Comparative Evaluation of Time-of-Flight Depth-Imaging Sensors for Mapping and SLAM Applications

Lance Fang [1], Alex Fisher [1], Stefan Kiss [2], James Kennedy [2], Chatura Nagahawatte [2], Reece Clothier [1], and Jennifer L. Palmer [2]

[1] RMIT University
Bundoora, VIC 3083
{s3379130, alex.fisher}@rmit.edu.au

[2] Defence Science & Technology Group
506 Lorimer St, Fishermans Bend, VIC 3207
{Stefan.Kiss, James.Kennedy, Chatura.Nagahawatte, Jennifer.Palmer}@dsto.defence.gov.au

Abstract

Autonomous robotic systems rely on simultaneous localisation and mapping (SLAM) algorithms that use ranging or other sensory data as input. Numerous algorithms have been developed and demonstrated, many of which utilise data from high-precision ranging instruments. Small unmanned aircraft systems (UAS) have significant restrictions on the weight of sensors they can carry, and light-weight ranging sensors tend to be subject to more error than their larger counterparts. The effect of these errors on SLAM effectiveness will depend on the algorithm in use. Our current work is focussed on evaluating different combinations of sensor and algorithm. This paper presents an evaluation of the performance of three SLAM algorithms that are freely available in the Robot Operating System (ROS), in conjunction with ranging data from two different time-of-flight imaging cameras: a commercially available Mesa Imaging sensor and a prototype sensor based on single-photon avalanche diode (SPAD) technology. Based on the results of this data collection, a difference with respect to the ability of the SLAM algorithms to handle noisy odometry data can be seen. GMapping is able to generate maps consistently when compared with KartoSLAM and Hector Mapping algorithms. However, KartoSLAM was able to create maps that represented the ground truth more accurately.

1 Introduction

The Defence Science and Technology (DST) Group and its academic collaborators are developing technologies to enable the effective use of autonomous land-based and airborne vehicles in complex, urban environments. Small ground robots and 'micro' unmanned aircraft systems (UAS), typically defined as aircraft with a maximum dimension of 1 m or less, are of interest for missions in urban environments such as: emergency response; intelligence, surveillance, and reconnaissance; humanitarian assistance; and disaster relief.

In areas where global positioning system (GPS) signals are unavailable or unreliable, autonomous robots typically navigate using simultaneous localisation and mapping (SLAM) algorithms with range or imagery data acquired on-board the platform. Range data is commonly provided by 2D or 3D laser-based light detection and ranging (LIDAR) or radar sensors. However, high-precision ranging sensors are often bulky and heavy; and this makes them impractical for application on very small platforms, such as micro UAS.

The goal of the present work is to examine the mapping and SLAM performance of several time-of-flight (ToF) cameras that function as solid-state LIDAR sensors: a single-photon avalanche-diode (SPAD) camera fabricated by Politecnico di Milano and a Mesa Instruments SwissRanger (SR) 4000 sensor with a calibrated range of 8 m. The study is motivated by the fact that sensors of this type are being produced in microchip-based form and are thus capable of meeting the size, weight, and power constraints of multi-rotor micro-UAS. SPAD sensors are also useful as low-light cameras [Zappa *et al*., 2007]. A better understanding of their use with SLAM algorithms is thus desired.

In conducting SLAM, range data is typically combined with odometry from on-board or off-board systems (*e.g*., inertial-measurement units or GPS data, if available). Pose estimates and maps created by SLAM algorithms are thus susceptible to errors from each of these sources; and each has different characteristics. How these errors propagate to the final pose and map estimates also depends on the nature of the SLAM algorithm and how it is configured.

In the experiments described here, the effect of errors inherent in the range measurements has been isolated from errors in the odometry through the use of a motion-capture system to supply the necessary odometry data. The motion-capture system provides sub-millimetre positional accuracy and small angular errors (small pose errors), while range errors on the order of centimetres or 100's of centimetres are produced by the ToF sensors. This 'nominally perfect' odometry data can also be perturbed with random noise to simulate real-world conditions, to test both the localisation and mapping performance of the SLAM algorithm.

