Name: Fanzhe Lyu
Name of Project Advisor: Fumin Zhang
Group Name: XXLs(team 30)
Visual-Based Navigation and Obstacle Avoidance for Robots

1 Introduction

Collision-free navigation is an essential component of autonomous robots, ensuring that the robot gets from one point in space to another in a safe manner. Cameras, including depth cameras, RGB cameras, and stereo cameras, have been adopted to detect obstacles and to help robots avoid them. Because cameras are more accessible than sensors such as LIDARs, visual-based navigation, and obstacle avoidance have increasing popularity nowadays. This technical review briefly summarizes some commercial applications of visual-based navigation, explains underlying technology, and provides information of implementation for ideal operation.

2 Commercial Applications of Visual-Based Navigation

2.1 Quadcopters

Many quadcopters leverage cameras to provide features of navigation and obstacle avoidance. DJI equips most of its products with cameras and provides the capability for visual-based obstacle avoidance. One of its newly launched product, DJI Spark($400-$500), equips a single RGB camera with 81.9 °field of view, and 25mm focal length [1]. It could sense obstacles with a distance between 0.2m and 1.5m in the forward direction under sufficient lighting condition, and could avoid sensed obstacles when it performs tasks such as return-to-home and object tracking. One of DJI's competitors is Yuneec Typhoon H($900-$1000). Equipped with an Intel RealSense stereo cameras module, Yuneec Typhoon H could sense and avoid obstacles with distances between 9.8ft(3m) to 23ft(7m), and perform similar task sets [2].

2.2 Driving assistance

Existing automobiles incorporate cameras to detect possible collisions. Tesla Model 3, with a starting price of $35000, equips 8 surrounded cameras to provide 360°visibility with 250 meters of range [3]. Based on camera feedback, Tesla autopilot software provides semi-autonomous driving assistance and can control thrust and brake. Other companies such as Waymo [4] and Cruise [5] have autonomous driving products under development. They combine cameras and LIDARs to achieve collision-free autonomous driving.

3 The Technology involved in Visual-Based Navigation

Visual-based navigation and obstacle avoidance require methods and algorithms for robots to detect obstacles and then react to the identified obstacles [6]. This section reviews technology involved in detecting possible collisions, and responding towards possible collisions.

3.1 Perception

Cameras are the primary sensors used in visual-based navigation and obstacle avoidance. Commonly used perception pipeline requires distances of seen objects. For different types of cameras, methods slightly differ [7]. Depth cameras can provide ranges directly. Stereo camera systems exploit binocular disparities and triangular geometry to provide an estimation of distances. RGB cameras can be combined with IMU sensors or GPS sensors to generate measurements [8, 9]. Cameras are not always perfect: noisy and uncertain sensor data might decrease the accuracy of object detection and impair the robustness of the system. Mapping obstacles into either 2-D or 3-D occupancy grid addresses noise [10]. Strategies such as image segmentation and image classification are developed and incorporated to enhance the accuracy of obstacle detection.

3.2 Decision

After sensing possible collision objects and mapping them into occupancy grids, robots require control policies to avoid detected obstacles. Commonly used methods include reactive methods, velocity space methods, and optimal trajectories synthesis.

Reactive methods, such as the Vector Field Histogram (VFH) [11]method, use the locally sensed obstacle space and the current goal point to identify an immediately applied control policy. VFH uses a local polar (histogram) representation for local points in the global map relative to the current robot pose. Processing of the polar histogram generates new steering commands.

Velocity space methods, such as Dynamic Window Approach (DWA) [12], involve sampling and scoring of possible control commands. DWA samples from a discrete set of relative control or velocity changes from the current control or velocity. It performs scoring and collision checking on each trajectory regarding the occupancy grids.

Bridging the gap between reactive planners and velocity space planners are continuous trajectory synthesis implementations on the occupancy grids. Optimal trajectory synthesis planners such as Elastic Bands(EB) [13] and Timed Elastic Bands(TEB) [14] utilize occupancy grids as cost functions, together with kinematics and dynamics constraints, to generate optimal paths.

