Path Guiding for People with Visual Impairments for Short-Range Rendezvous with a Static Target

Jirachaya "Fern" Limprayoon
jlimpray@andrew.cmu.edu

Fall 2021 - Spring 2022

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

SCS Faculty Advisor:
Aaron Steinfeld
steinfeld@cmu.edu

Submitted in partial fulfillment of the requirements for CS Senior Honors Thesis.

Abstract

Navigating in complex indoor spaces like train stations, shopping malls, and university buildings can be challenging for people with visual impairments, since current path planning approaches rely on users to spontaneously adjust their trajectories based on their visual knowledge. To help a user with a visual impairment rendezvous with a static object (e.g., a waiting robot, a kiosk, or an autonomous car at the curbside) in a seamless way, we propose an approach that can incorporate information about the user's position and orientation from sensors attached to the environment, nearby robots, or cars to dynamically plan and give users instructions of how to reach their goal.

Acknowledgments

This work was funded by the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR grant number 90DPGE0003). NIDILRR is a Center within the Administration for Community Living (ACL), Department of Health and Human Services (HHS). The contents of this paper do not necessarily represent the policy of NIDILRR, ACL, HHS, and you should not assume endorsement by the Federal Government. We would also like to thank the O&M Specialist for the invaluable feedback.

Xiang Zhi Tan provided continuous and extensive mentoring and technical advice during this work. The work related to rendezvous with a mobile target was done in collaboration with Prithu Pareek (Chapter 3). Additional thanks to Elizabeth Carter, Samantha Reig, and Katherine Shih for their assistance on this work.

Contents

Chapter 1

Introduction

Navigating alone in unfamiliar, complex, indoor spaces like airports, shopping malls, and university buildings can be challenging for people with visual impairments. Since a sighted guide might not always be available in all public indoor spaces when requested, many researchers have investigated the use of mobile service robots for providing navigational assistance [3, 7, 13, 15]. Although these mobile robots are helpful in helping user navigate public spaces, discovering and rendezvousing with these robots without relying on visual knowledge can be challenging. Before the mobile robot can guide the user, the mobile robot needs to approach and rendezvous the user in a safe and socially appropriate manner.

Even when the user can locate the service robots, they might need to end the navigation early since the hallway to the goal is too narrow for both the user and the robot to go through or the robot cannot exit the building or enter certain spaces, so the user might need to navigate to the target by themselves. Finding and rendezvousing with a static object is not a difficult task to do when users can see the physical target since they can visually search the scene in a top-down manner [16]. However, the task becomes much more challenging when they are blind or have low vision. Locating a vehicle without relying on visual knowledge, even when there is a driver at the wheel to communicate with, is known to be the most difficult process for people with visual impairments when using ridesharing services [19]. This is expected to be even more difficult when finding an autonomous vehicle (e.g., Figure 1.1).

Although researchers have explored how to use smartphones [1, 6] to provide path guiding towards buildings or rooms, these systems are not precise enough to solve the *last three meter problem* (the *last-mile problem* is a transportation term for the path from a transit stop to a destination). This is an important piece to help people with visual impairments navigate and explore unfamiliar spaces independently whether that be finding the right ticket machines to top-up their ticket at train stations or finding an autonomous vehicle at the curb. We believe leveraging sensing already present in the space or object, such as the sensors on an automated vehicle the user is trying to find or surveillance cameras in a transit station, can help bridge this gap.

In this thesis, we investigated how to solve the last three meter problem, the problem of helping a user with visual impairment rendezvous with an object. We explored the conditions of rendezvousing with a moving robot (Chapter 3) and static object (Chapter 4). Our system perform most of the localization and pinpoint the user's exact coordinates, identify the correct

Figure 1.1: A person who is blind navigating to the correct door of a car (Google, 2012)

path, and guide either the mobile robot or the user to their target. To evaluate our systems, we conducted qualitative analysis with experts and simulator output.

Chapter 2

Related Work

Since a sighted guide might not always be available in all public indoor spaces when requested, many researchers have investigated the use of mobile service robots for providing navigational assistance [3, 7, 13, 15]. One of the challenges is how the user will rendezvous with these robots. Some prior studies begin the evaluation with the user already holding on to the mobile robot [15]. Other approaches rely on the users spontaneously adjusting their trajectories based on their visual knowledge of the robot's trajectory when rendezvousing with it [5, 17], but this method might be unsuitable for users who are blind or have low vision.

In order to help a user with a visual impairment rendezvous with the robot in a seamless way without relying on the user's visual knowledge of the robot's orientation and position, the robot should perform most of the localization to pinpoint exactly where the user is in the space and drive up close to them. Furthermore, the robot needs to approach the user in a way such that the user can easily rendezvous with it without startling or creating additional difficulty for the user. Recent participatory design research suggests that one effective way to initiate an interaction with a mobile navigation robot is to let a user with a visual impairment summon it through a smartphone application when they need escort assistance upon an arrival to a new indoor space [3]. Then, the robot should respect the user's autonomy and independence by allowing them to have ultimate control of the overall interaction [3].

However, service robots might not always be available or reachable to all spaces (e.g., smaller corridors, cluttered office). This led researchers to explore how to use smartphones to provide turn-by-turn navigation to help people with visual impairments navigate more autonomously in unfamiliar and complex spaces [1]. However, such systems often guide users to a general location (e.g., office of the person) instead of specific object such as a specific kiosk or vehicle. This is especially important when there are multiple similar objects to the target object in the scenes. For example, there may be multiple ticket kiosks with similar physical attributes at a public transportation station (e.g., Figure 2.1), but they might have different usage constraints like cash-only, cards-only, or coin-only. Thus, finding and reaching the most suitable kiosk becomes more complex for people with visual impairments.