The data utilised here is part of a larger dataset collected with ten different ranging sensors. The sensors employed include laser-based scanning instruments that

create 2D and 3D clouds of returns from surfaces in the laboratory (*e.g.*, a Hokuyo URG-04LN and a Velodyne Puck sensor, respectively). Other sensors utilise structured light to create 2D images from which depth information is extracted (*e.g.*, a 1st-generation Microsoft Kinect sensor). The broader goal of the project is to compare the performance of different SLAM algorithm and sensor pairings; however, the current focus is on the ToF cameras mentioned above.

2 Background

The literature is rich with datasets consisting of a set of measurements from a single depth sensor and the associated ground truth [Abdallah *et al*., 2006; Balaguer *et al*., 2007; Handa *et al*., 2014; Nguyen *et al*., 2007; Sturm *et al*., 2011]. Common datasets are essential for comparing the performance of different SLAM algorithms. However, relatively few exist that contain measurements from a variety of sensors under the same conditions and in the same environment. Such a dataset is useful not only for comparing SLAM algorithms, but also for comparing different *combinations* of algorithm and sensor. The broader aim of this work is to address this gap by supplying datasets for ten ranging sensors operating in an environment with essentially error-free odometry and a well-characterised, quasi-2D layout.

Evaluating the localisation performance of a SLAM algorithm against ground truth is relatively straightforward. It is typically done using the mean squared translational and rotational errors of the computed pose as a function of time, compared with the actual pose (the ground-truth pose).

Comparing mapping performance is more difficult, because there is no obvious way to generate an error metric for a map. Santos *et al*. [2013] performed a comparison of several SLAM algorithms that are available as part of the Robot Operating System (ROS), using a simulated environment and sensor measurements. The algorithms tested included GMapping, KartoSLAM, and Hector Mapping. To enable comparisons, an error metric for occupancy-grid maps was proposed based on the mean nearest-neighbour distance error. We adopt this metric in the present work. The error is calculated as follows. For each occupied cell in the ground-truth map: (1) find the nearest occupied cell in the map under test, (2) calculate the distance between these cells, (3) average this value over all occupied cells in the ground-truth map. Due to the arbitrariness of the coordinate origin in SLAM, an alignment step is first performed by minimising the error metric with an unknown translation and rotation of the entire map.

3 Experimental Overview

3.1 Laboratory Layout

The datasets are collected in a laboratory with well-characterised quasi-2D features. Its layout is shown in Figure 1(a); and photographs of the room are provided in Figure 1(b) and (c).

The laboratory has garage doors on opposite sides (labelled (1) and (2) in Figure 1), as well as three entrance doors ((3), (4), and (5)). Refrigerators (labelled (6) and (7)) are placed in front of a fire hose reel and an eye-wash station to create quasi-2D features that can easily be distinguished by the sensors.

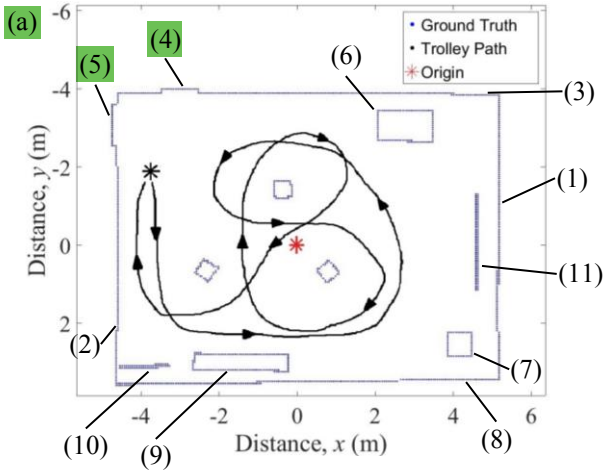

Figure 1: Laboratory with quasi-2D features. (a) Floorplan, (b) view along the *x* axis, (c) view along the *y* axis.

A net (labelled (8) in Figure 1) hangs on one side of the room. Bookshelves and a moveable partition wall ((9) and (10), respectively) are placed in front of a portion of the net; and in front of one of the garage doors, a transparent partition (labelled (11)) is placed to generate additional complexity in the feature set. In the middle of the room, tall obstacles are placed to create more features.

