds during day one. These values were then collectively averaged to determine the mean total completion time for each of the respective testing conditions. Based on the Kruskal–Wallis H test, there is a statistically significant difference between the average completion time of Condition 1 and 2 with a p-value of 0.017. In fact, the average time of Condition 2 is almost three times that of Condition 1, which is not surprising given the extra requirement of physically walking the through course. Additionally, the standard deviation for Condition 2 is more than twice that of Condition 1. A possible explanation for this disparity is the stronger influence each individual participant has on the final average due to the smaller sample size. This explanation is further supported by the range for each condition: Condition 1 has an individual minimum value of 3 minutes 50 seconds and a maximum of 8 minutes 33 seconds, while Condition 2 has a comparatively larger span from 8 minutes 43 seconds to 22 minutes.

Condition	Sample Size (n)	Percentage Facing Correct	SD	Avg. Pass Overs	SD
1	21	59.74	0.08	44.05	19.41
2	9	57.21	0.12	145.78	47.45

Table III. Percentage Facing Correct and Avg. Pass Overs over Three Test Rounds

Metric	Test Statistic	df	Asymptotic Significance
Percentage Facing Correct	0.117	1	0.732
Avg. Pass Overs	0.732	1	0.017

Table IV. Statistical Analysis of Percentage Facing Correct and Avg. Pass Overs

“Percentage Facing Correct” represents the average total percentage of time that the participant was facing within fifteen degrees of the correct orientation during the first day’s testing session. “Avg. Pass Overs” signifies the average number of additional turns that the participant made beyond the fifteen required to navigate through the testing course. The Kruskal–Wallis H test provided mixed results: it found no statistical significance between the two conditions for “Percentage Facing Correct”, but it did find evidence of statistical significance between them for “Avg. Pass Overs” with a p-value of 0.117. There was less than three percent difference in participants’ correct orientation time between Condition 1 and 2. Condition 2 participants on average made more than three times as many unnecessary turns during the first day's testing session. A possible explanation of this difference may be due to participants having to physically walk towards the checkpoint rather than wait for the system to automatically move them to it in the virtual environment. During transit, Condition 2 users may have shifted their direction, which required them to reorient themselves using the system.

User confidence was assessed on a numerical scale with one signifying lowest confidence and five signifying the highest confidence. All ten participants completed the same training session, so their confidence ratings were calculated together. Average participant confidence after utilizing the VR system was 3.5, with a standard deviation of 1.92. The range of confidence values is from 1 to 5, but the mode was 3 which is more indicative of the sample values.

Condition	Type	Sample Size (n)	Avg. Confidence	SD
1	Study Day 1	7	4.14	0.69
	Study Day 2	4	1.5	0.58
	Route Recall	4	2	1.41
2	Study Day 1	3	5	0
	Study Day2	3	3.67	1.53
	Route Recall	3	3.67	1.53

Table V. Avg. Confidence Rating for Day 1, Day 2, and Route Recall

Metric	Test Statistic	df	Asymptotic Significance
Study Day 1	3.571	1	0.059
Study Day 2	3.431	1	0.064
Route Recall	2.113	1	0.146

Table VI. Statistical Analysis for Avg. Confidence Ratings

While all participants completed the first day's testing session, three participants in Condition 1 opted not to complete the second day's testing session. In Table 5, "Study Day 1" and "Study Day 2" signify the participant's confidence in navigating the course during the respective day's testing session. "Route Recall" is an indication of the participants' confidence with their recall of the course layout separate from their confidence in navigating the route. For all three questions, participants in Condition 2 reported higher confidence than participants in Condition 1; however, the difference in sample size makes the numbers difficult to directly compare. Similarly, the standard deviation for Condition 1 participant responses about their confidence during the second day's testing session is smaller than that of Condition 2. The Kruskal–Wallis H test showed that none of the differences in confidence ratings between conditions reached the level of statistical significant for a p-value threshold of 5%.

