Stationary LIDAR Sensors for Indoor Quadcopter Localization

Marcell Rausch, Gabor Feher
Department of Telecommunication and Media Informatics
Budapest University of Technology and Economics, Budapest, Hungary

***Abstract*—The abstract goes here.**

I. Introduction

A prerequisite for future expansion of mobile robots is the ability to navigate autonomously and safely in their ever-changing environment without collisions. Simultaneous Localization And Mapping (SLAM) algorithms discover the surroundings, create a map and localize the robot in the map with the help of ranging or visual sensors. In an indoor scenario, this sensor is often the Light Detection and Ranging (LIDAR) sensor.

Many papers can be found about SLAM and LIDAR sensors that are capable of building highly accurate maps of building-interiors or even outdoors. (references!) However, they are using precise, long range rotating planner scanners, which can be carried by large an industrial-grade hexacopter or octocopter only, and due to their weight, they are not suitable for smaller quadcopters. Another frequent choice for drones are multiple cameras and image processing. The issue here is the rather intense computation demand, which is hard to satisfy on a small quadcopter with limited power supply.

In this paper, the evaluation of a SLAM system is introduced that uses stationary VL53L1X LIDAR sensors to scan the drone's environment and uses Cartographer SLAM to construct maps. This solution might provide lightweight, cheap, and computationally light alternative to currently used planar scanners.

Each setup is evaluated based on data collected from Gazebo simulator and LIDAR sensors are simulated with parameters determined by measurements with VL53L1X sensors. On the collected data multiple sensors layouts and settings were evaluated to find the optimal number of sensors and achieve balance between accuracy and simplicity.

II. Ranging sensor-based SLAM

Cameras are cheap, popular but they demand high processing capabilities. In contrast, ranging sensor-based approaches produce a lower amount of data and thus they are less resource-intensive. Ultrasonic sensors have been used for distance measurements for many years, but they suffer from significant measurement noise. LIDAR sensors on the other hand uses light to measure distance, offer more accurate measurements, typically lower form factor, higher resolution, and update rates. Planar scanners for 2D or 3D scanning typically use a rotating LIDAR sensor that makes evenly distributed measurements.

Stationary LIDAR sensor-based modules can be found for collision avoidance. TeraBee offers an off-the-shelf system with 8 selectable LIDAR sensors. The company offers long-range and fast-ranging versions, that can range up to 60m or 600Hz.

Bitcraze has developed a Multi-ranger deck for their mini drone that weights only 27g. This deck carries 5 VL53L1X lidar sensors, 4 horizontally and 1 vertically. The company website has a SLAM demo, but provided no information of the SLAM algorithm in use.

III. The VL53L1X LIDAR sensor

VL53L1X is described as a long-distance ranging Time-of-Flight sensor by STMicroelectronics [1]. The size of the module is 4.9x2.5x1.56mm, it emits 940nm invisible laser and uses a single-photon avalanche diode (SPAD) receiving array with a field-of-view(FOV) of 27 degrees. The module runs all the digital signal processing on a built-in low power microcontroller that frees up the host processor from these tasks. According to the datasheet, the sensor is capable of 50Hz ranging frequency and up to 400cm distance measurements. The size of the laser receiving array can be programmatically changed by setting the region-of-interest(ROI) size.

The most relevant settings of the sensor are the timing budget, distance preset and the resolution. The timing budget sets the maximum time available for the sensor to complete ranging. It can be adjusted between 18.5ms to 1s. Distance preset sets the expected range distance and by moving the ROI across the whole field-of-view, higher scan resolution can be achieved. Using this method the resolution can be extended to 4x4, without overlapping segments.

To simulate the VL53L1X sensor accurately, a series of measurements were conducted on distances from 0.5m to 4m in 0.5m steps with selected sensor settings. On each distance, the average measured distance, standard deviation, sampling rate, and typical range status were measured with the selected settings. Measurements were done with sensor resolutions 1x1, 2x2, 3x3, 4x4, and all distance presets: short, medium, long. It is unnecessary to test the sensor with all possible timing budget values, so 4 values have been selected. 18.5 ms provides the highest sampling rate of 50Hz, but can only be used with short-distance preset. 33ms is the lowest timing that can be used with long-distance preset. Based on the datasheet, 140ms is the lowest budget that can provide ranges up to 400cm and lastly, 200ms is expected to provide less noisy measurements. The last three values are also used in the VL53L1X datasheet[1] to demonstrate ranging accuracy.