There are existing systems that are designed to help people with visual impairments find objects and navigate their hands to objects at a smaller scale [4, 8]. For these, the user turns on their smartphone camera to capture images and perform object recognition [8]. Since these systems [8, 18] can only tell whether the object is in the field of view of the user's phone camera,

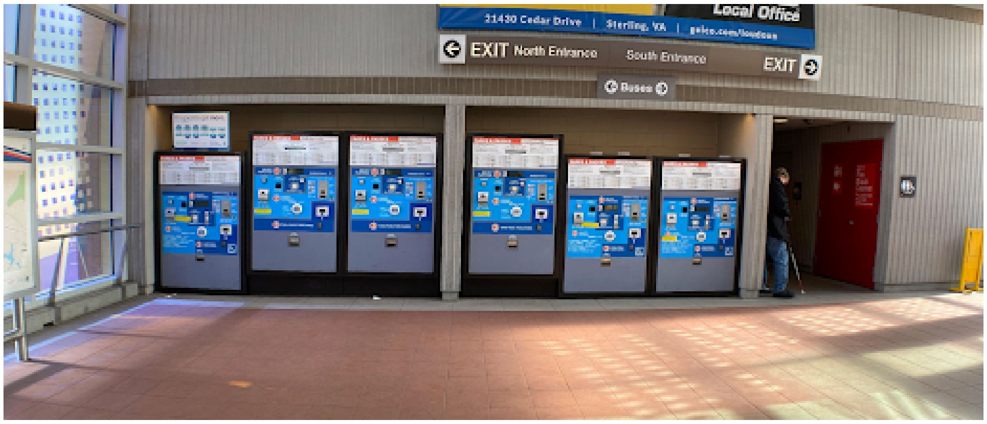

Figure 2.1: Finding one of the shorter kiosks from this camera position can be challenging.

they cannot be extended to guide users with path-finding to approach kiosks or meeting a vehicle that is our of field of view or occluded.

This led researchers to explore the use of crowdsourcing to provide real-time path guiding navigation instructions [2, 4]. However, this process removes the autonomy of users with visual impairments who might want to explore the scene without relying on another person. More importantly, they might be in a private space where they might not want to video-call and let the crowdworker see the physical space. Therefore, our research explores how to provide more detailed and adaptive path-guiding information to people with visual impairments without crowdsourced assistance. We assume the system (1) knows the position of the desired target as well as other similar objects in the scene and (2) continuously tracks the user and the dynamic obstacles from sensors already present in the scene. In the following chapters, we examined the rendezvous problem in two conditions, when the device with sensors moves (a traditional robotics scenario) and the user guided by instructions derived from static environmental sensors. In general, we hope to understand (1) how a mobile robot should approach a person with a visual impairment and (2) how a person with a visual impairment should be guided to a stationary robot, kiosk, or autonomous vehicle in a socially seamless manner.

Chapter 3

Rendezvous With Mobile Sensing

The work described in this chapter was started prior to the thesis [11]. It is included here since aspects of this work were updated and extended in support of the thesis effort.

3.1 Introduction

While prior work has explored how robots can detect target users [20], one of the most important design questions that has yet to be answered is how the robot should travel to the person when they are within a few meters of the robot. In this chapter, we examined how a mobile robot should approach and rendezvous with a user with visual impairment.

3.2 Implementation

Our system starts after the user summons the robot via their mobile phone when they arrive at a public indoor space. Assuming the robot is within a few meters of them, the app asks the user to position their smartphone screen displaying a fiducial marker facing outward at their desired robot handle position. The marker is then used to label the position and orientation of the user's preferred hand, which is used to calculate the robot's goal position and orientation.

After we establish the goal position and orientation for the robot, we use motion planning with hand-tuned parameters to help the robot approach the user seamlessly. Our navigational system is based on the ROS navigation stack [12], which plans and controls the robot in two stages. It uses a global planner that plots a path to the goal in a predefined map followed by a local planner to execute the movements on the robot while accounting for any obstacles missing from the map. The default global planner does not account for the robot's orientation and often results in the robot turning in place at the goal position (the second column from the left in Figure 3.1).

Allowing the robot to turn in place close to the user poses a serious safety risk of bumping the person's hands and/or knocking their cane over. However, making the turn radius too wide can limit the robot's ability to get to the user in smaller corridors. In order to remedy this problem, we invalidate turning in place motion on a global planner level to discourage the robot from turning in place for most of the time while allowing the local planner to take over and turn the