Qualitative Results

Qualitative data was collected via post-study questionnaires in which the participants provided feedback at the conclusion of the first and second day of the study. All participants provided answers to the first day questionnaire while eight participants gave responses to the second day questionnaire. Two participants declined to complete the optional second day testing session.

The most consistent comment from the questionnaires was that the speed at which the verbal directions were given was too fast. Participants reported that they felt rushed to turn quickly to match the pace of the instructions. These responses are corroborated by both the researcher's observations and the recorded data; many participants became temporarily stuck because they turned back and forth too quickly and repeatedly passed over the orientation adjustments.

The speed of the verbal directions in the experiment was chosen based on survey data involving screen reader usage by people with visual impairments, which claimed that users increased the device's speaking rate to be much faster than natural speech as they acclimated to the system [12]. The results of this study, however, suggest that the speed of verbal directions required to accurately direct active physical movements may differ from the speed of screen readers for passive information processing.

Many participants also provided feedback about the semantic and acoustic clarity of the instructions. Participants reported that major directions to "turn left" and "turn right" as full ninety degree turns, were "simple" and "easy to follow," and that the overall voice feedback was clear and pleasant. These results corroborate the preferences stated by participants in Lahav et al.'s study [11].

However, many participants stated that the supplemental commands of “slight left” and “slight right” were ambiguous. Participants were unsure of how much or how little to turn based on those instructions. There was no consensus among the participants on what alternative instructions would have been more adequate for this type of slight adjustment in orientation. Several participants mentioned that they were eventually able to determine the degree of movement by “trial and error”, such as slowly turning until the verbal feedback changed.

Another finding was the impact that the test equipment had on the participant’s current usage of their senses for navigation. A few of the participants stated that the over-the-ear headphones and verbal directions from the software muffled or interfered with audial environmental cues that they typically use to navigate. Participants were particularly concerned about using the system while navigating outdoor locations like street intersections with this device because the verbal directions could prevent the user from hearing environmental cues. While the scope of this study was strictly indoor navigation, this could still be an issue in crowded or busy indoor spaces like train stations, airports, or malls.

While discussing the drawbacks of the verbal feedback delivery using over-the-ear headphones, participants suggested bone-conduction headphones. However, bone-conduction headphones may not be an improvement to over-the-ear headphones since the verbal instructions would still interfere with reception of audial cues from the environment. Another suggestion was to have haptic feedback delivered via the walking cane, similar to the *Canetroller* system detailed in Zhao et al.'s report [29]. However, other participants who had tested similar equipment found that the additional vibrations were confusing and holding the cane steadily was more challenging.

Some participants noted that the system only provided isolated directions without associating them to physical landmarks on the path. Those participants stated that the lack of connection to landmarks in the path made it harder to determine their current position in the path. While some participants were able to create their own landmarks using their senses and walking cane, this was noted as an area of improvement for the software. These findings do not lend support to the feasibility of a completely generalized navigation system with no human-defined landmarks.

On the second day of the study, only two participants were able to recall any part of the path from the first day's testing session. Of those two participants, the highest number of checkpoints reached was four out of fifteen total checkpoints. Several factors potentially could have influenced this outcome. The greatest impediment may have been the researcher's limited access to the test facility. Due to the daily operations of the Center for the Visually Impaired, the study could only be conducted on one day each week. Several participants stated that they may have better remembered the course if the second day of the study had followed the first day more closely. A related issue may have been the number of repetitions of the route with the testing equipment. As this was not a long-term study, participants only repeated the route three times on the first day, before being asked to retrace the route from memory on the second day. So, in conjunction with the length of time between the first day and second day, there may not have been enough repetitions for most participants to successfully recall the path.

Additionally, many participants noted that not explicitly being asked to memorize the route they took during the first day's testing session negatively affected their ability to remember the path later on. The researchers did not inform participants of this due to concerns that participants would rely on their memory rather than the test equipment during the first day. While potential interference may have been a valid concern, it also may have resulted in the observed under-performance on the second day of the study.

