2013 IEEE/RSJ International Conference on
Intelligent Robots and Systems (IROS)
November 3-7, 2013. Tokyo, Japan

Odometry-driven Inference to Link Multiple Exemplars of a Location

Stephanie Lowry, Gordon. Wyeth, *Member, IEEE,* and Michael Milford, *Member, IEEE*

***Abstract*— A major challenge for robot localization and mapping systems is maintaining reliable operation in a changing environment. Vision-based systems in particular are susceptible to changes in illumination and weather, and the same location at another time of day may appear radically different to a system using a feature-based visual localization system. One approach for mapping changing environments is to create and maintain maps that contain multiple representations of each physical location in a topological framework or manifold. However, this requires the system to be able to correctly link two or more appearance representations to the same spatial location, even though the representations may appear quite dissimilar. This paper proposes a method of linking visual representations from the same location without requiring a visual match, thereby allowing vision-based localization systems to create multiple appearance representations of physical locations. The most likely position on the robot path is determined using particle filter methods based on dead reckoning data and recent visual loop closures. In order to avoid erroneous loop closures, the odometry-based inferences are only accepted when the inferred path's end point is confirmed as correct by the visual matching system. Algorithm performance is demonstrated using an indoor robot dataset and a large outdoor camera dataset.**

I. Introduction

A key issue facing vision-based Simultaneous Localization And Mapping (SLAM) systems operating in real-world environments is the management of appearance change. Any localization and mapping system used for navigation will have to update stored location appearance data in order to remain relevant and continue providing both high precision and sufficient recall over extended periods of time. Furthermore, since many environments exhibit cyclic change (for example, between day and night; or winter and summer) a map could usefully contain more than a single representation of each location.

A number of vision-based localization methods have been proposed that allow multiple visual representations of the same location. These systems assume that change is gradual and thus localization is always possible [1] or simply do not attempt to link dissimilar appearances together spatially or topologically, instead performing localization in appearance space [2]. The question remains of how to match dissimilar environmental configuration in the face of imperfect localization.

In this paper we demonstrate a method of linking dissimilar appearance representations together, by using available visual loop closure information (when available) in combination with knowledge of robot odometry. Our system uses a two-step process: a prediction step where a probabilistic particle filter method is applied to robot odometry to infer possible location matches, and a validation step in which the inferred links are confirmed by the visual localization system when it is able to successfully localize. This allows the system to create complex representations of a location over time and has the potential to improve graph connectivity (see Fig. 1) and enable more effective map pruning.

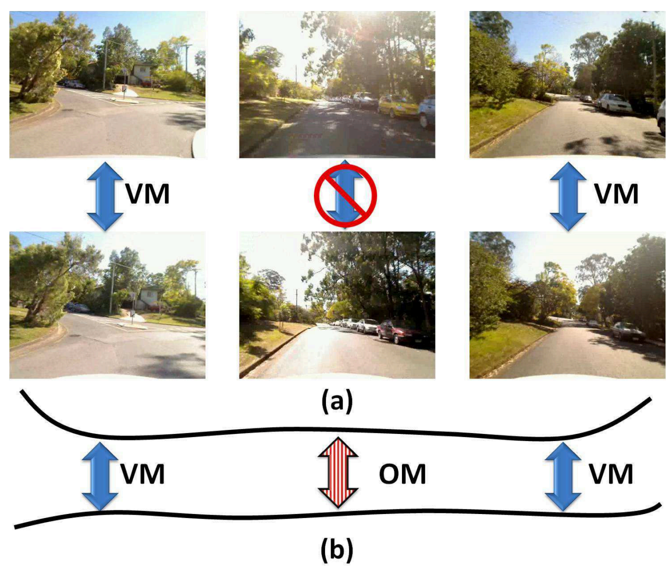

Figure 1. (a) Example path segment where visual matching succeeds at either end but fails due to appearance change in between. (b) Our system exploits the visual matches and dead reckoning information to link the intervening locations despite the visual changes.

The performance of the method is demonstrated in two studies. The first study uses an indoor dataset where parts of the environment experience visual change and demonstrates the system both correctly linking locations based on dead reckoning even when there is no visual similarity and not erroneously linking locations when a new path is traversed. The second study uses a larger outdoor dataset where over a 14km run our system increases the number of linked locations by an order of magnitude whilst the average linking error only increases by 15%.

