Video-based human-machine interface for intuitive mobile robot control using depth camera

Max DeSantis, Dr. He Bai

Introduction

As mobile robots become more ubiquitous, there is a rising need for the development of intuitive and reliable human-machine interfaces (HMIs) that facilitate cooperation and safety. The HMI, which can range from the command line to touchscreens to voice control, directly influences the skill level necessary to guide a remote vehicle.

Example applications may include emergency rescue and space exploration. Both require operators that may be preoccupied, thus benefiting from a control mechanism that reduces cognitive load.

This work developed a prototype interface with the goal of significantly lowering the required skill for a human operator. It is modelled after Google Streetview, a familiar browser tool to explore cities at street level; it requires only point and click interaction. A depth camera is used to compute a 3D goal position from a user-inputted location on the camera image.

Objectives

Develop an intuitive interface allowing basic guidance of mobile robot with minimal cognitive load or necessary training.

Integrate into multiple mediums such as a web browser, discrete program, smart phone, wearable device.

Method

Robot Operating System (ROS)[1]

- Popular robotic development software framework
- Handles communication between interface and robot
- Community maintained packages - avoids reinventing wheel

Gazebo Simulation Environment [2]

- Simulates dynamics and sensors
- Stereo camera, planar lidar, inertial measurement unit
- Turtlebot3 robotic platform - standardized, supports ROS and Gazebo