All but one participant reported at least a medium amount of confidence in their ability to navigate to checkpoints using the system after the training session. The sole participant who reported low confidence stated feeling as though they had made mistakes through the course and that they had "messed up". This likely indicates the participant was self-conscious about their performance, which may have affected their stated confidence level. Furthermore, all participants reported the same level of, or an increase in confidence while navigating after the first round of the testing session. All but two individuals reported below a medium amount of confidence for recalling the route during the second round of the testing session. An interesting note is that one of the participants who reported high confidence at recall during the second day was unable to retrace any part of the path on that day. This indicates that participants' stated confidence may be independent of their actual performance.

Navigation Mesh Generation Findings

This section was previously published as part of the paper "Recreating Virtual Environments from User Traffic Pattern" in the proceedings of the 10th International Conference on Virtual Worlds and Games for Serious Applications [30].

During this study, six sessions of data from Area 1, three sessions from Area 2, and one session from Area 3 were collected. Each data file consisted of a variable number of recordings, which ranged from 474 to 18762 participant positions. From these findings, both observation-based and quantitative results were obtained.

Qualitative Results

The first impression from the image comparison is that the Gaussian filter provides a less definitive walkable model of an area than either the Kalman filter or the raw data. It is assumed this is because the Gaussian output is a tenth of the size of the others. While this winnowing may result in less errant position data to remain in the final set, it also reduces the overall position density, even for the more stable readings. Because of this, more areas of the map are susceptible to removal by the second stage of filtering, which results in a much more incomplete navigation mesh.

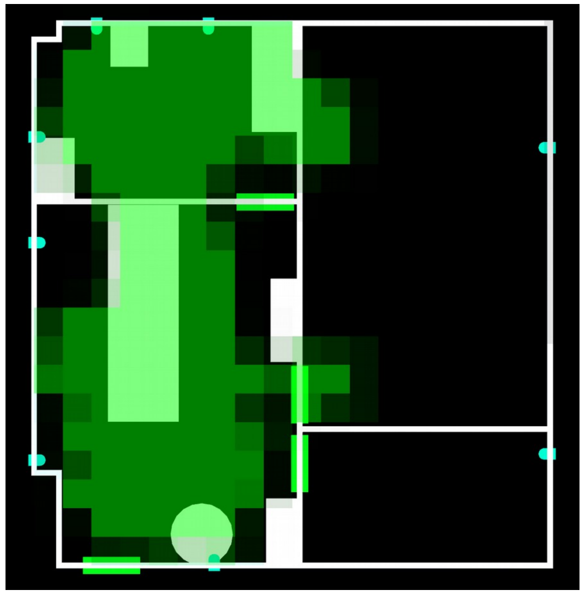

Figure 7. Area 1, Kalman Filter data set

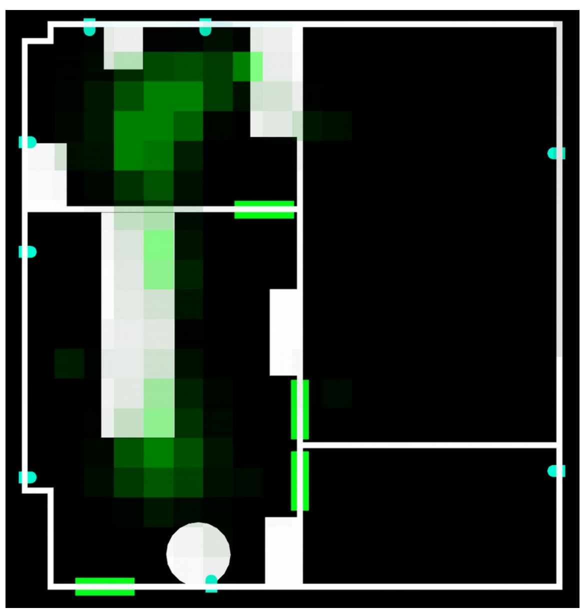

Figure 8. Area 1, Gaussian Filter data set

The proposed method seemed to provide less defining information about Area 1 compared to Areas 2 and 3, with more errant position data and “overflows” of the area’s borders and interior objects. It is assumed that this was largely because Area 1 is smaller and more confined compared to the other two. The manufacturer indicated that position data recorded by their beacons could vary by roughly one meter because of signal fluctuations. This naturally has more of an impact on smaller spaces such as Area 1, where one meter is a greater percentage of the total room space. Similarly diminished result quality in the narrower hallways on the sides of Area 2 were observed, in comparison to its more expansive central area.