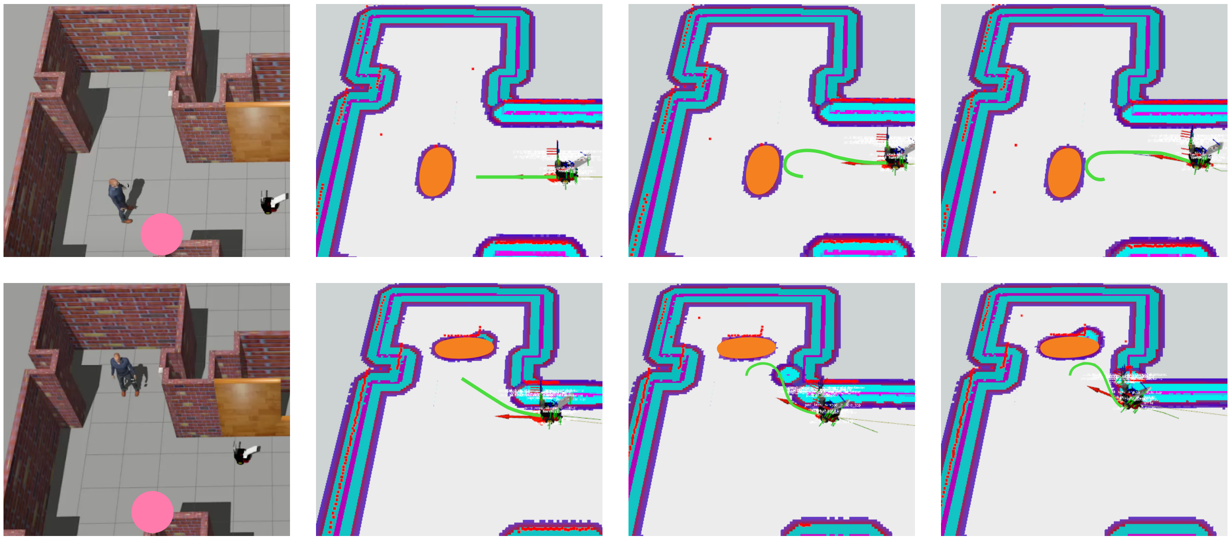

Figure 3.1: The leftmost images show the scenes from the simulator where the participant imagines themselves standing at the pink circle to witness the interaction. The other six images show the path planned by the different global planners for both scenes, from left to right: baseline, narrow, and wide. Notice that there is no visible turn on the path planned by the baseline approach since the baseline global planner does not account for the orientation at goal state, causing the robot to turn in place when it reaches the goal position when the local planner takes over.

robot in place when necessary. As such, we would like to plan for the navigation of the robot with orientation in mind in $< x, y, \theta >$ space.

For simplicity, we assume that the environment is static. We also fine-tune the ROS navigation's local planner parameters to force it to follow our global plan more closely when possible.

To control the turning radii and plan with orientation in mind, we use the lattice graph-based Anytime Repairing A* (ARA*) global planner [10] in order to explicitly hand-tune the different turn radii that our non-holonomic robot can perform. The state space of the lattice graph is constructed by applying a set of specified turn radii, called motion primitives, as a transition from each state. Then ARA* is applied on the lattice graph to get a trajectory from a starting state to a goal state. We use ARA* since it allows the planner to output a sub-optimal solution if an optimal solution cannot be found within 5 seconds.

To generate motion primitives for the lattice planner, we use the SBPL [9] library with a discretization of 0.05m. We created two different sets of motion primitives: one that encourages the robot to make a wide turn and another that restricts the robot to a narrow turn while still preventing it from turning in place.

3.3 Evaluation

3.3.1 Expert Feedback Procedure

Due to the COVID-19 pandemic, we were unable to conduct in-person usability testing. Our initial design was based on our previous experience and research with people with visual impairments [14]. To obtain feedback for our prototype, we conducted a 30-minute semi-structured interview online with an Orientation & Mobility (O&M) Specialist. At the beginning of the interview, they were asked about their usual method to approach a person with a visual impairment in general and their definition of a socially acceptable way to approach a person with a visual impairment before beginning sighted escort assistance for navigation. They viewed two sets of simulated videos of the robot approaching someone in a wide hallway (the two images in the left-most column of Figure 3.1). In one set, the user called the robot from a corner of the room, and in the other set, the user called the robot from the middle of the room. Each set had three videos showing three different trajectories generated from the baseline ROS default global planner and our ARA* planner with two different sets of motion primitives. After each video, the expert was asked to imagine that they physically witnessed the interaction and rate their agreement on a 5-point scale with the following six statements:

1. This robot behaves in a socially acceptable manner.
2. I would interrupt this robot before it reaches a person with a visual impairment.
3. This robot makes me uncomfortable.
4. I would recommend people with visual impairments call for this robot whenever it is available in a place they are unfamiliar with.
5. I would recommend the person with a visual impairment call for robot assistance only when human assistance is not available
6. I would discourage the person with a visual impairment from using this robot.

After each set, they were asked to provide qualitative comparisons among the three different videos. At the end of the interview, they were asked about what the robot should do after arriving within the person's reach and their overall feedback for the system.

3.3.2 Findings

The expert's ratings to the six statements did not differ across the videos. They agreed with statements 1 and 4 that the robot behaved in a socially acceptable manner and would recommend it to people with visual impairments if available. They disagreed with statements 2, 5, and 6 that they would interrupt this robot from assisting people, advise people to only use the robot when no humans were around, and actively discourage people from using this robot. They also strongly disagreed that the robot makes them feel uncomfortable. They explained that their expectations for robot guide assistants' behavior are different from human assistants' in that their opinion depends more on whether the robot respects the person's personal space and approaches the person in a smooth trajectory than on whether the trajectory is unobtrusive or human-like. The expert felt that the difference between the two paths generated with different radii is subtle, but the differences between the baseline path and the other two paths are much more noticeable. They believed that the planner that allows the robot to start reorientation earlier when possible provides a less obtrusive path to the user than the one that forces the robot to reorient itself at the goal position. This was because the robot spends less time in the user's personal space, lowering the chance of crashing into the person.