3.3 Improvements and Trends

A lot of research has focused on improving either perception or decision for robots. Custom Hardware implementations of algorithms have been developed to provide a faster and more robust perception [15]. State-of-the-art research also leverage deep learning and neuroscience to create a better understanding of the sensed scene, and control decision towards possible collisions [16].

4 Implementation of Visual Navigation and Obstacle Avoidance

Implementation of visual navigation and obstacle avoidance involve both hardware and software components. Hardware implementation consists of cameras, embedded computers. Selection of camera should balance weight and resolution. An Existing synchronized stereo camera module from ELP [17] might be ideal for its weight. Embedded computers should be powerful enough to execute algorithms. Raspberry Pi, commonly used as robot controllers, is ideal and easily accessible.

Software implementation requires developing a robust perception and decision algorithms. Software components run on embedded computers, process data from cameras, and send a decision to actuators. Developing and testing algorithms in a real environment is not ideal because cameras are fragile, and simple collision may break the module. Thus the development of software should be under a simulated environment such as Gazebo.

References

[1] DJI, "Dji spark – specs, tutorials & guides." [Online]. Available: https://www.dji.com/spark/info [Accessed: 2019-09-30]

[2] Yuneec, "Yuneec typhoon h." [Online]. Available: https://us.yuneec.com/typhoon-h-overview [Accessed: 2019-09-30]

[3] *Tesla 3 Owner's Mannual*, 2019th ed., Tesla, May 2019. [Online]. Available: https://www.tesla.com/sites/default/files/model_3_owners_manual_north_america_en.pdf

[4] Waymo, "Technology - waymo." [Online]. Available: https://waymo.com/tech/ [Accessed: 2019-09-30]

[5] Cruise, "Technology." [Online]. Available: https://getcruise.com/technology [Accessed: 2019-09-30]

[6] G. N. Desouza and A. C. Kak, "Vision for mobile robot navigation: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 2, pp. 237–267, Feb 2002.

[7] R. Szeliski, *Computer vision: algorithms and applications*. Springer Science & Business Media, 2011.

[8] Z. Yang and S. Shen, "Monocular visual-inertial fusion with online initialization and camera-imu calibration," in *Proc. 2015 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, Oct 2015, pp. 1–8.

[9] Z. Yang and S. Shen, "Monocular visual–inertial state estimation with online initialization and camera–imu extrinsic calibration," *IEEE Transactions on Automation Science and Engineering*, vol. 14, no. 1, pp. 39–51, Jan 2017.

[10] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics*. MIT Press, 2005.

[11] J. Borenstein and Y. Koren, “The vector field histogram-fast obstacle avoidance for mobile robots,” *IEEE Transactions on Robotics and Automation*, vol. 7, no. 3, pp. 278–288, June 1991.

[12] D. Fox, W. Burgard, and S. Thrun, “The dynamic window approach to collision avoidance,” *IEEE Robotics Automation Magazine*, vol. 4, no. 1, pp. 23–33, March 1997.

[13] S. Quinlan and O. Khatib, “Elastic bands: Connecting path planning and control,” in *Proc. IEEE International Conference on Robotics and Automation*. IEEE, 1993, pp. 802–807.

[14] C. Rösmann, F. Hoffmann, and T. Bertram, “Timed-elastic-bands for time-optimal point-to-point nonlinear model predictive control,” in *Proc. 2015 European Control Conference (ECC)*, July 2015, pp. 3352–3357.

[15] S. Jin, J. Cho, X. D. Pham, K. M. Lee, S. Park, M. Kim, and J. W. Jeon, “Fpga design and implementation of a real-time stereo vision system,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 1, pp. 15–26, Jan 2010.

[16] M. Pfeiffer, M. Schaeuble, J. Nieto, R. Siegwart, and C. Cadena, “From perception to decision: A data-driven approach to end-to-end motion planning for autonomous ground robots,” in *Proc. 2017 IEEE International Conference on Robotics and Automation (ICRA)*, May 2017, pp. 1527–1533.

[17] ELP, “Elp synchronous usb stereo camera module double undistortion lens.” [Online]. Available: http://www.webcamerausb.com/elp-synchronous-usb-stereo-camera-module-double-undistortion-lens-960p-ov9750-usb20-otg-webcam-compatible-with-windows-linux-android-mac-p-285.html [Accessed: 2019-09-30]