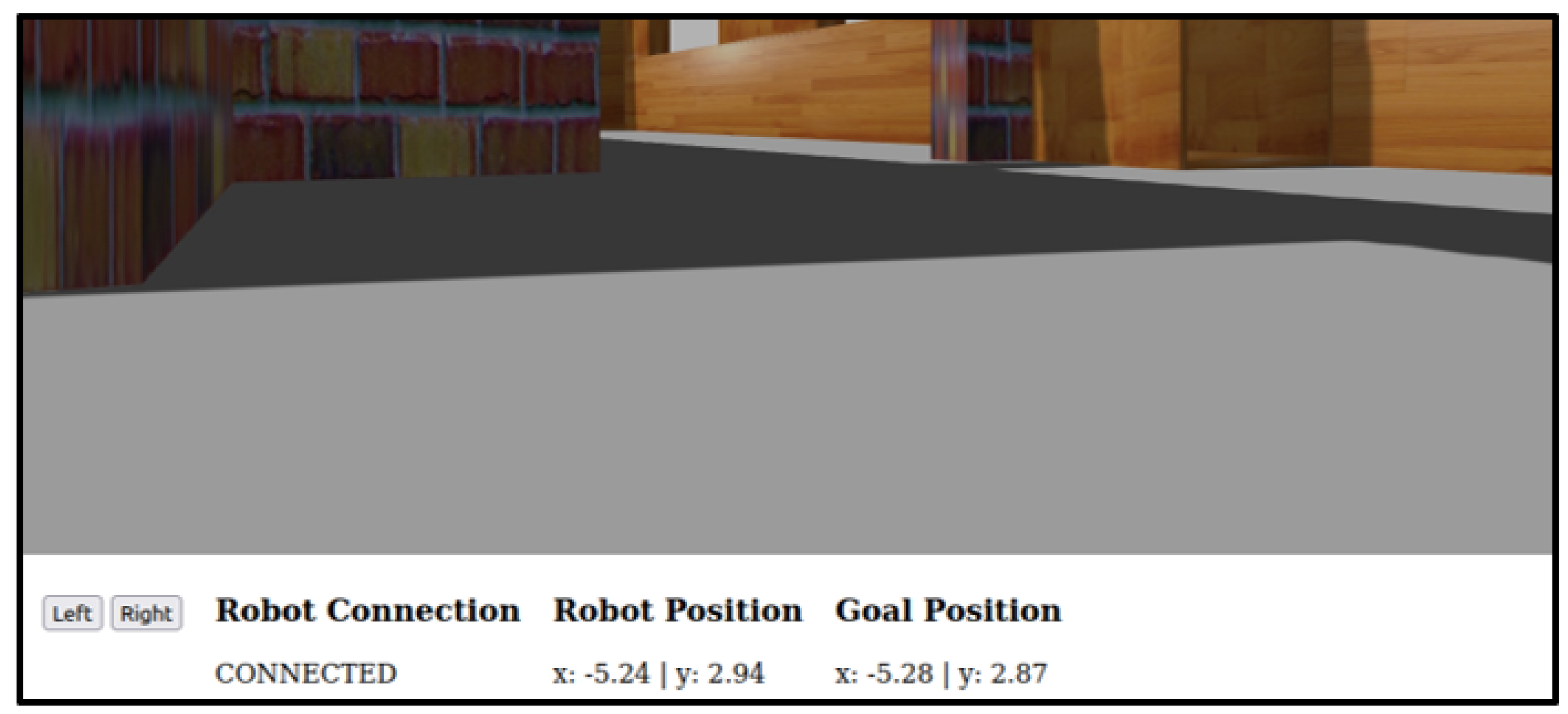
Interface Results

Implemented

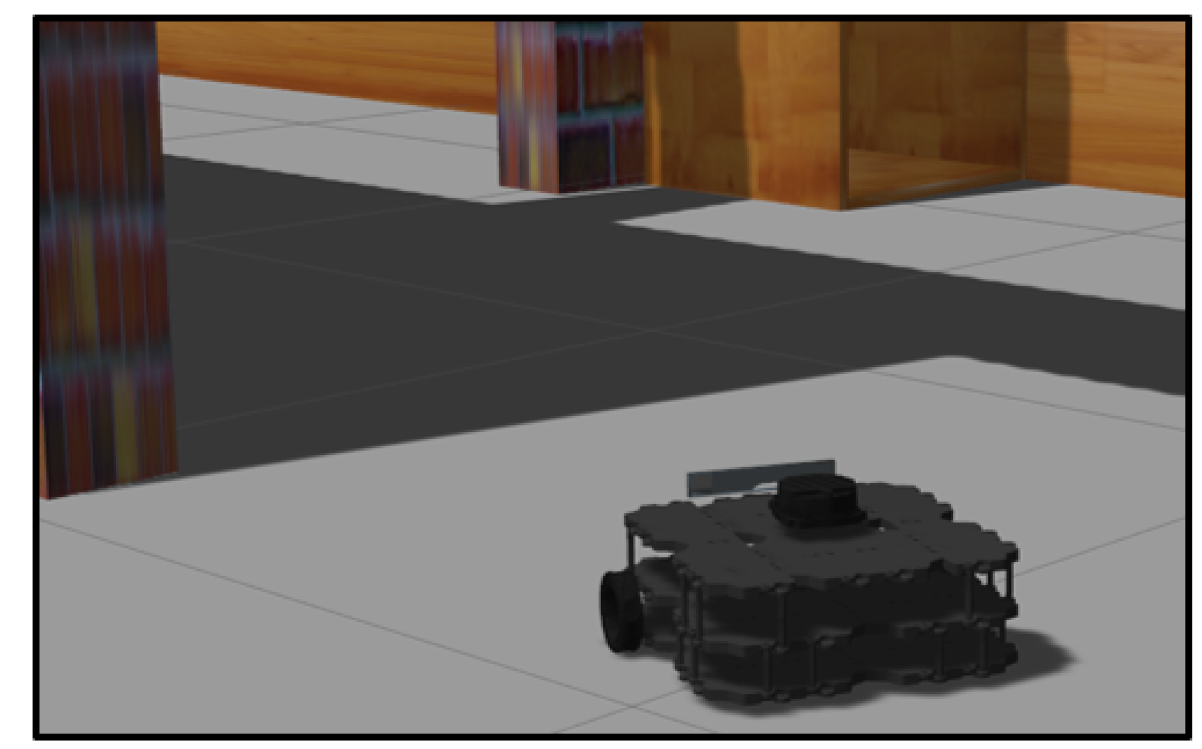
- Camera feed streamed over network to browser
- Click on visible location to select goal
- Manually rotate in place if necessary
- Robot's path planner handles the rest