The O&M specialist shared that they usually approach a person with a visual impairment

Statement	In the corner			In the middle of the room		
	baseline	narrow	wide	baseline	narrow	wide
1 (Socially acceptable)	4	4	4	4	4	4
2 (Would interrupt robot)	2	2	2	2	2	2
3 (Uncomfortable)	1	1	1	1	1	1
4 (Recommend for unfamiliar place)	4	4	4	4	4	4
5 (Recommend only when no human)	2	2	2	2	2	2
6 (Discourage using the robot)	2	2	2	2	2	2

Table 3.1: Ratings from Expert (1 = Strongly disagree, 2 = Disagree, 3 = Neutral; 4 = Agree; 5 = Strongly agree).

from the side and tap them on the shoulder before introducing themselves. They then ask which arm the person with a visual impairment prefers to hold onto and which side the person prefers them to be on. Then, they usually bump their arm into the person with a visual impairment's to allow them to determine where that person should hold, typically right above their elbow. In general, the expert recommended that a guide assistant should indicate that they are talking to the person with a visual impairment (e.g., "by tapping them on the shoulder") and then introduce themselves before beginning the navigation together to make the interaction socially appropriate. They further suggested that the assistant should also be courteous of the user's primary navigation guidance tool, whether it is a guide dog or a white cane.

The O&M specialist liked that the robot respects the independence of the user by letting them grab onto its handle, which also allows the user to let go of the robot whenever they want. They also appreciated that the robot allows the user to use their primary guidance method (e.g., a guide dog or a white cane) alongside the robot. This allows the person to begin their navigation with the robot and continue to navigate by themselves when they reach their destination. More importantly, this feature allows the person to let go of the robot and continue navigating on their own if the robot malfunctions.

The expert expressed confusion when the robot approached the person from the front rather than the side unlike a trained sighted assistant. However, they said that people may adapt to this behavior if they use it on a regular basis. They also noticed that the robot did not consistently stop on one particular side of the person. They were worried it might be confusing for the user and noted that it is crucial that the robot stops at a consistent spot agreed upon with the user and then signals the user when it is ready to be gripped.

They commented that the speed shown in the video seemed slow and safe but that they would want to know more about how that speed would be perceived in a physical space, how the robot works, and how much noise it makes, because those factors cannot be fully grasped through simulation and they impact whether the interaction is socially appropriate. Overall, they found our prototype to be promising and believed it has the potential to allow people with visual impairments to travel in an unfamiliar indoor environment more independently.

3.4 Summary

This chapter described proof of concept robot functionality that helps people with visual impairments rendezvous in an indoor space with a mobile navigation robot in a socially appropriate and seamless fashion. We used a lattice graph-based ARA* planner as a global planner to encourage the robot to approach a person with a smooth trajectory. We conducted a semi-structured interview with an O&M specialist who provided feedback about our system to better understand the design considerations for future robotic escort systems. Specifically, the O&M specialist advised that the robot should approach the user from the side rather than the front, and it should approach the person consistently from the side they indicated. While the expert provided valuable feedback on our initial design, the insights from that interview do not necessarily reflect how a person with a visual impairment would use the system in real-life. To address this, we plan to further test our system in-person with users with visual impairments. In the next chapter, we discuss how we extended this for the inverse approach where the system helps people with visual impairments rendezvous with stationary objects, like autonomous vehicles. For example, the vehicle could map the scene, generate the planned path for the user, and relay guidance back to the user's phone.

Chapter 4

Rendezvous With Static Sensing

In this chapter, we look at the inverse of the prior chapter. Here the user is attempting to rendezvous with static objects in a sensed environment. We expand upon the system we used in the previous chapter to guide the user. In other words, instead of moving a robot, the planned trajectory is converted into instructions that are presented to the user. In this chapter, we describe this extension and subsequent evaluation of the system.

4.1 Implementation

We assume that the system (1) knows the positions of the desired target as well as competing objects in the scene according to a pre-computed map of the environment where all the point of interests and static obstacles are labeled and (2) continuously tracks the user and the dynamic obstacles from sensors in the environment (e.g., LIDAR on waiting autonomous cars, a static robot, building cameras, etc).

Although humans can move like a holonomic robot, treating them like a non-holonomic robot with constraints on turning radii can make their path more legible to other pedestrians. This is important as it informs other pedestrians where the user may be going and allows them to change their current paths so as to not intersect the user. Therefore, we expand our lattice graph-based ARA* to allow the system to generate a global plan within a specified time constraint (for this work, set to 5 seconds) that discourages the user from turning in place near the goal.

Since the current global planner requires up to 5 seconds to plan, there may be time sync problems when the user has moved away from the starting position used for planning. For mobile robot navigation (e.g., Chapter 3), this is often addressed through a local planner that uses odometry and laser scanner data to localize itself, estimate how far the robot is from the start position, and send information to the robot controller at the correct time. This moves the robot in the desired direction. However, estimating one's own position is challenging even for users with visual impairment that may have been trained to do so in step counting. Human gait is more variable step-by-step than odometry and smaller errors can propagate, making the path guiding too inaccurate to be useful. Instead, we can incorporate information about the user's real-time position and orientation from the sensors in the scene into our local planner instead.

We designed a system architecture, depicted in Figure 4.1, that can extract the information on where the user is from the phone, extract the path-guiding information from the ROS planner, and allow the phone and the planner to communicate. After several design iterations, we decided to use the HTTP GET request protocol to allow the phone to initiate the request and receive the human-readable path-guiding instructions. We chose this because this allows the system to separate the planning portion from the feedback generation, which is dependent on the hardware of different personal and wearable devices. This design approach can be expanded to allow the user to customize the frequency of each type of instructions given. Currently, the default behavior is to give verbal instructions about every 30 cm ($\approx$ 1 ft) and vibration about every 90 cm ($\approx$ 3 ft). The duration of vibration changes inversely with the number of time steps the user is away from their goal.