S. Lowry, G. Wyeth and M. Milford are with the School of Electrical Engineering and Computer Science, Queensland University of Technology, Brisbane, Australia. stephanie.lowry@student.qut.edu.au. This work was in part funded by the Australian Research Council through Discovery project DP1113006. S. Lowry is supported by an Australian Postgraduate Award and a QUT Vice-Chancellor's Scholarship.

The paper proceeds as follows. In Section 2, we provide an overview of current approaches to managing map information in changing environments and performing localization under drastic appearance change. Section 3 describes the odometry-driven inference method. Section 4 describes the experimental setup used to test the algorithm, and Section 5 presents results showing odometry-driven inference on a large outdoor dataset. Section 6 discusses these results and directions for future work.

II. Background

Management of changing environments is a crucial part of any real-world persistent navigation system. One method of managing changing environments used by laser-based mapping systems is to create multiple maps of the environment and use a selection process to choose the best one at a given time. In [3], an environment is represented by multiple maps that are updated on different timescales, allowing transient changes to be filtered out whilst static elements remain. The optimal number of configurations for each local submap [4] can be determined using fuzzy k-means clustering. If a map of the static part of the environment is available, semi-static maps [5] can be used for localization. These temporary maps are deleted when they no longer match the current observation sufficiently well.

A number of vision-based SLAM systems provide map updating capabilities to manage changing environments. Map updates can be achieved via node pruning in a topological grid [6] or using biologically inspired models of memory to decide what information gets remembered and what gets forgotten [1]. The map update assumes the robot is correctly localized and thus any new sensor information is due to environmental change rather than pose error. Another vision-based approach avoids the localization problem entirely by not attempting to match dissimilar appearances of the same location [2]; instead the system creates and localizes within plastic maps of visual experiences. Similarly to [7] this can be considered a form of localization and mapping in appearance space.

When an environment has changed so much that existing image matching techniques fail, it is possible to use topological or sequential data to infer information about locations regardless of appearance. Sequences of images have been used to match locations despite drastic changes in lighting and weather [8]. Topological information can also be used to correct data association errors caused by faulty loop closures [9, 10].

The system presented in this paper follows the final approach of using topological information to infer information about locations. This system (denoted CAT-Graph+) matches dead reckoning along a robot path and uses occasional visual loop closures to determine whether to accept or reject the odometry-based inferences thereby providing a mechanism to generate multiple representations of locations within the environment. CAT-Graph+ does not perform SLAM, but is intended to work alongside an existing SLAM system, enhancing its capability to build persistent, coherent maps.

III. CAT-Graph+

In this section we introduce the method for linking visual representations of location via odometry-driven inferences and demonstrate its use with the existing topological SLAM system CAT-Graph [11, 12]. Although CAT-Graph+ was developed to work closely with CAT-Graph and many of its key elements are similar, it is capable of being decoupled and used with other SLAM systems: the only requirements are access to odometry and to the reported loop closures of the underlying SLAM system. The system can either be run as a parallel process to the underlying SLAM or during an offline update phase.

CAT-Graph+ is designed to operate in environments where successful visual loop closures are rare due to significant appearance change to the environment. Using an odometry-only particle update scheme, CAT-Graph+ can propagate information from an assured visual loop closure along a known path. The knowledge of current location combined with a dead reckoning sensor allows the system to continue to predict its location along routes that are visually unrecognizable due to environmental change. If these predictions are validated by a later visual loop closure, the system map is then updated with the inferred links.

A. Initialization

CAT-Graph+ uses a particle filter to predict the current robot location when visual information is unavailable. The particles are initialized whenever a visual loop closure is reported by the underlying SLAM system (see Fig. 2(a)). These particles represent the localization probability and are distributed according to the reported loop closure position with added Gaussian noise. Each particle p_i is assigned to a particular spatial location x_i and has an associated weight w_i.