3.2 Sensors

The array of sensors is mounted on a trolley that is wheeled manually around the room. The movement introduces some run-to-run variation into the path;

Figure 2: Trolley carrying sensors

however, odometry data is captured with the motion-tracking system during each circuit with the trolley. The array includes a variety of sensors, including infra-red (IR) ToF cameras, stereo-vision-based depth cameras, and laser ranging sensors. A summary of the ToF sensors, the focus of the current work, can be seen in Table 1.

As shown in previous research [Chan and Lichti, 2015; Chiabrando *et al*., 2010; Pascoal *et al*., 2008], nearly all of the sensors require a warm up to reduce drift due to thermal instability. To avoid introducing this error, all sensors are allowed to warm up for a minimum of 90 minutes before any data is collected.

3.3 Motion-Capture System

A twenty-camera motion-capture system (manufactured by Optitrack) is used to track the position and attitude of the sensors. Typically, the system triangulates the positions of retro-reflective markers (balls) placed on an object. Illuminators on the cameras provide the source of IR light. However, the wavelength emitted by the Optitrack illuminators (850 nm) is found to interfere with many of the sensors, because they operate at similar wavelengths; so the illuminators have been switched off. Instead of tracking retro-reflectors, five high-brightness IR LEDs are attached to the sensor array and tracked by the motion-capture system. This eliminates the interference, and improves tracking performance, particularly at the extremities of the test volume (*e.g*., the corners of the room).

3.4 Room Characterisation

The ground truth is obtained by use of the Optitrack system. To determine the exact location of each wall and object, a special wand is used. This wand has several retro-reflective balls attached, and a sharp point at one end. The location of the point relative to the balls is known, hence coordinates of the various key points on the obstacles can be found by placing the point upon them and measuring the locations of the balls. This is done because placing a retro-reflective ball at each interrogation point is not always feasible. Obstacles are characterised by measuring the positions of each of their corners and assuming planar faces. Walls are characterised by measuring at least five points and then fitting a plane using a least-squares technique. Based on residuals of the plane-fitting process and measurements of obstacle surfaces, the error in the ground truth is approximately ±1 mm.

3.5 Data-Collection Procedure

Data is collected as the sensor trolley moves along two types of path: forward-looking and forward-looking with an occasional 360° pan. Both types of path start and end at the same point; and the basic path is shown in Figure 1(a). A minimum of five laps around each path is performed with each sensor. The addition of the 360° pans is aimed at increasing the amount of data available for the SLAM algorithms. Most ToF cameras have a fairly narrow field of view (FoV) in contrast with laser scanners, which can be a hindrance to scan matching within a SLAM algorithm. Hence it is reasonable to assume that when using this type of sensor the robot may have to periodically "look around the room" when exploring a new area.

3.6 SLAM Algorithms

Data from each of the ToF imaging sensors is individually input to three SLAM algorithms available in ROS: GMapping, KartoSLAM, and Hector Mapping. Because they are all 2D algorithms, a single horizontal slice of the 3D sensor data (*i.e*., a single row of pixels) is used. The GMapping and KartoSLAM packages have a range of options that can be used to tune their behaviour. The majority of the settings are left as defaults, with the exception being the odometry variance levels. These are set to zero in the case where the "nominally perfect" odometry data (directly from the motion-capture system) is used, in which case the algorithm is effectively performing mapping only, without localisation. We then introduce white noise into the odometry-derived poses (Gaussian, with a standard deviation of 0.1 m in translation and 1° in rotation). In this case, the variance settings of the SLAM algorithms are tuned to provide the best map. It is important to note that the

Table 1: Summary of the ToF sensors

Sensor	Range (m)	Resolution (pixels)	Field of view	Manufacturer/designer
SwissRanger 4000	8	176 × 144	43.6° × 34.6°	Mesa Imaging
Single photon avalanche diode (SPAD) sensor	12	64 × 32	40° × 20°	Politecnico di Milano

noise is added to the odometry-derived poses, as opposed to the odometry itself (*i.e.*, velocities and angular rates). This is not completely representative of a real-life scenario in which the odometry-derived pose drifts in a random-walk manner. However, for these initial comparisons, it provides a simple way to introduce a known level of error into the odometry data.

