MiRODES: Mini Intelligent Robot for On-campus Domain-specific Event Support

Mila Manzano
Department of Computer Science
The College of New Jersey
manzanp1@tcnj.edu

Kristen O'Donnell
Department of Computer Science
The College of New Jersey
odonnek7@tcnj.edu

Eden Espinosa
Department of Computer Science
The College of New Jersey
espinoe2@tcnj.edu

Sejong Yoon
Department of Computer Science
The College of New Jersey
yoons@tcnj.edu

Abstract

Life with robots in everyday life is no longer a picture of science fictions. Robots are expected to interact with people in more natural, real-world contexts, such as crowded indoor spaces beyond controlled environments. This paper introduces our ongoing project of building a robot tour guide that aims to cope with crowds of visitors during on-campus recruiting events. We present a user interface that can effectively interact with guests while answering domain specific questions. Additionally, we discuss challenges in our robot system design with the aim of safely navigating through crowded indoor environments.

1 Introduction

As the world continues to develop pioneering technology, the progress of robotics has emerged as a defining chapter. Beyond revolutionizing industries, robots are augmenting human capabilities and paving the way for a future where machines seamlessly integrate into the fabric of everyday human life. For example, it is now fairly common to see guide robot, including commercial solutions, in various places, such as museums, airports, or stations [1, 2]. Due to wide availability of these service robots, methods to evaluate such robots were also developed [3]. However, commercial service robots are expensive, cost at least a few thousand US dollars. Moreover, there are still challenging open problems such as making appropriate locomotive plans by understanding dense crowd motion, or effectively cooperating with human tour guide.

In this project, we aim to build a Mini Intelligent Robot for On-Campus Domain-specific Event Support (MiRODES), which is a tour guide type of service robot. We are particularly interested in building a robotic that can assist student tour guide at on-campus recruiting events at higher education institutes. MiRODES aims to achieve the following design goals: (a) designed specifically for a particular domain of on-campus events, (b) capable of cooperating with and lessening the load of human tour guide, (c) able to smoothly navigate through dense crowd, and ideally (d) cost-efficient. To accomplish the goals, we focus on software aspect of the system design and rely on existing platforms for hardware aspects of the system. A typical use case scenario we hope to accomplish is depicted in Fig. 1 and more details of the scenario will be explained in Section 2.

This is the accepted version of the manuscript. Copyright held by authors. Do not redistribute. To cite this article: "Mila Manzano, Kristen O'Donnell, Eden Espinosa, and Sejong Yoon. 2024. MiRODES: Mini Intelligent Robot for On-campus Domain-specific Event Support. In Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (HRI '24). Association for Computing Machinery, New York, NY, USA, 1245–1248. https://doi.org/10.1145/3610978.3641272"