Based on the completed measurements, it can be seen that sampling time is independent of the distance preset and is directly proportional to the resolution and the timing budget. On each tested scenarios, the most accurate measurements were always with resolution 1x1. By increasing the resolution, therefore lowering the receiver SPAD array size, the measured distance starts to deviate from the real distance, and standard deviation increases. This can be seen at 2.5m distance with medium preset. Long-distance preset however produces more accurate measurements even on 3m, so it is preferred over medium preset. Measurements with long-distance preset and 140ms timing budget follow the real distances even on 3 meters but start to deviate at 3.5 meters with 2x2-resolution and above. Raising the timing budget to 200ms reduces the error and standard deviation, but the error is still significant.

18.5ms timing budget with short distance preset can produce the fastest ranging measurements up to 1.6m according to the datasheet. The short preset also has the best ambient light immunity, a favorable quality. Surprisingly, in the tests, the sensor with these settings was able to measure up to 2m with 7mm standard deviation only. Above 2m no ranges were successful. Compared to other presets, the sampling time for each resolution decreased significantly. In the case of 4x4-resolution, the sampling time has decreased from 390ms to 265ms and the maximum distance decreased from 2.5m to about 1m. The increase in the update rate might be beneficial for the SLAM algorithm.

TODO

IV. 2D SLAM EVALUATION

The fair way to compare the performance of the SLAM algortihm with different LIDAR layouts is to use the same dataset for every setup. For this reason, there were simulated LIDAR scanners on the top and bottom of the drone that evenly scanned the top and bottom hemispheres. The complete scan cover the whole sphere around the drone with ranges evenly distributed along all axes, capable of even 3D SLAM evaluation. The drone was flown around inside a simulated building with corridors, doors, windows, small and big rooms, while all sensor data was recorded. Later, the recorded data was filtered according to the given sensor configuration and played back to the SLAM algorithm for map creation and pose-estimation to the evaluation. Cartographer also used the IMU data as an initial guess for the pose and the accelerometer to determine the direction of gravity. In the evaluation the RMS error was calculated between the trajectory calculated by the SLAM and the ground-truth trajectory, which was provided by the simulator. The total length of the recorded flight was 144 seconds. The LIDAR distance measurements were published on 30Hz each and IMU data was published on 50Hz.

The LIDAR parameters were changed incrementally one by one. In the first simulation unfiltered range data was used, which had the maximum possible sampling rate, 4m maximum distance, and 128x128-resolution in a total of the top and bottom hemispheres. This produced the most accurate map and localization estimates, compared to the other down-sampled dataset. Second, the effect of the LIDAR resolution was tested. The initial measurements were filtered as if 13 LIDAR sensors would have been placed evenly on the horizontal axis, covering the full 360 degree. The resolution of the sensors were set to 4x4, 3x3, 2x2 and 1x1. To simulate the behavior of VL53L1X, ranges inside the sensor's field of view were averaged. This filtering caused artifacts like curved corners and transition points that occured when an object was only partly in the field of view and another distant object covered the rest of the ranging. These smooth transitions hide the sharp edges of walls, making it harder to insert points and to find loop-closures. Third, the resolution with the best performance from the second step was used with various sampling rates of the actual sensor. Forth, the maximum distance was lowered to realistic, while keeping the same layout of the sensors. Finally, the number of sensors was reduced until the lowest number, where the setup was still capable of producing a decent quality map.

A. *Effects of LIDAR resolution*

First, unfiltered data ranges with maximum rate and distance were used. Because there was no noise at all, even without a global SLAM, a decent quality map could be built. The extracted trajectory closely followed the ground truth trajectory. The RMS error on the x- was 6.3cm and a bit less 6.5cm on the y-axis, resulting in a cumulated 12.8cm ¡¡¡????¿¿¿¿ RMS error.

1) 4x4-resolution: In this scenario the filter selected 4 times 52 points around the quadcopter in 4 rows. This scenario had many artifacts and transition points, still, the estimated trajectory followed accurately the ground-truth trajectory. There were no constant offsets on neither axes. The RMS error of this setup was only 6cm more than the error of unfiltered data. The RMS error was 8.12cm on the x- and 10.48cm on the y-axis, resulting in a cumulated error of 18.6cm.

2) 3x3-resolution: As the resolution is lowered to 3x3, the field of view of each region of interest had broadened. This caused the artifacts coming from averaging to be more significant. The number of ranges per scan was also lowered, so the ratio of transition points and wall hits increased heavily impacting the performance of the SLAM. Tuning on the SLAM algorithm helped to produce a decent quality map. The global SLAM was able to compensate drifting submaps. Although the mapping accuracy was reduced along the reduced sensor resolution, the localization accuracy had increased. The cumulated RMS error for this setup was 11.34cm. The RMS error on the x- was 4.04cm and on the y-axis it was 7.3cm. The accuracy of this setup is the result of the better SLAM parameter tuning.