4 Results

Figures 3–5 show the results of GMapping, Hector Mapping, and KartoSLAM computations with the ToF imaging sensor data overlaid on the ground-truth map. The maps have a resolution of 0.05 m/pixel with an image size of 896 × 896 pixels. The maps have the same resolution and image size as the ground-truth map to ensure that the nearest-neighbour map comparison performs optimally.

4.1 Quantitative Map Comparisons

Tables 2 and 3 present the average and standard deviations of the score obtained when a map generated by use of a SLAM algorithm is compared with the ground truth

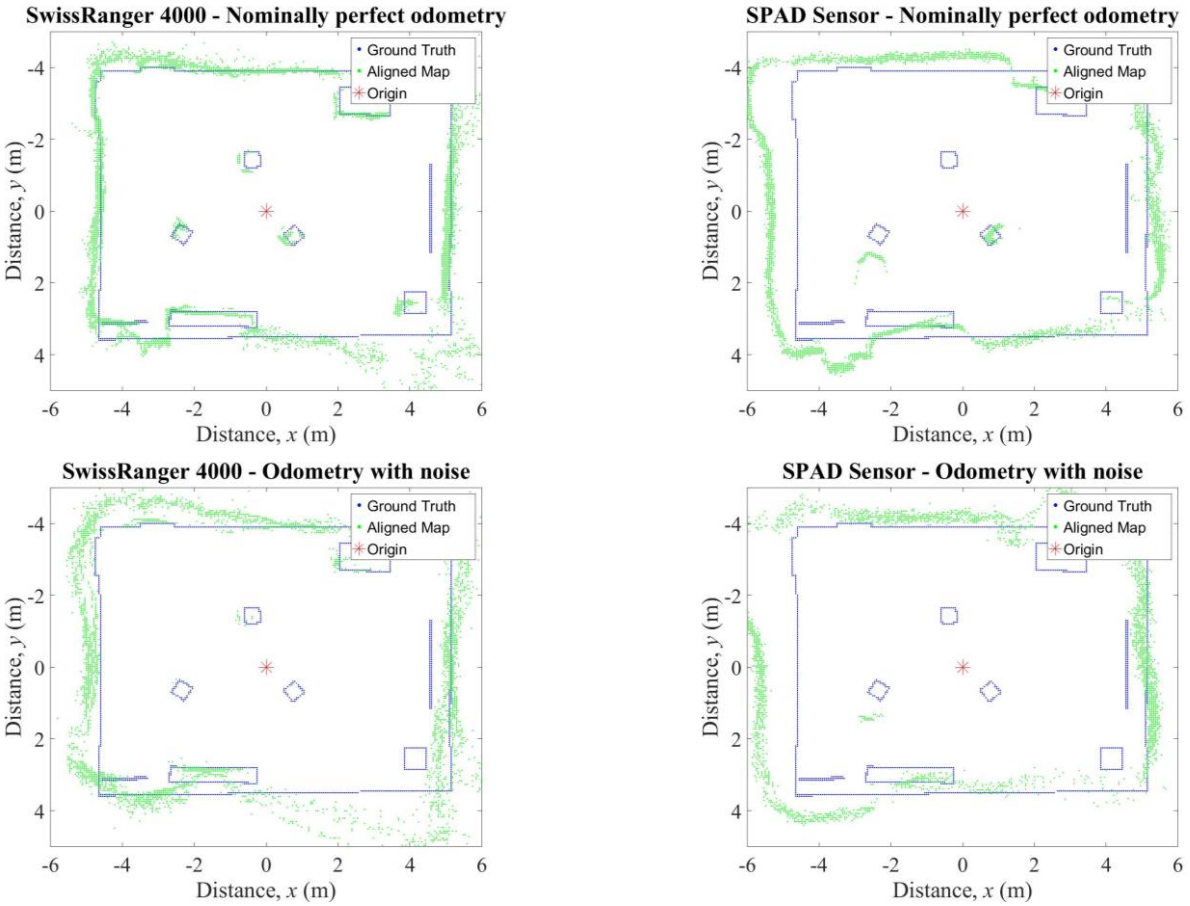

Figure 3: Maps generated using the GMapping algorithm with SwissRanger 4000 and SPAD sensor data, with nominally perfect odometry and with noisy odometry