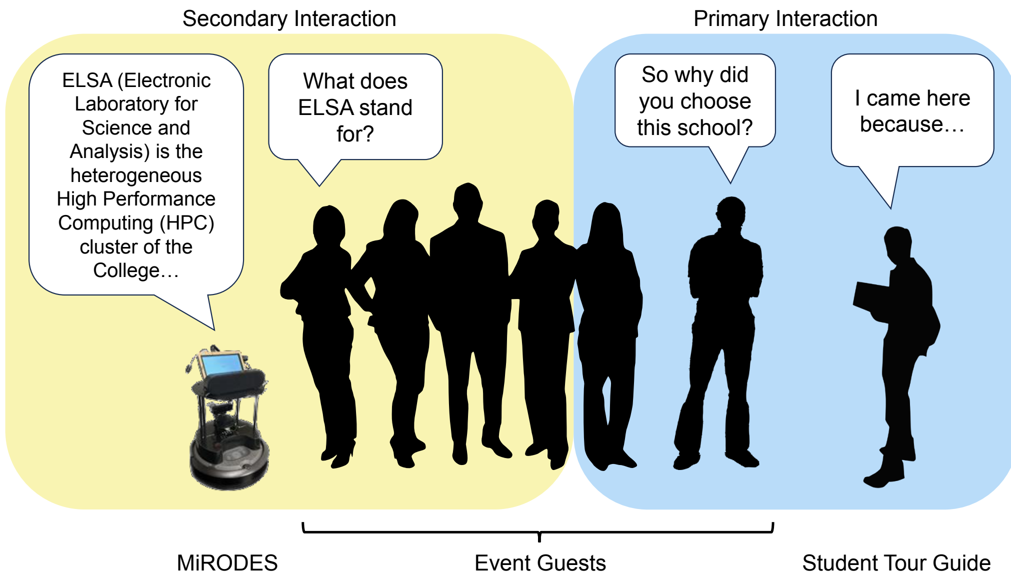

Figure 1: MiRODES Application Scenario 1. Our system can lead the introduction, interact with guests with light, though domain-specific questions, freeing the human tour guide to engage in in-depth conversations with a subset of guests.

In the rest of the paper, we first introduce the team, work distribution, and the use case scenarios of our service robot. Then, we introduce our implementation details followed by identified challenges in completing the project and future plans.

1.1 Team

The project team has three female undergraduate student members majoring in computer science, one junior (Manzano), one sophomore (O'Donnell), and one senior (Espinosa). The team is guided by a faculty mentor (Yoon). None of the team members had prior experience in robotics or human computer interaction research. We worked together on the project for seven weeks during the summer of 2023. Manzano led the intelligent UI system development, continued her project after the summer, and drafted this report. O'Donnell led the TurtleBot4 navigation module development. Espinosa participated mostly in a supporting role in this project while working on her own computer vision project that will be incorporated into this system later.

2 DESIGN AND APPLICATION SCENARIOS

On-campus recruiting events at higher education institutes are often packed with guests, including prospective students and families. Typically, they are partitioned into tour groups with a few human tour guides, usually consist of current students. The purpose of guests is to obtain as much information as possible about the school, while that of the student tour guide is to provide good information while making sure everyone is involved and engaged throughout the tour session. However, often, that is not the case, increasing the load and stress of the student tour guide. We identify three scenarios that our proposed service robot, MiRODES can aid support:

- When a student tour guide is engaged deeply with a subset of guests (denoted as primary interaction), MiRODES can provide a domain-specific question answering system (denoted as secondary interaction), so that no guest is left out in the tour experience (Fig. 1).
- MiRODES can check the schedule timeline and keep the timing of the tour by reminding the student tour guide to move on, when the engagement with guests got delayed (Fig. 2).
- With an ability to smart navigation and steering for indoor environment, MiRODES can lead the way in relatively less crowded area (Fig. 3).

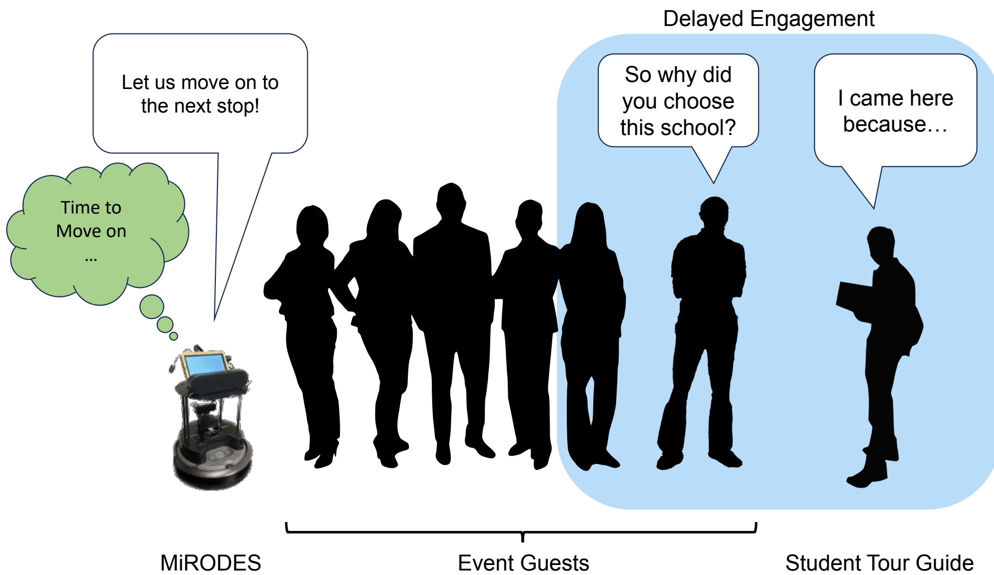

Figure 2: MiRODES Application Scenario 2. Our system can check the schedule and notify the human tour guide. Note that this requires not only time keeping, but also indoor localization.

