Quadrupedal Robotic Guide Dog with Vocal Human-Robot Interaction

Kavan Mehrizi[1], Zhongyu Li[2], Koushil Sreenath[2]
[1]Department of Computer Science, Diablo Valley College, Pleasant Hill, CA
[2]Department of Mechanical Engineering, University of California, Berkeley, Berkeley, CA

2021 Transfer-to-Excellence Research Experiences for Undergraduates Program (TTE REU Program)

Abstract

Guide dogs play a critical role in the lives of many, however training them is a time- and labor-intensive process. We are developing a method to allow an autonomous robot to physically guide humans using direct human-robot communication. The proposed algorithm will be deployed on a Unitree A1 quadrupedal robot and will autonomously navigate the person to their destination while communicating with the person using a speech interface compatible with the robot. The speech interface utilizes cloud based services such as Amazon Polly and Google Cloud to serve as the text-to-speech and speech-to-text engines.

Background

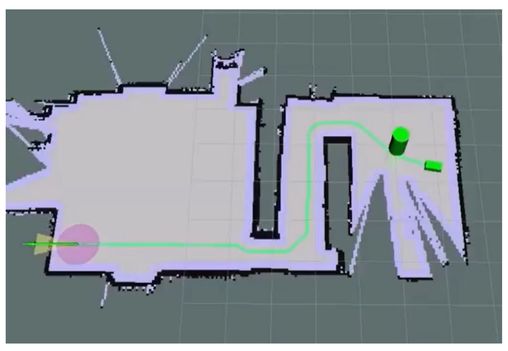
The training and maintenance of traditional guide dogs presents challenges to the elderly, frail, and visually-impared. Each guide dog has to be trained individually in a time and labor intensive process and the skills gained from one dog cannot be implemented into another dog. In addition, guide dogs may get ill or need to retire, which creates an issue of training and finding of a replacement dog, which may not be a good match for the user [1].

Most previous robotic guides are bunglesome and are limited in maneuvering in narrow and complex spaces or in uneven terrain due to their bulky size or rely on physical interaction between the robot and the user, by having them physically hold a leash or rigid arm, without any way for the user to verbally give commands such as to reroute, or stop the robot. In addition, they also do not have the ability to create a customizable wake word [2]–[4].

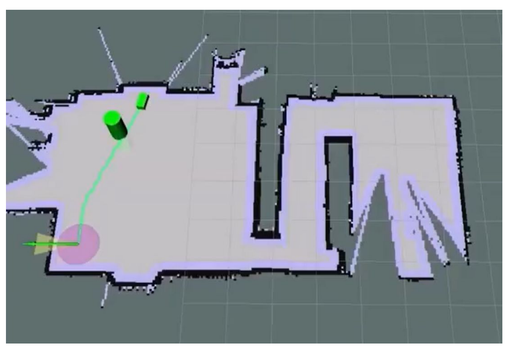
Methodology

- The robot vocally communicates with and understands the user using text-to-speech (TTS) and speech-to-text engines (STT)
- Installed a speaker onto robot to send and receive audio data
- Found basic open source codes [5], [6] that allowed for the integration of Amazon Polly to be used to TTS engine and Google Cloud's STT API, then configured each into the robot's infrastructure
- Created algorithm to allow for vocal commands and to act accordingly
- Would publish to navigation node's target goal if user wants to navigate somewhere based on preset coordinates

Results



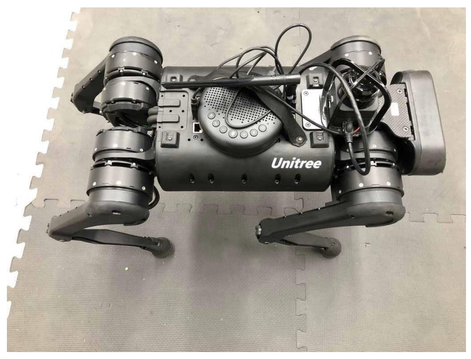
Simulated navigation map after receiving coordinates from speech interface. Green figures show final position of user and robot, while green line is path created to final position.



Navigation map after receiving new navigation goal coordinates from speech interface. Purple circle is initial position of user and robot.

- STT ignored speech with no wake word
- STT understood user's command and translated into text
- Algorithm published coordinates to navigation goal and published response to TTS
- TTS received text from node and translated response into audible voice
- Guide dog's navigation obtained coordinates from goal node and created path to goal

Conclusion



A1 Robot with speakerphone installed

- Developed a successful speech interface
- Interface algorithm communicates with TTS and STT engines
- We were able to integrate the different engines and scripts to communicate with the guide dog navigation pipeline
- The main advantages of this work is we are able to customize the wake up word to anything and are able to create custom commands fairly easily by adding them to the word dictionary
- Unlike the previous guide robots, we are able to have a custom wake up word and are able to integrate this speech interface with a leash while using a maneuverable robot

Future Work

- Currently developing the guide dog to operate an elevator to allow for multi-floor navigation
- In order to facilitate multi-floor navigation, restructuring the robot's navigation to take floors into consideration
- Need to further develop the speech interface to send coordinates that can relate to what floor level the navigation goal is at
- Interface will be developed to allow for more commands such as telling the robot to stop at its current position as well as giving the user instructions when needed such as to tell user to avoid and move around an obstacle
- Need to further optimize the speech interface such as making it easier to input new commands and new locations into the algorithm

References

[1] J. Lloyd, C. Budge, S. La Grow, and K. Stafford, "An Investigation of the Complexities of Successful and Unsuccessful Guide Dog Matching and Partnerships," *Frontiers in Veterinary Science*, vol. 3, 2016.

[2] Z. Li and R. Hollis, "Toward a Ballbot for Physically Leading People: A Human-Centered Approach," *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019.

[3] A. Wachaja, P. Agarwal, M. Zink, M. R. Adame, K. Möller, and W. Burgard, "Navigating blind people with walking impairments using a smart walker," *Autonomous Robots*, vol. 41, no. 3, pp. 555–573, 2016.

[4] D. R. Bruno, M. H. de Assis, and F. S. Osorio, "Development of a Mobile Robot: Robotic Guide Dog for Aid of Visual Disabilities in Urban Environments," *2019 Latin American Robotics Symposium (LARS), 2019 Brazilian Symposium on Robotics (SBR) and 2019 Workshop on Robotics in Education (WRE)*, 2019.

[5] AWS Robotics, "aws-robotics/tts-ros1," *GitHub*. [Online]. Available: https://github.com/aws-robotics/tts-ros1. [Accessed: 02-Aug-2021].

[6] X. Tan, "SUCCESS-MURI/success_google_stt," *GitHub*. [Online]. Available: https://github.com/SUCCESS-MURI/success_google_stt. [Accessed: 02-Aug-2021].

Acknowledgements

The TTE REU Program is sponsored by NSF and the Center for Energy Efficient Electronics (NSF #0939514). I would like to thank my mentor Zhongyu Li, my Principal Investigator Koushil Sreenath, Nicole McIntyre, Tony Vo Hoang, Sam Mountain, my peers, and the Hybrid Robotics Group for their constant support.

Contact Information
Kavan Mehrizi
kavanmehrizi@[berkeley.edu],[gmail.com]

Support Information
This work was funded by the Hopper Dean Foundation and National Science Foundation Award #1757690.