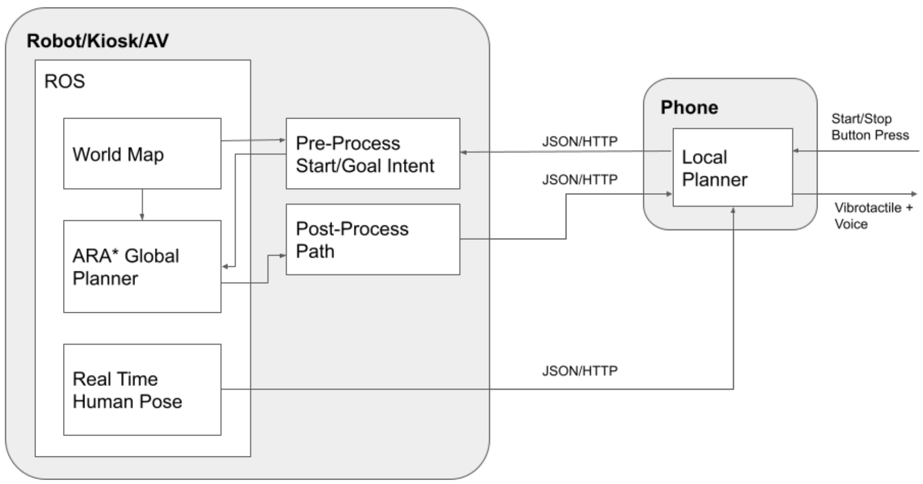

Figure 4.1: System Diagram: The user initiates interaction with the system by pressing start on the phone screen: then, the local planner system on the phone sends two HTTP GET requests to the software system of the static target to get (1) the global plan and (2) the real time human position and orientation from the sensors. The local planner then sends vibrotactile signals and voice instructions according to the algorithm described in Algorithm 1.

One purpose of our local planner is to decide when the appropriate time to tell user their next direction is. Timely guidance is needed for the user to correct their path when they make a mistake or are approaching a planned turn. In the best case scenario, the user might be just a little bit off from the desired plan (i.e. deviance $< \alpha = 10$ cm ≈ 4 in), and the local planner would just guide the user to the next step of the global plan. However, if the user deviates from the desired plan too much to be directed to the next step with the same instruction (i.e. $\alpha <$ deviance $< \beta = 20$ cm ≈ 8 in), then the local planner will need to introduce a correction. It should guide the user from their current location to where they are supposed to be before continuing the rest of the navigation instructions. However, if the user made a bigger mistake, the local planner will need to send a request to the global planner for re-planning from the user's current location and replace the original navigation plan.

The pseudocode of the local planner is outlined in Algorithm 1.

Algorithm 1 Local Planner

Input: Current Human Pose (h), Global Plan (P), Goal Pose (g)
Output: Local Plan Instruction

```
for p ∈ P do
    if |p − h| < α and orientation in clock angle of h is the same as that of p then
        output instruction at step p
    else if |p − h| < β then                                  ▷ α, β ∈ ℝ and α < β
        output instruction from h to p
        output instruction at step p
    else
        P = Global_planner(h, g)
        Local_planner(h, P, g)
        break
    end if
end for
```

It is possible for the local planner to become stuck and run forever if the user decides to completely deviate from the plan, so we arbitrarily set a timeout when the local planner is called more than 5 times in a trip. We also adjusted the system by moving computation of the local plan from the static system to inside the phone to avoid expensive re-planning and information communication with the global planner.

While mobile robots can be guided by providing sequences of specific absolute positions and orientations with respect to the world map, humans need different kinds of information to seamlessly navigate to the goal. Instead of using poses, we communicated the sequence of turns and distances needed to travel from the current position to the next. We chose to communicate sequences of turns in units of twelve clock angles since (1) it provides more precision compared to the four and eight cardinal directions and (2) is a common convention when providing orientation guidance to people. The distances from each time steps returned by the global planner were represented to users in feet since we plan to work with users in United States.

4.2 Evaluation

To evaluate our system, we simulated different human behaviors and observed the reaction of our system. We used the same simulated world setup as the previous chapter. We synthesized real-time human pose based on the generated global plan and added a desired amount of noise to some portions. We evaluated our system on two sets of static target positions: (1) in the corner of the room and (2) in the middle of the room. We also evaluated our system on two levels of human pose noise: (1) noise $< \alpha$ and (2) $\alpha <$ noise $< \beta$ for all time steps, where noise is defined as the norm of the difference between human position and the position returned by the global plan at a particular time step.

