2018 IEEE International Conference on Cognitive Computing

Autonomous Scooter Navigation for People with Mobility Challenges

Rajath Swaroop Mulky, Supradeep Koganti, Sneha Shahi, Kaikai Liu
Computer Engineering Department
San Jose State University (SJSU)
San Jose, CA, USA
Email: {rajathswaroop.mulky, supradeep.koganti, sneha.shahi, kaikai.liu}@sjsu.edu

***Abstract*—Despite the technical success of existing assistive technologies, for example, electric wheelchairs and scooters, they are still far from effective enough in helping the blind and elderly navigate to their destinations in a hassle-free manner. Riders often face challenges in driving scooters in some indoor and crowded places, especially on sidewalks with numerous obstacles and other pedestrians. People with certain disabilities, such as the blind, are often unable to drive their scooters well enough. In this paper, we propose to improve the safety and autonomy of the navigation by designing a cutting-edge autonomous scooter, which allows people with mobility challenges to navigate independently and safely in possibly unfamiliar surroundings. We focus on the localization and navigation challenges for the autonomous scooter where the current location, maps, and nearby obstacles are unknown. Solving these challenges will enable the scooter to both travel within buildings and perform tight maneuvers in densely crowds automatically.**

***Keywords*-Autonomous system, Sensor fusion, Scooter, SLAM**

I. Introduction

An intelligent mapping system is essential to provide a high-resolution understanding of the real physical world with dynamic obstacle avoidance. A variety of sensing and mapping technologies have been proposed for the autonomous navigation application scenarios, for example, Ultrasound-based [1], Lidar-based [2], and vision-based approaches [3]. Microsoft's Kinect and Google's Project Tango Tablet Development Kit are two very famous vision-based development platforms. Though these systems have many advantages, they also have their downsides. Vision-based depth information obtained from the region exposed to sunlight or covered by high reflective materials (tiles and glass doors) will not give an accurate measurement of depth. If the scene is featureless, for example, white walls with uniform texture, there will not be enough features detected for estimation of visual odometry. The other situation when this can happen is in a large environment where only a small area near the camera has depth readings and the other area is too deep, for example, in corridors and hallways. Self-driving cars from Google and Uber heavily rely on the Velodyne Lidar for their 3D mapping [2]. The personal mobility scooter developed in [4] also utilizes this kind of Lidar. However, the existing 3D Lidar devices are bulky and expensive at a minimum of \$10,000 per unit, which is not suitable for small and affordable mobility scooters (around \$1000 each). The performance of low-cost Lidar modules [5] can only achieve the coverage of several meters and not efficient enough to navigate a scooter through small scale places or avoid "on-road" dynamic obstacles, therefore lacking convenience and resilience.

In this paper, we propose to develop an intelligent autonomous scooter that assists people with independent transportation challenges towards their independent and dignified lifestyle in previously unmapped surroundings. Existing Lidar and multi-camera devices used in outdoor self-driving cars are too bulky and expensive for the scooter. The small safety margin and highly dynamic pedestrian walking patterns are more challenging than outdoor roads in terms of mobility. We propose to design and implement a new hybrid far-field and near-field mapping solution, targeting various cases from near-field fine-grained resolution to long range sparse coverage.

II. System Overview

A. System Module Design for the Scooter

To achieve the automatic steering control of the scooter, we need to install additional motors to control the steering wheel, speed, and direction.

Steering Control. As shown in Fig. 1, we made a few vendor-independent modifications in the mobility scooter to automate the steering control. We used a linear actuator (capable of 25 lb thrust with 4-inch movements) to push and pull the steering rod to the desired steering angle. This mechanism achieves better torque with simple installation than the servo motor. However, a major disadvantage of this mechanism is that the linear actuator has a higher error in linear actuation when compared with the precision servo motor. The same input to the linear actuator cannot promise the same position of the scooter.

To improve the control accuracy, a Proportional Integral Derivative (PID) controller has been used to actuate and turn the steering to the desired angle step-by-step with high degree accuracy. We utilize the Inertial Measurement Unit (IMU) to measure the actual angle and drive the linear actuator according to the angle differences. We mount the MPU-9250 sensor on the steering rod in the

978-1-5386-7241-9/18/$31.00 ©2018 IEEE
DOI 10.1109/ICCC.2018.00020

x-y plane. The angle can be calculated via $Angle = Atan(a_x/\sqrt{(a_y \times (a_y + a_z) \times a_z)})$, where a_x, a_y, a_z are sensor readings from the accelerometer in x,y,z directions. This mechanism ensures the scooter moving to the desired direction and reduces the chances of under-steering or over-steering. In the case of interruptions or road obstacles, for example, when the wheels tend to turn because of uneven road conditions, the PID and feedback loop would help the actuator to pull the steering back to its original direction in a short interval.