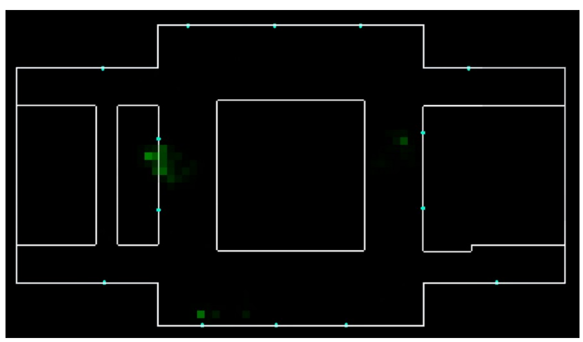

Figure 9. Area 2, Low Data Density

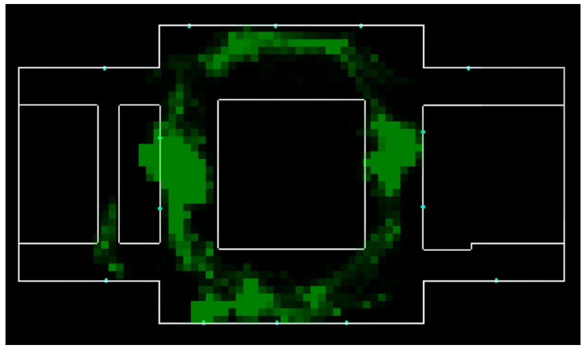

Figure 10. Area 2, High Data Density

It was also observed that the more position data collected during a session, the more defined the final navigation mesh turned out to be. When Area 2's mesh which covers 858 readings is compared with the one made from 11815 readings, only two short paths are visible in the less data dense map. On the other hand, the entire central area and parts of the left and right hallways are clearly defined when the full set of data is included in the mesh generation. This is in part due to the method of filtering data points – less dense areas (such as the hallways) are more likely to be considered "errant" values and thus removed. This risk can be mitigated by the collection of more readings, which increases the density of those spots on the map and makes them less susceptible to deletion during the data clean-up process.

Quantitative Results

The following table provides quantitative comparisons between each area's actual walkable space and the navigation mesh's ability to cover it. "Ideal" maps of the three areas were created with all walkable areas filled in. Values were then found via conversion of the ideal area map and the navigation mesh into two color images, where green represents walkable space, and black indicates non-walkable space. Pixels that occupy the same X and Y position on both images were compared to calculate the ratio of matching to non-matching sections. A summary of the results follows. “Total Error Rate” represents the percent of invalid positions the final navigation mesh covers. “Percentage from Excess” is the percentage of those errors caused when the mesh covers non-walkable areas. “Percentage from Lack” is the percentage of errors when the navigation mesh does not include walkable positions of the map.

Filter	Total Error Rate	Percentage from Lack	Percentage from Excess
None	31.34	18.65	81.35
Gaussian	54.52	56.02	43.98
Kalman	28.81	12.84	87.16

Table VII. Selected statistics for Area 1

These findings gave a different perspective on the visual observations. Area 1’s total error rates (except those from the Gaussian filter) were far lower than what was expected. When the map was compared with the statistics, it was determined that the total map coverage was able to balance out the interior errors. As mentioned in the Observations section, Area 1 is much smaller than either Area 2 or 3. Because of the confined space, it was much easier for the data to fully cover a large portion of the room. The trade-off compared to the larger rooms is the loss of interior details – the non-walkable objects such as tables were covered by the navigation mesh in the same way as the walkable spaces around them. Since the majority of the area’s space was walkable, the over coverage of the objects inside it had a smaller impact on Area 1’s score. In practice however, Area 1’s lower error rate is less useful to this work than the more defined pathways shown in Areas 2 and 3.