From the simulation, we observed that when the user deviated from the plan less than α for all steps, the local planner gave verbal instructions at approximately every 1 foot and haptic

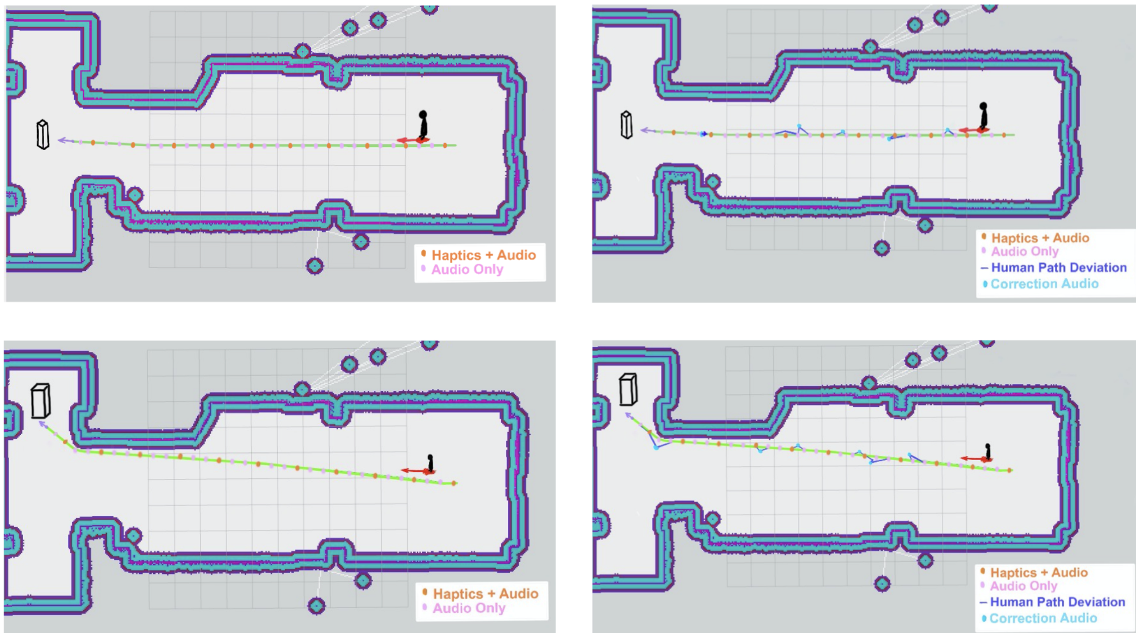

Figure 4.2: The rectangular prisms in the four images depict where the static targets are. The black figure with red arrow at the bottom depicts the user's position and orientation. The images in the left column depict the scenario where the user deviates from the plan less than α for all steps. The images in the right column depict the scenario where the user deviates from the plan more than α for some steps but less than β for all steps.

feedback approximately every 3 feet. When the user deviated from the plan more than α but less than β, the system sent audio to correct the user's path from where they were at to where they were supposed to be at the current time step. We noticed that this could be inefficient sometimes, especially when the user's current location was already closer to the position than they were supposed to be at the next time step. In that case, the user would be guided to walk back before they could walk towards the goal again.

This shows our system is robust to small deviations from the planned path by the user (< 10 cm $\approx$ 4 in). However, it may create problems when users make more mistakes as the path gets more difficult (i.e., contains multiple turns in a row) since the frequency of the instructions stay constant regardless of the path complexity. Moreover, this current setting might annoy experienced users who might not want to be notify every so often about keep going straight. In order to validate these hypotheses, our system should be further tested in-person by end users.

4.3 Summary

This chapter described a proof of concept robotic system that helps people with visual impairments rendezvous with static objects (e.g., a waiting robot, a kiosk, an autonomous car at the curbside) using external sensing. We extended our system from a prior project ([11], Chapter 3) which used a lattice graph-based ARA* global planner to encourage smooth trajectories. We evaluated our system with four simulated scenarios, and observed that the system guided a simulated user with verbal instructions and haptic feedback exactly at the specified frequency when the user's deviation from the path was sufficiently small. Future work will further test our system's behavior when (1) simulated users make mistakes, not only for position but also for orientation, and (2) when simulated users deviate from the path so much that the local planner needs to ask the global planner for a new plan.

Chapter 5

Conclusion

This thesis explores how adjustable motion planning can help (1) a mobile robot approach a person with a visual impairment and (2) guide a person with a visual impairment to a stationary robot, kiosk, or autonomous vehicle in a socially seamless manner. We used a lattice graph-based ARA* planner as a global planner for both systems. This allows the planner to return a path in a specified duration with provable bounds on sub-optimality [10] while discouraging the moving party from turning in place when approaching the goal, reducing risk of collisions.

In the first project, we examined how a mobile robot should rendezvous with people with visual impairments. In this scenario, the device with sensors moves to a stationary, waiting user. We developed a proof of concept of a robotic system that can handle short-range rendezvous with people with visual impairments [11]. To evaluate our system, we conducted an interview with an O&M specialist. They observed that (1) our planner produced less obtrusive trajectories for the user than the ROS default global planner and (2) recommended that our system allow the robot to approach the person from the side, as opposed to the front examples we showed.

In the second project, we focused on the inverse, where the user is moving towards a static device that has sensors. We expanded the system from the first project to design a path guiding plan for the second system. Since it is difficult for humans to seamlessly navigate to a goal using robotic pose information, we converted the information into a sequence of turns (in clock angles) and distances (in feet) needed to travel from the current position to the next. We modified our prior system by moving the local plan inside the user's phone to avoid expensive re-planning and information communication with the ROS Global Planner in the goal device. We evaluated our system by simulating different human behavior and observed the reaction of our system. We observed that our system is robust to small deviations from the planned path by the user (< 10 cm ≈ 4 in). However, our system can be inefficient when users deviate from the path more. This is due to the need to direct users to recover the intended path before moving forward.