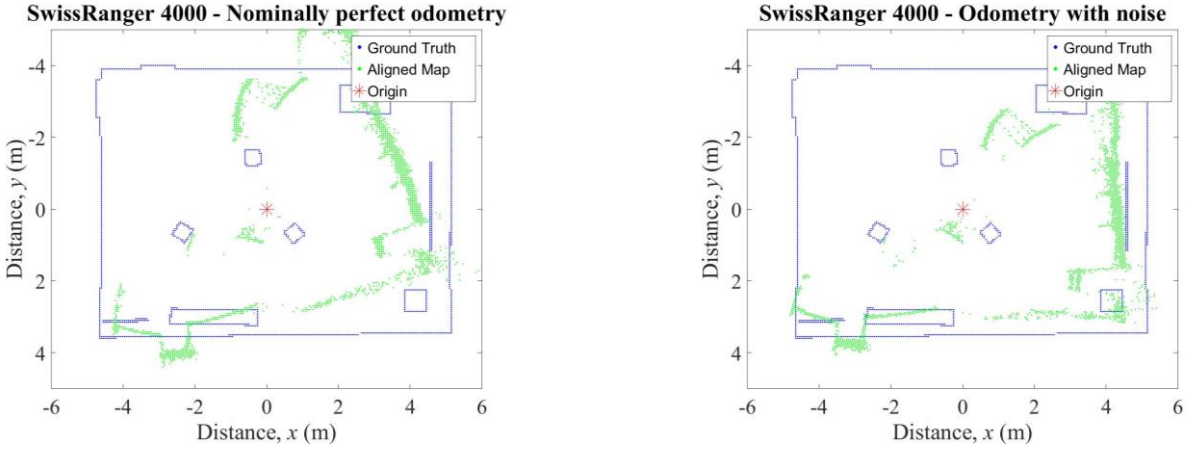

Figure 4: Maps generated using the Hector Mapping algorithm with SwissRanger 4000 data, with nominally perfect odometry and with noisy odometry