B. Particle Propagation and Weight Updates

The position x_i^k of each particle p_i at time step k is sampled from the motion model distribution with probability:

$$P(x_i^k \mid x_i^{k-1}, u^k) \tag{1}$$

where x_i^{k-1} is the particle position at time step $k-1$, and u^k is the odometry input (see Fig. 2(b)). The particle weights are updated according to the likelihood $P(x_i^k|x_i^{k-1},u^k)$:

$$w_i^k = w_i^{k-1} P(x_i^k \mid x_i^{k-1}, u^k) \tag{2}$$

C. Odometry Hypothesis and Location Prediction

At each time step, the odometry particle filter performs a hypothesis calculation that represents the best guess of the inference system as to the current location. Each particle p_i is assigned a location hypothesis value determined by the sum of all the particle weights within a user-defined radius d^{odo}:

$$P^k(p_i) = \sum_{|p_j - p_i| < d^{odo}} w_j^k \tag{3}$$

The particle p_i with the largest P^k is considered to be the most likely match, and represents the "best guess" of the location at the current time:

$$P_k = \max_i P^k(p_i) \tag{4}$$

The maximum hypothesis value P^k is tested to see if it exceeds a user-defined threshold T^{odo}, that is, to see if:

$$P_k > T^{odo} \tag{5}$$

If Equation 5 holds we have a potential candidate for an odometry-inferred match (see Fig. 2(c)). However, this candidate will not be confirmed as a true match until the inferred odometry path has been validated by the visual localization system.

D. Path Validation

Since unrestricted odometry propagation will inevitably lead to inaccurate linking, the odometry particle filter can only form links when the odometry hypotheses are proven to be valid by the visual system. The validation step happens at each visual loop closure, and all the odometry hypotheses since the previous loop closure are assessed for inclusion in the map (see Fig. 2(d)).

The path validation process is as follows. Suppose LC is the spatial location of the visual loop closure. The hypothesis calculation from Section C is repeated, but comparing the odometry particles against the successful loop closure:

$$P^{path} = \sum_{|p_i - LC| < d^{path}} w_i \tag{6}$$

Once again we test to see if P^{path} is greater than a certain user-defined threshold value:

$$P^{path} > T^{path} \tag{7}$$

If Equation 7 holds, then we consider that CAT-Graph+ has correctly predicted the robot's path. The inferences that successfully passed the validation step described in Section C above are accepted as true map links. If P^{path} does not exceed T^{path}, all the odometry inferences from the path section are discarded. Regardless of the outcome of the path validation step, the odometry filter will also be re-initialized as described in Section A above.

An assumption implicit in CAT-Graph+ is that at least some loop closures can be achieved in parts of the environment. These loop closures are then used to validate the intervening odometry-driven inferences. This is a weaker requirement than in other systems where confirmed loop closures are generally required at a specific location before the visual representations at that location will be linked.

CAT-Graph+ is not susceptible to false positives generated by incorrect visual loop closures, as the reported visual loop closure needs to match the location predicted by the odometry. If it does not, no odometry linking will occur.

It is not sufficient to assume that two path segments that start and end at the same place traverse the same locations in between (see Fig. 6(a) for a simple counter-example). However, CAT-Graph+ also requires that Equation 5 holds at each time step. This requirement is to ensure that the robot actually is traversing the same path.

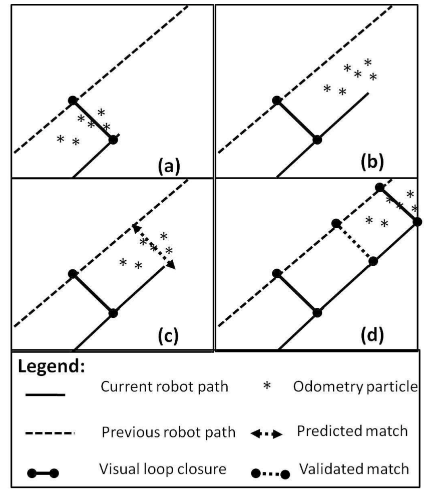

Figure 2. CAT-Graph+ schematic. (a) *Initialization*: when a visual loop closure occurs, particles are selected from the location distribution. (b) *Propagation:* particle positions are updated according to the odometry input. (c) *Prediction:* the current location is predicted from the particle distribution. (d) *Validation*: when a loop closure occurs, if the prediction location matches the actual (visually confirmed) location, the predicted path between loop closures is validated. The particles are then re-initialized (as in (a)) and the process repeats.