3) 2x2 and 1x1 resolutions: Compared to the 3x3-resolution, these scenario contains relevantly fewer points, which makes scan matching significantly less efficient. Tuning of this setup was noticeably harder and the quality of the produced map was highly decreased. The grid resolution of the map should be reduced to low resolution. While the configuration with a high-resolution grid map was able to follow the ground-truth trajectory in small rooms and the narrow corridor, it lost track in the big and spacious room, when the ranges were far apart.

In the case of the 2x2-resolution, the high-resolution trajectory was at some points as far as 1.4m from the ground-truth pose, but the cumulated RMS error was only 43.4cm. This shows that a seemingly small difference in RMS error can mean an error of more than a meter at a part of the trajectory. The low-resolution trajectory in comparison has a cumulated RMS error of 22.13cm. 7.95cm error on the x- and 14.17cm on the y-axis. The 1x1-resolution scenario did not produce any usable map or localization.

These scenarios were unusable for the SLAM, so they were not tested further.

B. *Effects of sampling rate*

For the 4x4-resolution 400ms sampling time was chosen, based on the real sensor measurements. The timing budget of 33ms was selected over the 18.5ms because it has a lower sampling rate. The dropped sampling rate from 30Hz to 2.5Hz was significant and made the tuning much harder. Two different grid resolution were tested as well. The quality of the map was decent in both cases, no major flaws were visible. The global SLAM had to compensate rotation drifts, that occurred when the drone turned around in the big room. Low grid resolution produced a smoother path, while on high-resolution grid there were some edges and sharper turns. In general, the configuration with a higher grid resolution felt very unstable.

The cumulated RMS error on a high grid resolution was 16.49cm, on low grid resolution it was higher: 18.95cm. Higher grid-resolution results in greater accuracy in position estimates, but its stability is lower in comparison to lower resolution.

As for the 3x3-resolution, the same timing budget was chosen, but this setup has a better sampling time of 220ms compared to the previous 400ms. This results in an 80% increase in update rate to 4.5Hz, which also means almost twice more often pose updates. Lowering the grid resolution was necessary here as well. The map quality was adequate, global SLAM almost hasn't done any corrections. The cumulated RMS error was 14.67cm, 6.4cm on the x-axis, and 8.2 on the y-axis. Because of stability and accuracy, 3x3-resolution seems like a better resolution for Cartographer SLAM.

C. *Effects of maximum distance*

The VL53L1X sensor measurements showed that each setup works accurately until 2.5m and each has about 25cm drift at a distance of 3m. This way 2.75m as maximum distance was chosen for the tests. The lowered maximum distance had a positive effect, that the number of transition points has been reduced. Because of the shorter distance the field of view could not open as wide as it could before. The downside of the limited range is that the algorithm loses track in the bigger spaces. This caused more flaws and a small tilt can be observed. The trajectory using sensors with 3x3-resolution has a significant offset on the y-axis, which was picked up due to less reference points. Similar scans were pulled together, resulting in a shortened corridor. The trajectory using the setup with 4x4-resolution sensors has a smaller offset. It seems like with more points per scan the SLAM algorithm was able to pick up small details and better track the movement of the drone on this mostly even corridor. On the other hand, it failed in the big room, where the same situation appeared.

The cumulated RMS error for 3x3-resolution is 45cm and 49.6cm for 4x4. In the end the mostly constant offset with otherwise correct shape results in a smaller error, than the pose estimation that follows ground-truth closely but failed in shape at a relatively small area.

D. *Reduced set of evenly distributed LIDAR sensors*

In this test, the goal was to find the minimum number of LIDAR sensors, with which the SLAM algorithm still works reliably and can produce a coherent map. The sensors were set to 4x4 and 3x3-resolution, 400ms update rate and 2.75m as maximum distance. The tuning of the algorithm proved to be significantly harder this time. Testing multiple setups, finally 8 sensors were successful in the case of 4x4-resolution sensors. The RMS cumulated error was 34.1cm. The 3x3-resolution failed. Therefore this setup is assumed to be the lowest number of VL53L1X sensors, that is still able to produce consistent maps with Cartographer SLAM in 2D mode.

V. Conclusion

Acknowledgment

The authors would like to thank...

References

[1] VL53L1X, "Datasheet," www.st.com/resource/en/datasheet/vl53l1x.pdf, accessed: 2019-11-30.