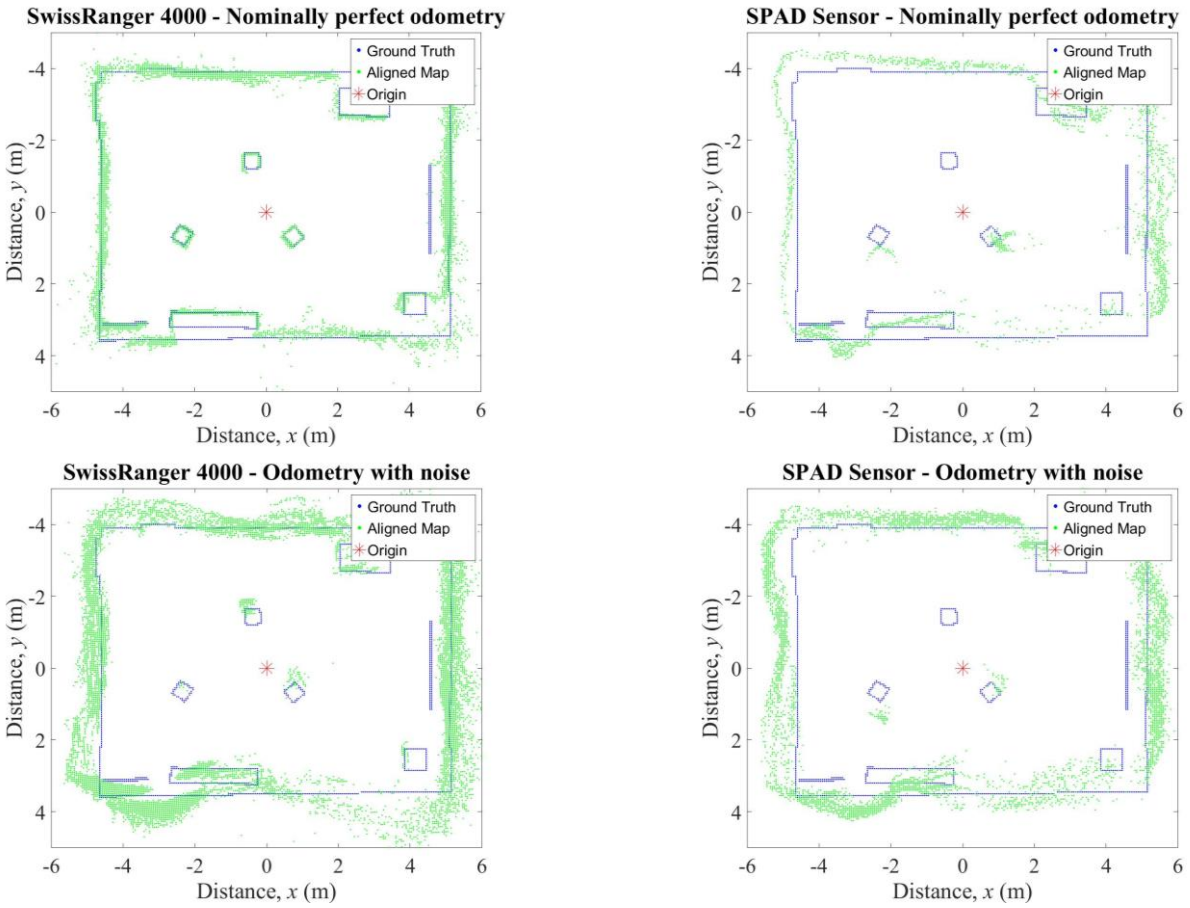

Figure 5: Maps generated using the KartoSLAM algorithm with the SwissRanger 4000 and the SPAD sensor data, with nominally perfect odometry and odometry noise

Table 2: Summary of values obtained when comparing the ground-truth map to maps generated with nominally perfect odometry data

Algorithm	SwissRanger 4000 data			SPAD sensor data		
	Average	Standard deviation	Number of failures	Average	Standard deviation	Number of failures
GMapping	0.127	0.0118	0	0.394	0.0671	1
Hector Mapping	0.759	0.0949	0	n/a	n/a	5
KartoSLAM	0.0889	0.00707	0	0.215	0.0661	1

Table 3: Summary of values obtained when comparing the ground-truth map to maps generated with noise present in the odometry data

Algorithm	SwissRanger 4000 data			SPAD sensor data		
	Average	Standard deviation	Number of failures	Average	Standard deviation	Number of failures
GMapping	0.195	0.0328	0	0.583	0.651	1
Hector Mapping	0.797	0.138	1	n/a	n/a	5
KartoSLAM	0.0796	0.00891	1	0.349	0.173	1

using the metric developed by Santos *et al*. [2013]. In some cases, no map can be obtained using a particular SLAM algorithm and sensor combination; and these are indicated by entries of "n/a" in the tables.

5 Discussion

In Table 2, the average score of the maps generated with GMapping are listed. As can be seen, data from the SwissRanger 4000 sensor produces maps that are more accurate and detailed than those obtained with the SPAD data. When noisy odometry data is used with GMapping, the SwissRanger and the SPAD sensor both are able to generate viable maps. Although the GMapping parameters are altered to compensate for the noisy odometry, the maps are of significantly lower quality than those with nominally perfect odometry. The lack of accuracy in the finer features is noticeable in these cases. This is also reflected when the data in Table 2 is compared with that in Table 3. The overall average scores for the maps generated by use of the SwissRanger and SPAD data with nominally perfect odometry are 0.127 and 0.394, respectively; while the overall average scores for the map with noisy odometry data are 0.195 and 0.583, for the respective sensors. It remains unknown how much odometry error GMapping can withstand before it fails to generate a map.

In Tables 2 and 3, the results from the maps generated by use of Hector Mapping are summarised. As indicated in the tables, Hector Mapping fails to produce a map with SPAD data, as well as with most of the SwissRanger data. When the SwissRanger data are analysed with Hector Mapping, with and without odometry noise, the algorithm is able to generate a map for half the room and then it fails. A possible reason for this failure to generate a full map is due to the algorithm's reliance on scan matching and utilisation of little to no odometry data. As the maps generated by Hector Mapping are found to be unusable, it was deemed that the performance of Hector Mapping with ToF imaging data is inferior to that of GMapping and KartoSLAM.