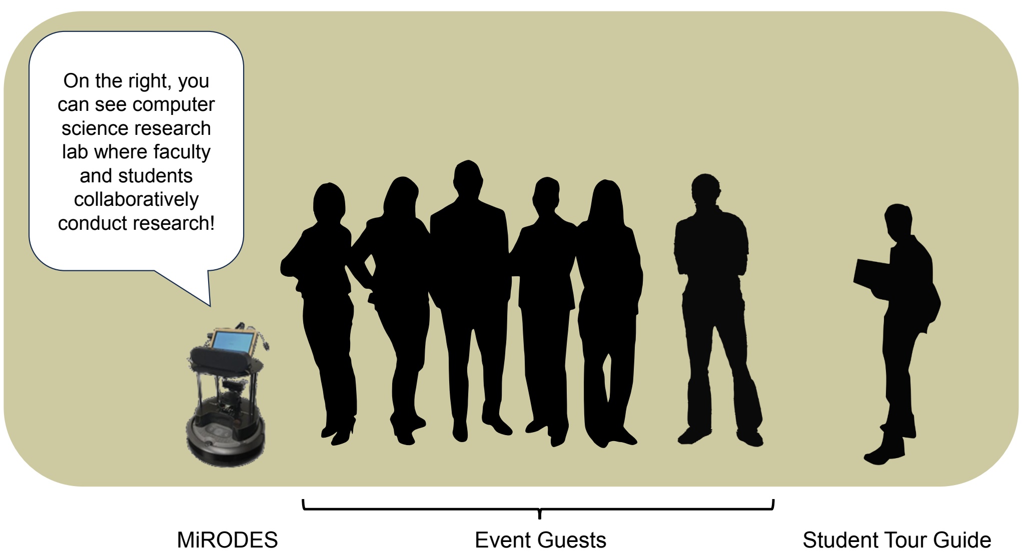

Figure 3: MiRODES Application Scenario 3. In less crowded areas, our system can lead the tour group, alleviating the load of human tour guide.

With the above, MiRODES can lessen the load of student tour guide while making sure that no important information is missing in the communication.

3 IMPLEMENTATION

Fig. 4 shows the overall system design. Our system at a high level, consists of three parts, user interface, Turtlebot4, and the core control server. Our robot's hardware configuration is shown in Fig. 5. Current system relies on pre-existing robot platform and hardware parts. The robot base is TurtleBot4 and consists of Raspberry Pi 4 with RGB-D and LiDAR sensors, with Create 3 to move

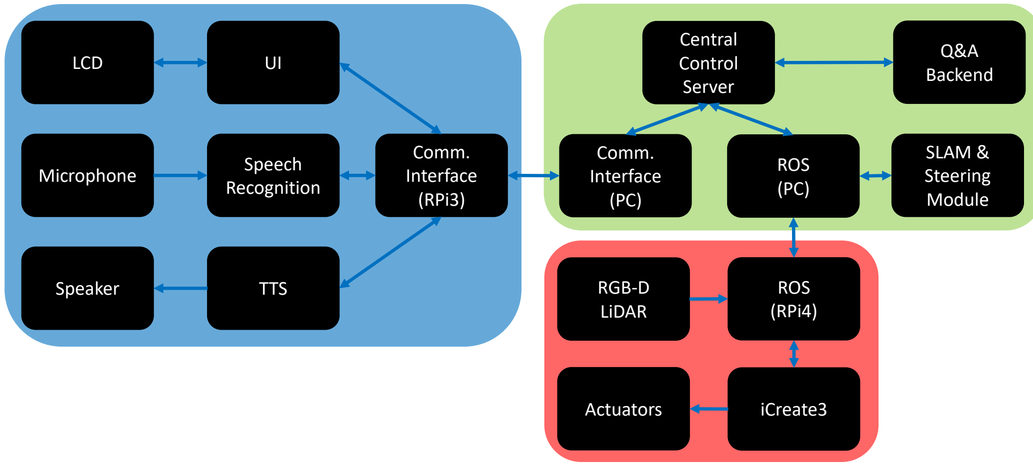

Figure 4: Overview of MiRODES system design. It has three major components, user interface in blue panel, actuator, TurtleBot, in red panel, and the central server in green panel. This paper focuses on blue panel (user interface) and a subset of green panel (main control).