Future work will further test the performance of both systems with end users. For the first system, we hope to determine the most safe and socially acceptable goal positions for the robot when approaching the user in different scenarios and to fine-tune a cost function for local planning. We also plan to explore how the mobile robot should introduce itself to the user to create trust before starting navigation. For the second system, we hope to find suitable parameters that allow the planner to perform adjustable guiding frequency based on path curvature. The user may be more likely to deviate from the plan when the target position is harder to reach or the planned path is more curvy. With usability testing, we plan to identify the most effective communication channel whether that be audio or vibrotactile or some combination of the two.

This work will support the larger project's goal of shared robot guides for people with disabilities. Research from this work will inform the rendezvous tasks when users enter a building or seek help.

Figure and Table Descriptions

Figure 1.1: The image depicts a person with visual impairment walking towards autonomous car waiting at a driveway.

Figure 2.1: The image depicts a row of six WMTA ticket machines. There are three of short and three of tall kiosks. The height of kiosk from left to right is as follows: short, tall, tall, tall, short, and short.

Figure 3.1: The illustration contains 8 images in a 2 by 4 grid. The four images in the top row depict a person standing facing east, and a robot facing west about 3 meters away from the person. The four images in the bottom row depict a person standing in the north corner and facing south, and a robot facing northwest about 4 meters diagonally away from the person. The images in the left most column depict the scenes from the simulator where the participant imagines themselves standing to the south of the interaction. The images in the second column from the left depict straight line trajectories from the robot to the user produced by the baseline planner, meaning that the global planner does not take into account the orientation at the goal state which would cause the robot to turn in place at the end as the local planner takes over. The images in the third column from the left depict trajectories produced by a narrow planner where the robot starts reorienting itself about 1.5 meters away from the person when the person is in the middle of the room and about 1 meter away from the corner when the person is in the corner. The images in the last column depict trajectories produced by a wide planner where the robot starts reorienting itself about 2 meters away from the person when the person is in the middle of the room and about 1 meter away from the corner.

Table 3.1: The table shows the O&M specialist's agreement ratings to 6 statements: 1) This robot behaves in a socially acceptable manner; 2) I would interrupt this robot before it reaches a person with a visual impairment; 3) This robot makes me uncomfortable; 4) I would recommend people with visual impairments call for this robot whenever it is available in a place they are unfamiliar with; 5) I would recommend the person with a visual impairment call for robot assistance only when human assistance is not available; and 6) I would discourage the person with a visual impairment from using this robot. The entries column from left to right corresponds to the videos of trajectories produced by the three planners in two different scenarios: baseline with the user in the corner, narrow with the user in the corner, wide with the user in the corner, baseline with the user in the middle of the room, narrow with the user in the middle of the room, and wide with the user in the middle of the room. For all videos, the ratings of the specialist for each of the 6 statements are the same. They are Agree, Disagree, Strongly Disagree, Agree, Disagree, and Disagree, respectively.

Figure 4.1: The illustration shows the system diagram used for rendezvous with static sensing task. The user can start or pause the interaction with the system by pressing start/pause button on their mobile phone screen. When the user press start, the local planner will send two HTTP GET requests to the software stack inside static sensing (e.g. stationary robot, kiosk, or autonomous car). One request is for obtaining real time human pose information in JSON object format, while another is for obtaining the global plan. There is a pre-processing and post-processing conversion that is done inside the software stack of the static sensing before and after the communication with ARA* Global Planner inside ROS navigation stack. To compute global plan, the global planner send the request to look-up the the (1) positions of static obstacles and (2) the position and orientation of the user's goal from the pre-computed world map. With information of global plan and real time human pose, the local planner decide when vibrotactile feedback and voice instruction should be output.

Figure 4.2: The illustration contains 4 images in a 2 by 2 grid. The two images in the top row depict a person standing facing west, and a kiosk facing east about 13 meters away from the person. The two images in the bottom row depict a kiosk in the north corner and facing south, and a person facing northwest about 13.5 meters diagonally away from the kiosk. The two images in the left column depict the scenes, where simulated human poses deviate from the plan minimally. The two images in the right column depict the scenes, where simulated human poses are more noisy.

Bibliography

[1] Dragan Ahmetovic, Cole Gleason, Kris M. Kitani, Hironobu Takagi, and Chieko Asakawa. Navcog: Turn-by-turn smartphone navigation assistant for people with visual impairments or blindness. In *Proceedings of the 13th International Web for All Conference*, W4A '16, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450341387. doi: 10.1145/2899475.2899509. URL `https://doi.org/10.1145/2899475.2899509`. 1, 2

[2] Mauro Avila, Katrin Wolf, Anke Brock, and Niels Henze. Remote assistance for blind users in daily life: A survey about be my eyes. In *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, PETRA '16, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450343374. doi: 10.1145/2910674.2935839. URL `https://doi.org/10.1145/2910674.2935839`. 2

[3] Shiri Azenkot, Catherine Feng, and Maya Cakmak. Enabling building service robots to guide blind people: A participatory design approach. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, HRI '16, page 3–10. IEEE Press, 2016. ISBN 9781467383707. 1, 2

[4] Jeffrey P. Bigham, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C. Miller, Robin Miller, Aubrey Tatarowicz, Brandyn White, Samual White, and Tom Yeh. Vizwiz: Nearly real-time answers to visual questions. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology*, UIST '10, page 333–342, New York, NY, USA, 2010. Association for Computing Machinery. ISBN 9781450302715. doi: 10.1145/1866029.1866080. URL `https://doi.org/10.1145/1866029.1866080`. 2