IV. EXPERIMENTAL SETUP

In this section we describe the two datasets used, ground truth measures, image pre-processing and parameter values. In these experiments CAT-Graph provided the underlying visual loop closure system, and both datasets had been tested using CAT-Graph or its predecessor CAT-SLAM [13]. Baseline operation capabilities under visually stable environmental conditions were known: CAT-Graph/CAT-SLAM achieved recall of over 70% at 100% precision on the indoor environment (Dataset 1)[12] and over 20% at 100% precision on the outdoor environment (Dataset 2) [13] . These results are considered "state of the art" in visual SLAM systems and so form an excellent baseline for comparison with CAT-Graph+ results.

A. Dataset 1: Pioneer dataset

The first dataset was an indoor dataset in which the environment was changed in a controlled manner. The experiment was designed to answer three specific questions, namely:

- Does CAT-Graph+ correctly link locations via odometry when visual matching fails?
- Does CAT-Graph+ erroneously link locations on different paths even when the paths start and end at the same place?
- What is the effect of parameter variation?

1) Testing Environment

The selected testing area was an indoor office environment (see Fig. 4(b)). A sample test path is shown on the floor plan in Fig. 3. On the first traversal, the environment surrounding the robot path was kept clear, whilst on the third traversal part of the environment (shaded blue in Fig. 3) was manually changed by adding and removing large amounts of furniture (see Fig. 4(c)). The second traversal started and ended at the same points as the first and third traversals, but took a different path in between.

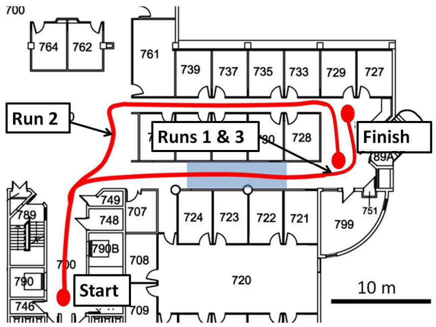

Figure 3. Sample test paths. Runs 1 & 3 follow the same path but under different visual conditions. Run 2 follows a different path with identical start and end points.

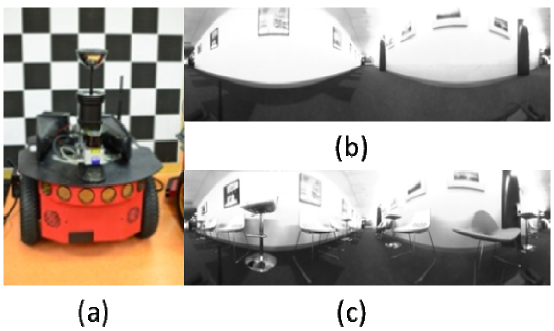

Figure 4. The Pioneer 3DX data capture platform (a), panoramic view of a section of the test environment (b), test environment with changed furniture layout (c).

2) Robot Platform and Image Processing

The experimental data was collected using an Adept Pioneer 3DX robot (see Fig. 4(a)) driven by remote control. The Pioneer's wheel odometry was logged at 10 Hz. Images were captured at 7.5fps using a Basler A310fc camera and a panoramic mirror, and were unwrapped to 960 × 240 panoramic view. This camera system was chosen to be identical to that used in an earlier experiment [12] in order to maintain equivalent baseline operation, but to ensure sufficient appearance change occurred, the image was restricted to a forward-facing patch of 261 × 101 pixels, representing a field of view comparable to many perspective cameras of approximately 100° × 45°. This image patch was tested to ensure that operation in the static parts of the environment was not significantly affected whilst localization in the altered parts of the environment was sufficiently difficult. Features were extracted from the image patch using SURF [14].

3) Ground Truth

Metric ground truth was provided using data from a Hokuyo URG04LX laser scanner logged at 10 Hz. Offline, this data was processed using AMCL [15] via the ROS framework [16]. A previously generated occupancy grid map was used. This occupancy grid map was generated using a SICK LMS-291 laser scanner, and processed using the Gmapping algorithm [17] to 5cm resolution.

B. Dataset 2: St Lucia dataset

A larger, outdoor dataset was used to evaluate the applicability of CAT-Graph+ to varied types of environments.