Density	Total Error Rate	Percentage from Lack	Percentage from Excess
Low	54.25	98.57	1.43
Medium	26.03	87.96	12.04
High	23.76	86.97	13.02

Table VIII. Selected statistics for Area 2

Area 2’s error rates also contradicted initial expectations. While the medium and high data density maps had acceptably low error rates, the initial expectation was that the error rate for the low data density map would have been much higher than 54%. This is explained by the significant amount of “empty” space in the center and sides of the map.

Since the low-density map provides relatively little coverage of the entire area, its total error rate benefits from the fact that large portions of the map should not be covered. Its lesser coverage in comparison to the other, more data rich maps is revealed in the error rate breakdown. The low-density map has approximately 98% of its total errors due to lack of coverage, almost the complete amount.

	Total Error Rate	Percentage from Lack	Percentage from Excess
Map	36.23	48.19	51.90

Table IX. Selected statistics for Area 3

The total error rating for Area 3 during the trials were higher than all but the worst performing maps in Areas 1 and 2. This was an interesting outcome and prompted further inquiry into the original navigation mesh. Visually, the model conforms to the general shape of the environment, in particular the corridor on the left side of the map. However, the mesh has a tendency to drift into the non-walkable center of the hallway, especially at the corners where participants were going around the curves. Because this drift causes walkable parts to not be included in the mesh, the ratio of over-to-under coverage in the Area 3 map is approximately equal.

Despite the higher error rating, however, the outline of the mesh provides a reasonable picture of the location. This raises the possibility that a greater focus on the definition of key unique aspects of an area, rather than attempts to extensively record the location in its entirety, may also produce good results for navigation meshes.

A final interesting note is that for the Area 2, the least error prone map, the vast majority of the navigation mesh errors occur when the navigation mesh did not cover enough area rather than too much. The proposed method uses a fairly aggressive filtering process in order to catch more of the errant values caused by signal fluctuation. These statistics indicate that it may be worthwhile to re-evaluate the strength of those filters; the under coverage of the map noted in the figures may be because relevant information does not survive the winnowing process.

Chapter V

Conclusion

Response to Reviews

Portions of this thesis were presented at the 10th International Conference on Virtual Worlds and Games for Serious Applications, and published in the proceedings under the title "Recreating Virtual Environments From User Traffic Pattern". Several comments and critiques of the work were made during the conference. Questions related to the subject matter itself will be addressed in this section.

At the conference, a question was asked about the potential use of video surveillance footage for generating navigation mesh models of environments. While this idea is an interesting concept, a computer vision approach to navigation mesh generation is outside the scope of this experiment. Additionally, it has several possible limitations compared to the method described in this thesis. For instance, existing stationary cameras may not provide complete coverage of an area, which could create constant blind spots in the navigation mesh map. Furthermore, having to correct for the difference in each camera's viewing angle adds another level of complexity to the feature extraction process with little apparent gain.

Another question was regarding the ability to gather and display information about the user's elevation within the environmental map – specifically, how to detect what floor they were on. While all experiments for this thesis were conducted on one floor only, it is possible to extrapolate how this model could be applied to a 3D space. The simplest option would be to treat each floor of a building as a separate 2D layer, and then switch which one is assigned to a user based on their proximity to that floors Bluetooth beacons.

Final Thoughts

As previously stated, the focus of this thesis was to determine the feasibility of an autonomous, virtual environment-based mobility assistance system. Two separate experiments were devised to examine the key components of such a system: dynamically generated navigation mesh models of an area, and a navigation system that could successfully convey that information to a user with visual impairments.