[5] Jake Bruce, J Wawer, and Richard Vaughan. Human-robot rendezvous by co-operative trajectory signals. In *Proc. 10th ACM/IEEE International Conference on Human-Robot Interaction Workshop on Human-Robot Teaming*, 2015. 2

[6] Manuel Castillo-Cara, Edgar Huaranga-Junco, Giovanny Mondragón-Ruiz, Andree Salazar, Luis Orozco Barbosa, and Enrique Arias Antúnez. Ray: Smart indoor/outdoor routes for the blind using bluetooth 4.0 ble. *Procedia Computer Science*, 83:690–694, 2016. ISSN 1877-0509. doi: https://doi.org/10.1016/j.procs.2016.04.153. URL `https://www.sciencedirect.com/science/article/pii/S1877050916301843`. The 7th International Conference on Ambient Systems, Networks and Technologies (ANT 2016) / The 6th International Conference on Sustainable

Energy Information Technology (SEIT-2016) / Affiliated Workshops. 1

[7] João Guerreiro, Daisuke Sato, Saki Asakawa, Huixu Dong, Kris M. Kitani, and Chieko Asakawa. Cabot: Designing and evaluating an autonomous navigation robot for blind people. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '19, page 68–82, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450366762. doi: 10.1145/3308561.3353771. URL `https://doi.org/10.1145/3308561.3353771`. 1, 2

[8] Taeyong Kim, Sanghong Kim, Joonhee Choi, Youngsun Lee, and Bowon Lee. Say and find it: A multimodal wearable interface for people with visual impairment. In *The Adjunct Publication of the 32nd Annual ACM Symposium on User Interface Software and Technology*, UIST '19, page 27–29, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450368179. doi: 10.1145/3332167.3357104. URL `https://doi.org/10.1145/3332167.3357104`. 2

[9] Maxim Likhachev. Search-based planning with motion primitives, 2010. 3.2

[10] Maxim Likhachev, Geoff Gordon, and Sebastian Thrun. Ara*: Anytime a* with provable bounds on sub-optimality. In *Advances in Neural Information Processing Systems*, pages 767–774, 2003. 3.2, 5

[11] Jirachaya "Fern" Limprayoon, Prithu Pareek, Xiang Zhi Tan, and Aaron Steinfeld. Robot trajectories when approaching a user with a visual impairment. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '21, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383066. doi: 10.1145/3441852.3476538. URL `https://doi.org/10.1145/3441852.3476538`. 3, 4.3, 5

[12] Eitan Marder-Eppstein, Eric Berger, Tully Foote, Brian Gerkey, and Kurt Konolige. The office marathon: Robust navigation in an indoor office environment. In *2010 IEEE International Conference on Robotics and Automation*, pages 300–307, 2010. doi: 10.1109/ROBOT.2010.5509725. 3.2

[13] Rajesh Kannan Megalingam, Souraj Vishnu, Vishnu Sasikumar, and Sajikumar Sreekumar. Autonomous path guiding robot for visually impaired people. In Pradeep Kumar Mallick, Valentina Emilia Balas, Akash Kumar Bhoi, and Ahmed F. Zobaa, editors, *Cognitive Informatics and Soft Computing*, pages 257–266, Singapore, 2019. Springer Singapore. ISBN 978-981-13-0617-4. doi: 10.1007/978-981-13-0617-4_25. 1, 2

[14] Amal Nanavati, Xiang Zhi Tan, and Aaron Steinfeld. Coupled indoor navigation for people who are blind. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '18, page 201–202, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450356152. doi: 10.1145/3173386.3176976. URL `https://doi.org/10.1145/3173386.3176976`. 3.3.1

[15] Amal Nanavati, Xiang Zhi Tan, Joe Connolly, and Aaron Steinfeld. Follow the robot: Modeling coupled human-robot dyads during navigation. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019. doi: 10.1109/IROS40897.2019.8967656. 1, 2

[16] Mark B. Neider and Gregory J. Zelinsky. Scene context guides eye movements during visual search. *Vision Research*, 46(5):614–621, 2006. ISSN 0042-6989. doi: https://doi.org/10.1016/j.visres.2005.08.025. URL `https://www.sciencedirect.com/science/article/pii/S0042698905004505`. 1

[17] Rainer Palm, Ravi Chadalavada, and Achim J Lilienthal. Recognition of human-robot motion intentions by trajectory observation. In *2016 9th International Conference on Human System Interactions (HSI)*, pages 229–235. IEEE, 2016. doi: 10.1109/HSI.2016.7529636. 2

[18] Manaswi Saha, Alexander J. Fiannaca, Melanie Kneisel, Edward Cutrell, and Meredith Ringel Morris. Closing the gap: Designing for the last-few-meters wayfinding problem for people with visual impairments. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '19, page 222–235, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450366762. doi: 10.1145/3308561.3353776. URL `https://doi.org/10.1145/3308561.3353776`. 2

[19] Sina Azizi Soldouz, Md Sami Hasnine, Mahadeo Sukhai, and Khandker Nurul Habib. Looking through the perceptions of blinds: Potential impacts of connected autonomous vehicles on pedestrians with visual impairment. *Transportation Research Record*, 2674 (5):183–195, 2020. doi: 10.1177/0361198120914299. URL `https://doi.org/10.1177/0361198120914299`. 1

[20] Xi Sun, Xinshuo Weng, and Kris Kitani. When we first met: Visual-inertial person localization for co-robot rendezvous. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 10408–10415, 2020. doi: 10.1109/IROS45743.2020.9341739. 3.1