Speed/direction Control. The original mobility scooter has a fuse to limit the current from the power supply in order to maintain the legal speed limit of 5 mph. The output analog voltage from the potentiometer controls the speed. A similar control has been used for direction control: a voltage beyond 2.5v drives the mobility scooter in the reverse direction and a voltage below 2.5v drives it in the forward direction. There is another internal fuse located on the control circuit which limits the scooter from turning on if the throttle is not at the zero position. Due to this mechanism, we could not hack the controller by changing the supply of the analog signal without bypassing the fuse. Otherwise, the current of the control signal may likely exceed the limit of the fuse and cause problems. We propose to obtain the speed/direction control without breaking into the internal circuits. Specifically, we use two 7.4v precision servo motors (capable of delivering a maximum of 10N-m torque) to control the potentiometer by rotating its head without bypassing the fuse. This design can be applied to other vendor's mobility scooters with minimal changes.

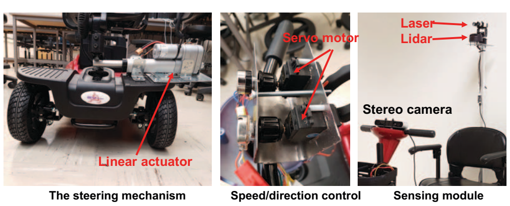


Figure 1. The steering mechanism, speed/direction control, and sensing module of the autonomous scooter.

Computing Module. We utilize existing open source computing boards, for example, the Nvidia Jetson TX2 module, Raspberry Pi, and Arduino. In this project, we make external sensors directly connectable with the Nvidia Jetson TX2. The Jetson TX2 runs Ubuntu 16.0.4 with ROS Kinetic. Most of the programs run on ROS platform are written in python or C.

To enable the open and easy development of the human interface module, we utilize an Android tablet as the front-end device. The Android tablet is connected to our hardware unit via the USB interface or Bluetooth-low-energy (BLE) interface. All the hardware commands and interactions will be wrapped as the Android APIs. To promote the innovation, we will open the Android APIs to other developers who want to program and develop new user applications for this autonomous scooter.

Sensing Module. We propose to utilize the long range eye safe laser ranging (up to 60 meters) to approximate a human's vision coverage and use stereo vision to help recognize the fine-grained world around them. The sensing module is shown in Fig. 1(c). Different from Lidar, laser ranging only works for a single point coverage and suffers under strong vibration by the motions. To achieve semi-Lidar functionality, we leverage the gyros-based pose data to compensate the laser motion in real time and create synthetic mapping of simple environments with regular shapes and deep hallways. Laser range finders are suitable for long ranges with limited resolution. Stereo vision, on the other hand, provides 3D structural data of nearby complex objects. To achieve simultaneous fine-grained resolution and long range coverage in the mapping of cluttered and complex environments, we dynamically fuse the measurements from the stereo vision camera system, the synthetic laser scanner, and the Lidar.

III. Indoor Mapping and Localization

A. *Simultaneous Localization and Mapping*

The large body of related work concerned with the task of mapping of the unknown environment is in the area of simultaneous localization and mapping (SLAM), which can be partitioned into two major sensor-based approaches: Lidar-based or vision-based. In terms of the SLAM algorithm, there are three major categories: Extended Kalman Filters (EKF) [6], Rao-Blackwellized particle filters (RBPF) [7], and graph optimization approaches [8]. Recent advances in incremental graph optimization allow for graph-based SLAM for online calculation, for example, KartoSLAM and Real-Time Appearance-Based Mapping (RTAB-MAP) [8]. Compared with other SLAM approaches, graph-based SLAM algorithms are usually more efficient, especially for large-scale environments.

B. *Hybrid 2D and 3D Map Fusion*

We utilize RTAB-MAP and stereo vision to get the fine-grained mapping of the 3D spatial world. However, in terms of the accuracy and robustness of the distance measurement, the stereo camera cannot compete with the laser ranger. John Leonard from MIT has a very good presentation and states "Elon Musk is Wrong: Why visual navigation of self-driving cars is far from solved". For example, Google's self-driving car utilizes the 3D Lidar to create the scene, while Tesla is heavily focused on using vision-based approaches.

We propose to perform data fusion for the long range laser and stereo vision. However, the location of the vision scene feature (ρ_m^p, where the upper script p means the screen's pixel space) is not in the same space and coordinate of the ranging results of the laser (d^{np}, where np is the polar coordinate of the navigation space). To facilitate the fusion process, we will convert these two measurements into the same navigation space (geodetic coordinate).