1) Testing Environment

The outdoor dataset was first presented in [18], and provides visual and GPS data from a car driven around suburban streets in St Lucia, Queensland over a 3 week period. Due to the varying times of day the loop was traversed, there is significant appearance change in certain parts, making it an ideal test dataset for CAT-Graph+. This paper used a section of this dataset that consisted of four visually different but spatially consistent loops around an approximately 3.6km route (see Fig. 5).

Figure 5. Outdoor test environment consisting of a 3.6km loop around a suburban road network. (Imagery ©2012 Cnes/Spot Image, DigitalGlobe, GeoEye, Sinclair Knight Merz & Fugro).

2) Image Processing

The camera data was captured from a forward-facing Logitech QuickCam Pro 9000 web camera, at 640 × 480 pixel resolution and at an average of 15 fps (see Fig. 1 for sample images). Features were extracted from the image patch using SURF [14].

3) Ground Truth

GPS was logged at 1Hz and used for ground truth. A simulated odometry input was generated using a linear interpolation of GPS data as in [18].

C. System Parameters

The key visual algorithm parameters and values for CAT-Graph were selected to mirror published results [12] and [13] on the same environments. An exception was made in the outdoor dataset, where due to the challenging nature of the dataset the visual hypothesis threshold T_h was lowered to 0.6 to increase the number of loop closure candidates. Table I lists the additional parameters required for odometry inference – namely the thresholds and distance values for odometry validation. These four new parameters (two threshold values and two distance values) have analogues T_h and d_h in CAT-Graph. Aside from the parameter d^{path}, which was chosen to be 10.0m in both cases, the other parameters were simply matched to their visual analogue.

TABLE I. CAT-GRAPH+ PARAMETER LIST

Parameter	**Indoor dataset**	**Outdoor dataset**
T^{odo}	0.8	0.6
d^{odo}	1.0 m	5.0 m
T^{path}	0.8	0.6
d^{path}	10.0 m	10.0 m

V. RESULTS

In this section we present results for the indoor and outdoor datasets. As CAT-Graph+ is a tool for linking locations based on localization information from an independent source, the focus here is on the connectivity of the graph created instead of conventional measures such as precision-recall curves. For this reason we measure our results in terms of the number of "possible" matches; that is, using the associated ground truth we define all map nodes within d_h of each other as "possible" matches, and then see what proportion of possible links are matched by either the visual SLAM system or the odometry-driven inference system.

A. *Dataset 1: Indoor Pioneer dataset*

1) *Linking of visually changed locations*

The results of running CAT-Graph+ along the same path through a visually changed environment are displayed in Fig. 6 (b) (this illustration omits Run 2 for clarity). Of the 499 possible links in this environment, the vision system matched 44 (or 9% of the total) with a maximum error of 0.35m and CAT-Graph+ matched 370 (75% of the total) with a maximum error of 0.53m.

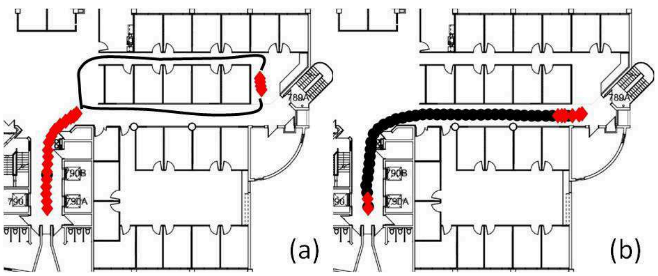

Figure 6. Visual (red diamond) and odometry-driven (black circle) matches imposed on ground truth for the 3 runs. (a) No incorrect odometry links are created when the system takes different paths between two locations. (b) Odometry links form when the environment has changed and visual matching fails.

2) *Different paths, same end points*

The results of running CAT-Graph+ along two different paths that start and end at the same points is displayed in Fig. 6(a) (this illustration omits Run 3 for clarity). It can be seen that CAT-Graph+ correctly identifies that these two paths take different routes between the start and end points and thus does not link them. This scenario demonstrates why a global validation check (using only T^{path} and d^{path}) is not sufficient and a probabilistic path prediction technique is also required.