Potential Modifications

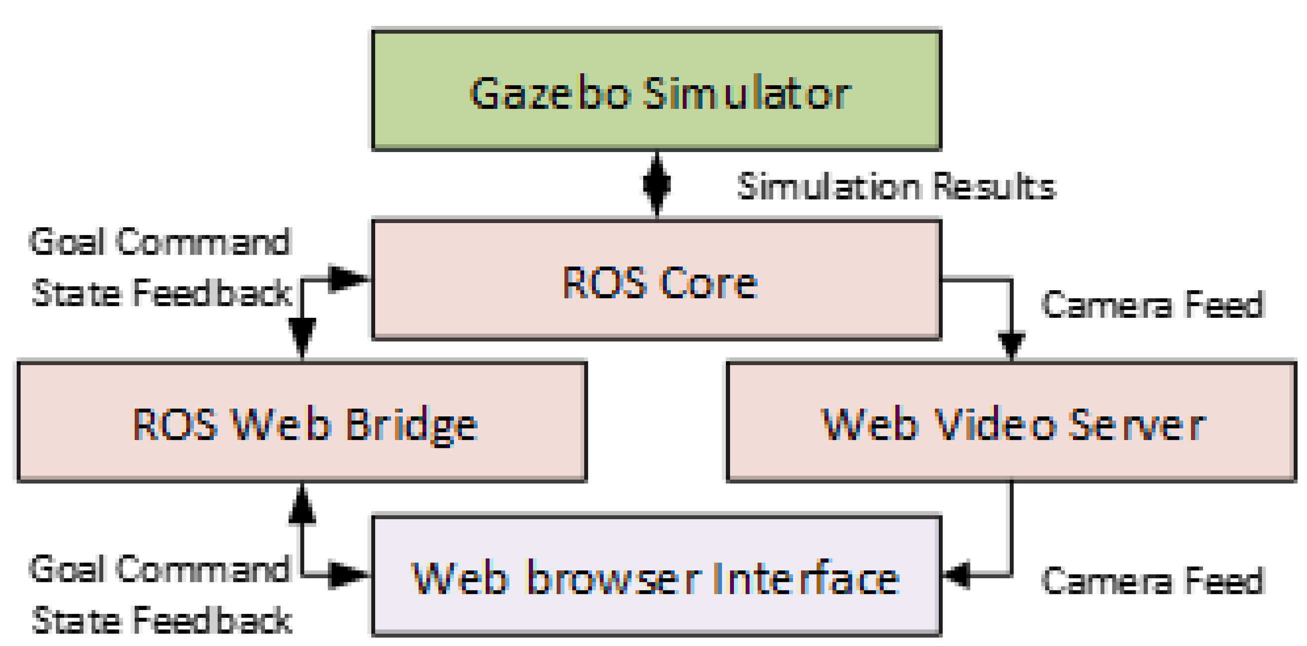
- Display overhead map generated by lidar to assist operator in localization
- Allow option for teleoperated control
- Relay additional state information



1: Operator view of cropped interface



2: Overhead view in Gazebo simulation environment



3: Software Architecture

Conclusions

- Functional prototype interface
- "Click and go" desired interactions
- Straightforward software/simulation implementation - hardware will require further work
- Browser functions well with ROS, other devices may require more attention
- Relies on performance of robot's on-board navigation stack

Future Work

- Embed visual indicators (heads up display) into camera feed
- Integrate with physical platform
- Develop smartphone interface
- Extend to other vehicles - Unmanned Aerial Vehicle, Unmanned Surface Vehicle
- Indicate navigable terrain to operator on camera feed
- Adapt AirSim simulator to support interface

References

[1] Morgan Quigley, Ken Conley, Brian Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, and Andrew Ng. Ros: an open-source robot operating system. volume 3, 01 2009.

[2] N. Koenig and A. Howard. Design and use paradigms for gazebo, an open-source multi-robot simulator. In *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566)*, volume 3, pages 2149–2154 vol.3, 2004.