Based on the system's geodesic location $\mathbf{p}^n$ in the indoor environment (n is for the navigation coordinate), we can generate the location of the laser detection point via the transformation function as $\mathbf{p}^n_{laser} = f^n_{np}(d^{np}, \mathbf{p}^n, G_R)$, where function $f^n_{np}()$ utilizes the laser range, the pose and location of the system for the transformation. The G_R is the gyroscope results when the laser ranging measurements are taken. Since we utilize one laser beam to extend the range of a low-cost Lidar, we need to use the motion and pose information (G_R) to synthesize the mapping results of the laser over time, i.e., the *synthetic laser mapping*.

The visual feature point (ρ^v_m) in the real physical world is in the view angle (polar) coordinate of the navigation space. The transformation process to the navigation coordinate is $\mathbf{p}^n_{vision} = f^n_v(\rho^v_m, \mathbf{p}^n)$. The relation between ρ^p_m and ρ^v_m can be constrained by the camera projection process as $\rho^p_m = \mathbf{P}^p_v \rho^v_m$ with the following three steps: 1) map the ρ^v_m into camera view space; 2) convert to the Canonical view volume, i.e., NDC (Normalized Device Coordinate); and 3) map back to the screen pixel coordinate as ρ^p_m. Overall, the translation process can be modeled as

$$\mathbf{P}^p_v = \mathbf{K}^G \mathbf{D} = \mathbf{T}^p_\pi \mathbf{T}^\pi_v \mathbf{D} \tag{1}$$

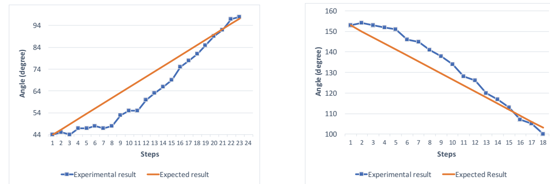
where $\mathbf{K}^G$ is the 3×3 modified camera intrinsic matrix in the OpenGL screen coordinate, and $\mathbf{D}$ is the camera extrinsic matrix that describes the camera's location coordinate transformation including rotation and translation. Thus, the intrinsic and extrinsic matrix combined as matrix $\mathbf{P}$ could illustrate the full perspective model which describes the relationship between a vision feature point and its projection in the OpenGL screen frame.

With all the projection model (1) and transformation functions $f^n_{np}()$ and f^n_v available, the laser measurement and the vision mapping can be transformed into the same navigation coordinate, i.e., fusing the $\mathbf{p}^n_{vision}$ and $\mathbf{p}^n_{laser}$. The data fusion process will be based on the occupancy grid mapping, where each grid represents the possibility that the area is occupied, empty or unknown.

IV. System Evaluation

A. Steering Control

To improve the steering angle control accuracy, we utilize the IMU to estimate the current pose of the steering rod, then apply the PID controller to turn the steering to the desired angle step-by-step. Fig. 2 demonstrates the angle convergence of the PID controller. The final angle of the steering rod is approaching to the desired angle in several steps with small error and no over-steering.



(a) Case 1 (b) Case 1

Figure 2. The angle convergence of the scooter steering rod.

B. Hybrid Near-field and Far-field 2D Mapping

The low-cost Lidar used in this project has a 360-degree horizontal view with a short range of 7-meters. This short coverage is not sufficient and reliable enough for an autonomous scooter. To overcome the coverage problem, we utilize the servo motor to drive the laser sensor and scan the front space with 120-degree coverage. We proposed a hybrid near-field and far-field mapping approach by fusing the long-range laser sensor and Lidar results. This hybrid mapping result is a synthetic 2D map with better coverage. We implemented the ROS driver for our hybrid approach and connected to the ROS rviz for the visualization.

1) Object mapping: To evaluate the effectiveness of our solution in mapping the object, we conduct experiments in Fig. 3 by detecting the obstacle at different distances (7-meter and 10-meter). The white dots in the Fig. 3 represent the Laser sensor mapping and the red dots represent the Lidar map in the ROS rviz. The results show that the Lidar can detect the object in 7-meter range, but missed the target in 10-meter range. Using the hybrid Lidar and long range laser result, we can detect the target in longer distance and extend the coverage of the 2D mapping.

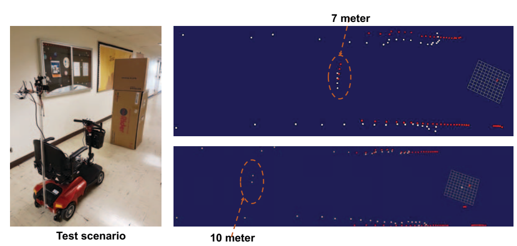


Figure 3. The object mapping experiment when the object is in 7 meter and 10 meter, respectively.