3) *Parameter Variation*

A parameter variation test was performed for the indoor dataset. Each parameter was varied whilst keeping the other parameters at the baseline values listed in Table I to test the effect of parameter selection on the system. Table II lists the range tested for each parameter.

TABLE II. RANGE OF PARAMETER VALUES

Parameter	Indoor dataset
T^{odo}	0.1 – 0.9
d^{odo}	0.5 m – 20 m
T^{path}	0.1 – 0.9
d^{path}	0.5 m – 50 m

The experiment was run 4 times with each parameter combination to account for the randomness inherent in the particle filter. Over all such experiments, visual matching successfully linked (on average) 3.68% of all possible matches with an average error of 0.2 m (max error of 0.96m). The results across all the parameter combinations account for (on average) 28.14% with an average error of 0.21m (max error of 0.82m). The results showed a clear increase in map connectedness using odometry and visual matches over using odometry matches alone, regardless of the parameter combination. Obviously, good odometry matching would be impossible without good visual matching to provide initialization and validation checks; the usefulness of the odometry-only inference is to take good visual matching and apply it more broadly across an environment.

B. *Dataset 2: St Lucia dataset*

In this section we present results from the outdoor St Lucia dataset to demonstrate the system behavior in a larger, outdoor environment. Fig. 7(a) displays the number of visual loop closures and odometry links as a percentage of the total possible number of graph links. The odometry links substantially increase the connectivity of the resulting graph.

Of particular note is the effectiveness of CAT-Graph+ over repeated traverses of the same loop, as the proportion of odometry-driven loop closures does not decrease with the number of traverses.

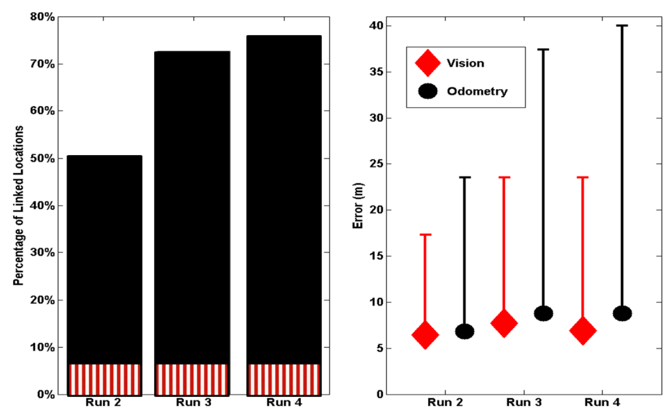

Figure 7. St Lucia experiment: (a) percentage of all possible links matched by visual (red stripes) and odometry (black solid) matching; (b) error (in m) of visual (red diamond) and odometry (black circle) matches

Fig. 7(b) displays the error distribution for both visual links and odometry links for the St Lucia datasets (average and maximum errors are displayed). In each case, the average error in the odometry matches is similar to the visual loop closure error, but due to the dependence of odometry links on visual loop closures, the maximum odometry match error consistently exceeds the maximum visual loop closure error.

Fig. 8 displays an example of a location correctly linked by CAT-Graph+ that was not matched using the visual loop closer. The CAT-Graph+ links provide the ability to perform later processing such as increasing graph connectivity or building up representations of each physical location using multiple image representations. Furthermore, later processing on the images could be performed as an additional confirmation of correctness.

Figure 8. Sample of location matching by CAT-Graph+. Changed lighting conditions and parked cars parked make visual recognition challenging.

VI. Discussion and Future Work

The work presented here provides a method for generating links between representations of a given physical location, even when the appearance of the location has changed significantly. This method complements existing SLAM systems as it makes different assumptions and offers a different capability, namely the ability to link dissimilar appearance representations. One application is the ability to develop more complex models of how location appearances change over time, using a method such as [1]. Another use for an odometry-driven inference model is to maintain the performance of the underlying vision recognition system by providing feedback about the new appearance of the world. The updated information can be integrated into the visual system's maps to minimize system degradation over time.

The increased graph connectivity that CAT-Graph+ provides also has the potential to improve optimal path planning in a topological mapping system such as CAT-Graph. For example, in the case of the St Lucia dataset a system that only linked the visual loop closures might plan and follow a path several hundred meters in length greater than one created by CAT-Graph+.