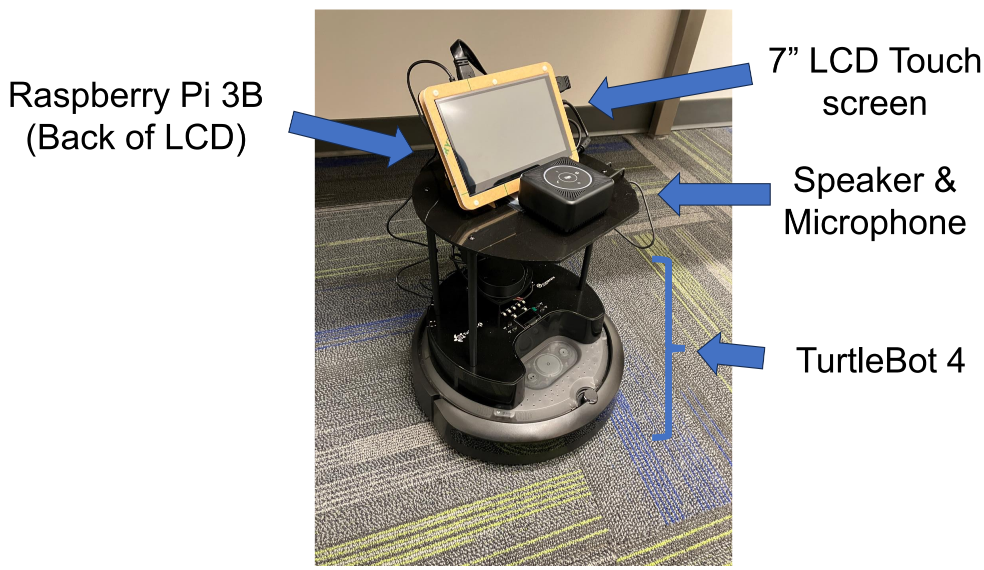

Figure 5: Hardware configuration of MiRODES, showing LCD and speakerphone installed on top of TurtleBot4.

around. The system also provides power to a 7-inch touch screen LCD and a Rasberry Pi 3B for a user interface. In this report, we focus on introducing the LCD user interface we developed, and introduce the progress and future work to complete the project implementation.

3.1 Interactive LCD User Interface

Attaching an LCD screen to the TurtleBot adds an interactive element that allows users to ask the system questions and return AI-generated responses. This application uses a basic client-server model in which an external computer acts as a host server and the LCD attached to a Raspberry Pi acts as the client. Dependencies used include open source programs such as, GPT4All, GTK+3.0 (for GUI), and WhisperCPP (for speach recognition).

GPT4All [4] is an open-source program used to develop AI generated text based on Large Language Model (LLM). GPT4All can run locally on consumer grade CPUs which is great for enforcing privacy and security. Since this program is Python based, the host program needed to be written in Python. However, this did not prove to be an issue since the host program runs on an external computer. GPT4All was modified in this interactive program to use PDF documents as its large language database. This allows answers to be domain specific and accurate.

GTK3 and gtkmm were used in the development of the client user interface. The interface was rather simple using a basic chat interface that displays messages sent between the user and the robot system (GPT4All backend running on the server). Given that the LCD was backed by a Raspberry Pi 3B, it was best to minimize the amount of CPU being used to ensure that the program was running at an optimal speed. To do so, C/C++ was used as the preferred language for the LCD to ensure low CPU usage. Initially, we used GTK for the GUI since it was compatible with C. However, since we needed the WhisperCPP for speech recognition, we migrated our code to use gtkmm, the C++ interface for GTK. Fig. 6 shows the design of our prototype user interface.