2) Wall mapping: To evaluate the mapping results in an indoor crowded environment, we conduct the experiment in Fig. 4 to estimate the front wall. The laser sensor is facing to the front and it can only detect the front wall; while the Lidar sensor is 360 degree with wide coverage. Fig. 4 demonstrates the aligned data from Lidar and Laser. The Lidar was able to capture the shape of a wall with high precision. The Laser missed few points of detection and the detected wall is not in its original shape. The problem is caused by rotating the servo motor. We have to use the motor to drive the laser to scan the wall in different time period, which causes some drift in terms of ranging accuracy. Thus, we can conclude that Lidar achieves accurate mapping with short range; while laser achieves long range detection with reduced accuracy in terms of shape. By fusing these two results, our hybrid solution can achieve better coverage with accurate shape estimation.

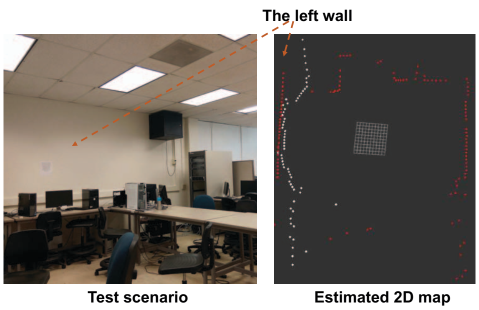


Figure 4. The indoor mapping results of the wall.

3) 2D Mapping: Fig. 5 shows the experimental results of estimating the 2D indoor map with three pre-defined obstacles. When comparing with the ground truth, our estimated 2D map is very close to the ground truth and all the obstacles have been detected with very accurate position estimation. The location trace point shown in Fig. 5 demonstrate accurate location estimation of the scooter with very small error gap when compared with the group truth.

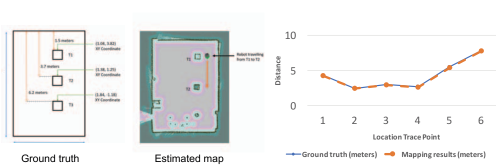


Figure 5. The comparison of the 2D mapping results and location trace estimation.

C. Hybrid 2D and 3D Mapping

(a) Environment (b) Point cloud

Figure 6. The 3D mapping (point cloud) result of the indoor corridor.

Visual SLAM (vSLAM) builds a dense 3D model of the scene as it moves through it and also creates a trajectory of the camera. Since 2005, a lot of research has been conducted on vSLAM because it has the capability to extract landmarks all over the scene and overcome the challenges of the 2D Lidar. We utilize RTAB-Map, which is a Graph-Based SLAM approach based on an incremental appearance-based loop closure detector. RTAB-Map is constructed using RGB images, depth data, and visual words. Fig. 6 shows our vSLAM result in one indoor scenario.

V. Conclusion

While mobility scooters may help to improve the quality of life of their users, operating them is still challenging in many scenarios. In this paper, we designed an autonomous system from the ground up to help people with mobility challenges to better navigate, explore the physical world, and connect with friends and family. We propose a novel intelligent autonomous scooter with short-range vision-based mapping and long range hybrid indoor mapping.

Acknowledgment

The work presented in this paper is funded by National Science Foundation under Grant No. CNS 1637371.

References

[1] Y. Jin, J. Kim, B. Kim, R. Mallipeddi, and M. Lee, "Smart cane: Face recognition system for blind," in *Proceedings of the 3rd International Conference on Human-Agent Interaction*. ACM, 2015, pp. 145–148.

[2] Velodyne lidar. [Online]. Available: http://velodynelidar.com

[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*, pp. 1–19, 2016.

[4] H. Andersen, Y. H. Eng, W. K. Leong, C. Zhang, H. X. Kong, S. Pendleton, M. H. Ang, and D. Rus, "Autonomous personal mobility scooter for multi-class mobility-on-demand service," in *Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on*. IEEE, 2016, pp. 1753–1760.

[5] Rplidar a2 360 laser scanner. [Online]. Available: https://www.slamtec.com/en/Lidar

[6] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: part i," *IEEE robotics & automation magazine*, vol. 13, no. 2, pp. 99–110, 2006.

[7] B. D. Gouveia, D. Portugal, and L. Marques, "Speeding up rao-blackwellized particle filter slam with a multithreaded architecture," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2014, pp. 1583–1588.

[8] M. Labbe and F. Michaud, "Online Global Loop Closure Detection for Large-Scale Multi-Session Graph-Based SLAM," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sept 2014, pp. 2661–2666.