Future work with CAT-Graph+ seeks to test two key assumptions. As noted previously, the first assumption is that there are at least some visual loop closures within the environment. In challenging environments this may not be the case. One option to resolve this issue is to use a matching technique that is more robust to variation in the environment and we are currently investigating this possibility. The second key assumption is the dependence on a reliable odometry input. We note that, whilst poor odometry will lead to fewer inferred matches, it will not lead to false positive matches due to the required validation step, which discards the inaccurate predictions. Further investigations are underway to test performance in more challenging real world conditions with noisy odometry.

Acknowledgment

We thank Arren Glover for access to the St Lucia dataset and Will Maddern for the use of the CAT-Graph codebase.

References

[1] F. Dayoub, *et al.*, "Long-term experiments with an adaptive spherical view representation for navigation in changing environments," *Robotics and Autonomous Systems,* vol. 59, pp. 285-295, 2011.

[2] W. Churchill and P. Newman, "Practice makes perfect? Managing and leveraging visual experiences for lifelong navigation," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, 2012, pp. 4525-4532.

[3] P. Biber and T. Duckett, "Experimental Analysis of Sample-Based Maps for Long-Term SLAM," *International Journal of Robotics Research,* vol. 28, pp. 20-33, 2009.

[4] C. Stachniss and W. Burgard, "Mobile Robot Mapping and Localization in Non-Static Environments," in *AAAI'05 - The 20th National Conference on Artificial Intelligence* 2005, pp. 1324-1329.

[5] D. Meyer-Delius, *et al.*, "Temporary Maps for Robust Localisation in Semi-static environments," presented at the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2010.

[6] M. J. Milford and G. F. Wyeth, "Mapping a Suburb With a Single Camera Using a Biologically Inspired SLAM System," *IEEE Transactions on Robotics,* vol. 24, pp. 1038-1053, Oct 2008.

[7] M. Cummins and P. Newman, "Appearance-only SLAM at Large Scale with FAB-MAP 2.0," *International Journal of Robotics Research,* vol. 30, pp. 1100-1123, 2011.

[8] M. J. Milford and G. F. Wyeth, "SeqSLAM: Visual route-based navigation for sunny summer days and stormy winter nights," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, 2012, pp. 1643-1649.

[9] N. Sunderhauf and P. Protzel, "Towards a robust back-end for pose graph SLAM," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, 2012, pp. 1254-1261.

[10] E. Olson and P. Agarwal, "Inference on networks of mixtures for robust robot mapping," presented at the Robotics: Science and Systems, Sydney, Australia, 2012.

[11] W. Maddern, *et al.*, "Towards Persistent Localization and Mapping with a Continuous Appearance-based Topology," presented at the Robotics: Science and Systems, Sydney, Australia, 2012.

[12] W. Maddern, *et al.*, "Towards Persistent Indoor Appearance-based Localisation, Mapping and Navigation using CAT-Graph," presented at the IEEE/RSJ International Conference on Intelligent Robots and Systems, 2012.

[13] W. Maddern, *et al.*, "CAT-SLAM: Probabilistic Localisation and Mapping using a Continuous Appearance-based Trajectory," *International Journal of Robotics Research,* vol. 31, pp. 429-451, 2012.

[14] H. Bay, *et al.*, "Speeded-Up Robust Features (SURF)," *Computer Vision and Image Understanding,* vol. 110, pp. 346-359, Jun 2008.

[15] D. Fox, *et al.*, "Monte Carlo Localization: Efficient position estimation for mobile robots," presented at the Sixteenth National Conference on Artificial Intelligence, 1999.

[16] M. Quigley, *et al.*, "ROS: an open-source Robot Operating System," presented at the Proc. Open-Source Software workshop of the International Conference on Robotics and Automation (ICRA), 2009.

[17] G. Grisetti, *et al.*, "Improved techniques for grid mapping with Rao-Blackwellized particle filters," *IEEE Transactions on Robotics,* vol. 23, pp. 34-46, Feb 2007.

[18] A. J. Glover, *et al.*, "FAB-MAP + RatSLAM: Appearance-based SLAM for multiple times of day," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2010, pp. 3507-3512.