Participants in the navigation experiment were able to properly orient and make their way to the destination by utilizing the provided verbal feedback in addition to their own senses. While the route memorization results were not as hoped, the qualitative data collected from participant interviews provided valuable insights which can be used to improve the navigation system for future studies.

The dynamic map generation experiment was able to produce navigation meshes for three distinct area layouts. While the map of the smaller room was less useful for navigation due to indiscriminate markings of interior furniture objects, the model of the larger environments provided a reasonable impression of the shape and dimensions of the space. Furthermore, the results suggested that the fit of the map would scale with increases in the amount of aggregated position data collected.

The findings of each study provide valuable insights about the potential of a completed system. By utilizing a virtual map that automatically builds itself based on an area’s traffic patterns, a mobility assistance system can provide users with useful updates about their surroundings. This information can allow individuals with visual impairments to more safely navigate through dynamically changing environments without having to wait for humans to manually repair the navigation mesh map.

Chapter VI

Future Work

Having separately shown the feasibility of the virtual navigation and map generation processes, the next step would be to combine the two aspects into one system for a comprehensive study. This would remove the need for researchers to manually update the participant's position in the mobility assistance system during operation, potentially improving the quality of data collected during the trial. Based on survey feedback from participants, this addition would also make the equipment be more in line with their expectations of a "real-time" environmental and current location aware system.

There is potential to further develop the individual components of this future system. As part of a long term study, different types of feedback for the navigation assistance system could be included as options alongside verbal speech. This would allow for further exploration of information conveyance methods discussed in Chapter 2 of this thesis, such as haptic vibrations or digital tones, and their relative effectiveness for this task.

Further exploration in the dynamic map generation experiment could address some of the issues described in Chapter 2 of this thesis. The current method for saving, storing, and visualizing the aggregated position data could be moved to an online database, allowing offsite access to anyone with a mobile device. This in turn would allow for larger and more frequent collection of user traffic patterns. As a result, more testing could be conducted on how to update the navigation mesh over time in response to higher or low traffic in sections of the map.

References

[1] Altintas, B. and T. Serif, *Improving RSS-Based Indoor Positioning Algorithm via K-Means Clustering.* Wireless Conference on Sustainable Wireless Technologies 2011.

[2] Deng, G. and L.W. Cahill, *An Adaptive Gaussian Filter For Noise Reduction and Edge Detection.* Nuclear Science Symposium and Medical Imaging Conference, 1933. **3**.

[3] Fernandes, H., et al., *Stereo vision in blind navigation assistance.* World Automation Congress, 2010: p. 1 – 6.

[4] Geronazzo, M., F. Avanzini, and F. Fontana, *Auditory navigation with a tubular acoustic model for interactive distance cues and personalized head-related transfer functions.* Journal on Multimodal User Interfaces, 2016. **10**(3): p. 273-284.

[5] Habulezi, J. and T.N. Phasha, *Provision of learning support to learners with visual impairment in Botswana: a case study.* Social and Behavioral Sciences, 2012. **69**: p. 1555 – 1561.

[6] Huang, J.F., G.Y. Chang, and G.H. Chen, *A Historical-Beacon-Aided Localization Algorithm for Mobile Sensor Networks.* IEEE Transactions on Mobile Computing, 2015. **14**(6).

[7] Izadi, S., et al., *KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera.* Proceedings of the 24th annual ACM symposium on User interface software and technology, 2011.

[8] Kacorri, H., et al., *People with Visual Impairment Training Personal Object Recognizers: Feasibility and Challenges.* Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 2017: p. 5839-5849.

[9] Kallman, M. and M. Kapadia, Navigation Meshes and Real-Time Dynamic Planning for Virtual Worlds. SIGGRAPH 2014.

[10] Kriz, P.., Maly, F., Kozel, T., *Improving Indoor Localization Using Bluetooth Low Energy Beacons.* Mobile Information Systems, 2016. **2016**: p. 11.

[11] Lahav, O., et al., *A Virtual Environment for People Who Are Blind – A Usability Study.* Journal of Assist Technology, 2012. **6**(1): p. 2 – 21.