When the SwissRanger and SPAD data are analysed with KartoSLAM, the maps generated are similar to those from GMapping. As shown in Tables 2 and 3, the overall scores obtained with the KartoSLAM maps with nominally perfect odometry are considerably lower than the scores for GMapping maps with the SwissRanger and SPAD data, indicating that higher-quality maps are obtained with KartoSLAM. It should be noted that KartoSLAM does not exhibit the same consistency of map generation as GMapping does when noisy odometry is utilised. KartoSLAM is only able to generate four out of five maps that are relatively accurate when odometry error is included. When the maps with odometry error are analysed, it is evident that, like GMapping, KartoSLAM could not detect many features in the room. One possible reason why KartoSLAM generates slightly different results to GMapping is due to KartoSLAM's employing a graph-based SLAM approach, while GMapping uses a Rao-Blackwellized particle filter [Santos, *et al*., 2013].

One source of error that is apparent in the generated maps is multi-path effects. Multi-path effects are acknowledged by the manufacturer of the SwissRanger as being a source of error that could prevent corners from being correctly visualised. Because the SwissRanger and SPAD sensors are both based on an indirect ToF principle, they are both susceptible to multi-path effects. The maps generated from the SPAD sensor show that the SPAD sensor is more susceptible to multi-path effects than the SwissRanger. This is evidenced by many of the corners of the SPAD-based maps being more rounded than the corners on the SwissRanger-based maps. Again, the SwissRanger may be less susceptible to multi-path effects because it has been very well developed over several iterations, whereas the SPAD sensor is a relatively immature prototype.

A likely reason why some SLAM algorithms do not perform as well as the simulation results obtained by Santos *et al*. [2013] is due to the FoV of the sensors. The simulation results obtained by Santos *et al*. [2013] used a simulated Hokuyo URG-04LX-UG01 with a 240° FoV. The Hokuyo was chosen because the ROS-based SLAM algorithms have been designed to be utilised with wide-FoV sensors. This is evident from the results obtained with the Hector Mapping algorithm. Hector Mapping has been designed to be utilised with 2D sensors that possess a wide FoV. When a narrow-FoV sensor is used, it is unable to capture sufficient features in a single frame. When Hector Mapping attempts to match these features with features of the next frame, there are too few features to make an accurate match. With the data used here, Hector Mapping often mistakenly matches an obstacle in the middle of the room with a feature that is next to the wall. This means that many frames are unsuccessfully matched to each other, and the map generated is an inaccurate representation of the room. This is one possible reason why Hector Mapping does not perform as well here as in the simulation results obtained by Santos *et al*. [2013]. With the other SLAM algorithms, scan matching can be turned off, enabling the algorithms to place more reliance on the odometry data and therefore to generate better maps.

6 Future Work

Based on these results, it is clear that the performance of several standard SLAM algorithms using ToF camera data is not as good as that obtained with more traditional sensors. Future work will include further analysis of the various ROS SLAM algorithms and how they can be modified to better suit ToF camera data. The consistently inferior maps generated by use of the SPAD data, while possibly only a consequence of the lower resolution, indicate that further assessment of the accuracy of the SPAD sensor itself may be needed.

Further analysis of the data from other sensor types (laser scanners and vision-based systems) will also be performed. We hope to make this large dataset public to aid researchers in this area with the testing, refinement, and validation of SLAM algorithms.

7 Conclusion

In this paper, we detailed an extensive data collection experiment aimed at evaluating SLAM algorithms with various types of sensor data. Ten sensors were used in all, with this paper focussed on the analysis of the performance of ToF camera sensors with existing SLAM algorithms. The data from two of these sensors (a SwissRanger and a SPAD camera) is input to three SLAM algorithms available in ROS (GMapping, KartoSLAM, and Hector Mapping). This is done firstly

with the accurate poses derived from a motion-tracking system, then with noise introduced into the pose measurements. The latter is designed to test the algorithms' ability to perform localisation, while the former is mainly a test of its ability to perform mapping. The results show that the KartoSLAM and GMapping SLAM algorithms have varying capability to handle noise that may be present in odometry. When presented with nominally perfect odometry, GMapping and KartoSLAM are able to generate relatively accurate maps. The Hector Mapping SLAM algorithm is shown to be of limited use. It is unable to generate maps with nominally perfect odometry or with noisy odometry.