Figure 6: LCD User Interface application window showing chat log and an input box with a submit button at the bottom.

WhisperCPP [5] is an open-source implementation inspired by OpenAI's Whisper, designed to be used for speech recognition tasks. We employed WhisperCPP over the original Whisper since WhisperCPP is in C++ while Whisper is in Python, and we need to minimize CPU usage on the Raspberry Pi, the client side. Additionally, WhisperCPP can be run locally without being connected to a network. We modularized our main user interface program so that the WhisperCPP is invoked as an external module. This implementation choice helped the core logic of our system more easily maintained. In addition, the core logic was implemented as a multi-threaded application so that GTK main loop and WhisperCPP can run concurrently. Dan Fountain's speech script [6] was used to develop the robot's "voice" as we found that this library provided the best quality audio after testing various options. As a result of this sub-project, a fully functioning user interface device is developed, that can recognize speech, send the text to the server running GPT4All engine, develop responses, send back to the client, and respond clearly from the client.

3.2 Navigation with TurtleBot4

TurtleBot has been a popular platform for educational robot projects. Over the course of our research we devoted most of our effort to understand and implement of the TurtleBot navigation. MiRODES' navigation module has several subtasks, each of those is an ongoing research area in robotics: (a) localizing itself within a given, but unseen, raster image map, (b) finding the optimal path from current location to a goal location, (c) planning the navigation, (d) while avoiding collisions to static and dynamic obstacles in the environment. During the span of the project period (the summer of 2023), we focused on the first task, the localization problem.

After researching over options, we chose to use the Monte Carlo Localization (MCL). MCL works by using particle filters and LiDAR scan data to estimate the robot's position in the given map. Each particle represents the probability of the robot being within that location on the map. As the robot moves in the environment, the particles begin to converge around a single location which will ultimately be its correct position to initialize the navigation module. At the end of the project period, we were able to have a simulated robot localize itself in a default map provided in the TurtleBot package. This experience expanded our knowledge about the Robot Operating System (ROS) and how robots can learn to adapt to their surroundings and navigate.

4 DISCUSSION AND FUTURE WORK

Through our research, we have garnered valuable insights into the world of robotics. Our efforts have culminated in the development of key components for a comprehensive system, laying a foundation for future researchers to build upon. Moving forward, we hope to focus on the following:

- Improve the robustness of the user interface and response quality generated by the domain-specific LMM.
- Improve the automatic localization system. It should work with the real maps of the campus buildings. We may use indoor localization methods to improve the accuracy.
- Implement the navigation module with path finding, planning, and collision avoidance.

ACKNOWLEDGEMENT

This project is supported by National Science Foundation Grant #1955365 and #1955404.

References

[1] Lindsey – A robot tour guide in Lincoln's "The Collection Museum". https://lcas.lincoln.ac.uk/wp/projects/lindsey-a-robot-tour-guide/, 2018. [Online; accessed 8-Dec-2023].

[2] Promobot. https://promo-bot.ai/use-case/tour-guide/, 2013. [Online; accessed 8-Dec-2023].

[3] A. Jihan KIM, B. Bo-Yeong Kang, C. JongWoon Woo, and D Hyung-taek Baek. A study on evaluation method for indoor guide robot. In *2020 20th International Conference on Control, Automation and Systems (ICCAS)*, pages 107–110, 2020.

[4] Yuvanesh Anand, Zach Nussbaum, Brandon Duderstadt, Benjamin Schmidt, and Andriy Mulyar. Gpt4all: Training an assistant-style chatbot with large scale data distillation from gpt-3.5-turbo. https://github.com/nomic-ai/gpt4all, 2023.

[5] Georgi Gerganov. Whisper.CPP. https://github.com/ggerganov/whisper.cpp, 2023. [Online; accessed 8-Dec-2023].

[6] Dan Fountain. Raspberry Pi Text to Speech. https://web.archive.org/web/20151016182428/http://danfountain.com/2013/03/raspberry-pi-text-to-speech/, 2013. [Online; accessed 8-Dec-2023].