[12] McCarthy, T., J. Pal, and E. Cutrell, *The "voice" has it: screen reader adoption and switching behavior among vision impaired persons in India.* Journal of Assist Technology, 2013. **25**(4): p. 222-229.

[13] McDowell, F.J., *Methods and problems in educating the deaf and blind and a proposed study.* Graduate Student Theses, Dissertations, & Professional Papers, 1956. **6029**.

[14] Mundy, J.L., *The relationship between photogrammetry and computer vision.* Integrating Photogrammetry Techniques with Scene Analysis and Machine Vision, 1993: p. 92-105.

[15] Na, S., L. Xumin, and G. Yong, *Research on k-means Clustering Algorithm An Improved k-means Clustering Algorithm.* International Symposium on Intelligent Information Technology and Security Informatics, 2010. **3**.

[16] Olivia, R. and N. Pelechano. Automatic Generation of Suboptimal NavMeshes. Motion in Games: 4th International Conference. 2014.

[17] Palumbo, F., et al., *A stigmergic approach to indoor localization using Bluetooth Low Energy beacons.* IEEE International Conference on Advanced Video and Signal Based Surveillance 2015. **12**.

[18] Paull, L., et al., *AUV Navigation and Localization: A Review.* IEEE Journal of Oceanic Engineering, 2014. **39**(1).

[19] Rezazadeh, J. and A.S. Ismail, *Superior Path Planning Mechanism for Mobile Beacon-Assisted Localization in Wireless Sensor Networks.* IEEE Sensors Journal, 2014. **14**(9).

[20] Romano, J.M. and K.J. Kuchenbecker, *Creating realistic virtual textures from contact acceleration data.* IEEE Transactions on Haptics, 2012. **5**(2): p. 109-119.

[21] Rönnberg, J. and E. Borg, *A review and evaluation of research on the deaf-blind from perceptual, communicative, social and rehabilitative perspectives.* Scandinavian Audiology, 2001. **30**(2): p. 67-77.

[22] Shao, H.J., X.P. Zhang, and Z. Wang, *Efficient Closed-Form Algorithms for AOA Based Self-Localization of Sensor Nodes Using Auxiliary Variables.* IEEE Transactions on Signal Processing, 2014.

[23] Souza, E.S.d., A. Cardoso, and E. Lamounier, *A Virtual Environment-Based Training System for a Blind Wheelchair User Through Use of Three-Dimensional Audio Supported by Electroencephalography.* Telemed J E Health, 2018. **24**(8): p. 614-620.

[24] Toll, W. G. v, A. F. Cook IV, and R. Geraerts. A Navigation Mesh for Dynamic Environments. Computer Animation and Virtual Worlds. 2012. 23(6).

[25] Tulashvili, I., *Method of training blind students to the construction of graphic images with the application of Autocad 2D* Edukacja-Technika-Informatyka. **5**(1): p. 546 – 550.

[26] Wang, S., et al., *An optimization based Moving Horizon Estimation with application to localization of Autonomous Underwater Vehicles.* Robotics and Autonomous Systems, 2014. **62**: p. 1581 – 1596.

[27] Welch, G. and G. Bishop, *An Introduction to the Kalman Filter.* SIGGRAPH, 2001. **2001**(8).

[28] Yang., S.C. and Y.C. Fan, *3D Building Scene Reconstruction Based on 3D LiDAR Point Cloud.* 2017 IEEE International Conference on Consumer Electronics, 2017.

[29] Zhao, Y., et al., *Enabling People with Visual Impairments to Navigate Virtual Reality with a Haptic and Auditory Cane Simulation.* Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 2018: p. 1 -14.

[30] N. Murphy, D. Patel, D. Savas, D. Martin, C. Mei and R. Guo, "Recreating Virtual Environments from User Traffic Pattern," 2018 10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)(VS-GAMES), Würzburg, Germany, 2018, pp. 1-4.