Even with a narrow-FoV sensor, GMapping and KartoSLAM are able to generate maps, unlike the Hector Mapping algorithm. Hector Mapping may fail to generate any maps because it has been designed to be utilised with wide-FoV sensors. The data captured with the other sensors utilised for data collection are comparatively more suitable for analysis with the KartoSLAM and GMapping algorithms.

Acknowledgements

This research is supported by DST's Strategic Research Initiative on Trusted Autonomous Systems (Program Tyche). The authors wish to acknowledge Mr Matt Clemente for his assistance in the conduct of the experiments and the provision of a SPAD camera by Professor Franco Zappa of the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano. The authors also thank Dr Ki Ng of DST's Land Division for the loan of two of the other instruments used in the experiment.

References

[Abdallah, *et al*., 2006] Samer M. Abdallah, Daniel C. Asmar and John S. Zelek. Towards benchmarks for vision SLAM algorithms. *Proc*. *Int'nal Conf*. *Robot*. *Autom*., Orlando, FL, USA, 15–19 May 2006. IEEE.

[Balaguer, *et al*., 2007] Benjamin Balaguer, Stefano Carpin and Stephen Balakirsky. Towards quantitative comparisons of robot algorithms: Experiences with SLAM in simulation and real world systems. *Worksh*. *Perf*. *Eval*. *Benchmark*., *21st Int'nal Conf*. *Intel*. *Robots Sys*., San Diego, CA, USA, 29 October – 2 November 2007. IEEE.

[Chan and Lichti, 2015] Ting On Chan and Derek D. Lichti. Automatic *in situ* calibration of a spinning beam LiDAR system in static and kinematic modes. *Rem*. *Sens*., 7(8): 10480–10500, August 2015.

[Chiabrando, *et al*., 2010] Filiberto Chiabrando, Dario Piatti and Fulvio Rinaudo. SR-4000 ToF camera: Further experimental tests and first applications to metric surveys. *Int'nal Arch*. *Photogram*., *Rem*. *Sens*. *Spat*. *Inf*. *Sci*., *Comm*. *V Symp*., Vol. XXXVIII(Part 5), pages 149–154, Newcastle upon Tyne, UK 2010.

[Handa, *et al*., 2014] Ankur Handa, Thomas Whelan, John McDonald and Andrew J. Davison. A benchmark for RGB-D visual odometry, 3D reconstruction and SLAM. *Proc*. *Int'nal Conf*. *Robot*. *Autom*., Hong Kong, China, 31 May – 7 June 2014. IEEE

[Nguyen, *et al*., 2007] Viet Nguyen, Stefan Gächter, Agostino Martinelli, Nicola Tomatis and Roland Siegwart. A comparison of line extraction algorithms using 2D range data for indoor mobile robotics. *Auton*. *Robots*, 23(2): 97–111, August 2007.

[Pascoal, *et al*., 2008] José Pascoal, Lino Marques and Anibal T. de Almeida. Assessment of laser range finders in risky environments. *Proc*. *Int'nal Conf*. *Intel*. *Rob*. *Sys*., pages 3533–3538, Nice, France, 22–26 September 2008. IEEE.

[Santos, *et al*., 2013] João Machado Santos, David Portugal and Rui P. Rocha. An evaluation of 2D SLAM techniques available in Robot Operating System. *Proc*. *Int'nal Symp*. *Safety Sec*. *Res*. *Robot*., Linköping, Sweden, 21–26 October 2013. IEEE.

[Sturm, *et al*., 2011] Jürgen Sturm, Stéphane Magnenat, Nikolas Engelhard, François Pomerleau, Francis Colas, Daniel Cremers, Roland Siegwart and Wolfram Burgard. Towards a benchmark for RGB-D SLAM evaluation. *RGB-D Worksh*. *Adv*. *Reason*. *Depth Cam*. *Robot*.: *Sci*. *Sys*. *VII*, Los Angeles, CA, USA, 27–30 June 2011.

[Zappa, *et al*., 2007] Franco Zappa, Simone Tisa, Alberto Tosi and Sergio Cova. Zappa Principles and features of single-photon avalanche diode arrays. *Sensors and Actuators A: Physical*, 140(1): 103–112, 1 October 2007.