Featureless Visual Processing for SLAM in Changing Outdoor Environments*

Michael Milford, Ashley George

Abstract. Vision-based SLAM is mostly a solved problem providing clear, sharp images can be obtained. However, in outdoor environments a number of factors such as rough terrain, high speeds and hardware limitations can result in these conditions not being met. High speed transit on rough terrain can lead to image blur and under/over exposure, problems that cannot easily be dealt with using low cost hardware. Furthermore, recently there has been a growth in interest in lifelong autonomy for robots, which brings with it the challenge in outdoor environments of dealing with a moving sun and lack of constant artificial lighting. In this paper, we present a lightweight approach to visual localization and visual odometry that addresses the challenges posed by perceptual change and low cost cameras. The approach combines low resolution imagery with the SLAM algorithm, RatSLAM. We test the system using a cheap consumer camera mounted on a small vehicle in a mixed urban and vegetated environment, at times ranging from dawn to dusk and in conditions ranging from sunny weather to rain. We first show that the system is able to provide reliable mapping and recall over the course of the day and incrementally incorporate new visual scenes from different times into an existing map. We then restrict the system to only learning visual scenes at one time of day, and show that the system is still able to localize and map at other times of day. The results demonstrate the viability of the approach in situations where image quality is poor and environmental or hardware factors preclude the use of visual features.

1 Introduction

Visual mapping and navigation on robots has advanced rapidly in the last decade. There are now many vision-based techniques including FAB-MAP [1], MonoSLAM [2], FrameSLAM [3], V-GPS [4], Mini-SLAM [5] and others [6-10] that are competitive with or superior to range sensor-based algorithms, with routes as long as 1000 km being mapped [1]. The majority of these systems have been developed and demonstrated largely under certain conditions: high quality imaging sensors have been used, on relatively stable vehicle platforms and in bright illumination conditions, minimizing problems such as motion blur and changes in appearance. However, these are restrictive constraints, especially as robots are expected to operate over longer

* Michael Milford and Ashley George are with the School of Electrical Engineering and Computer Science at the Queensland University of Technology. Correspondence should be addressed to michael.milford@qut.edu.au. This work was supported by an Australian Research Council Discovery Early Career Researcher Award DE120100995 to MM.

periods of time and with lower hardware costs. Many growing fields such as environmental monitoring could benefit greatly from the availability of a small, low cost robot platform with an all-day, all weather mapping and navigation capability that is not reliant on GPS or environmental beacons. Towards that aim, in this paper we seek to address two of the major challenges facing visual mapping systems:

1. The difficulty of obtaining high quality images required by feature-based techniques, when using low cost hardware at speed on off-road terrain and in poor lighting.
2. The problem of achieving reliable place recognition in an outdoor environment over the course of a day and during all types of weather.

Figure 1 illustrates these two challenges using camera images. Large changes in illumination (compare Figure panels 1a and 1d) or changes in the weather (see rain drops on lens in Figure 1c) can radically alter the types of features detectable by a state of the art algorithm such as Scale-Invariant Feature Transforms (SIFT) [11] and Speeded Up Robust Features (SURF) [12]. Furthermore, in poor lighting with low cost hardware and on off-road terrain, image blur is hard to avoid (Figure 1c, also Figure 8-4). Motion blur affects both the place recognition and odometry components of a mapping system, while change in appearance over the course of a day primarily affects place recognition.

(a) (b) (c) (d)

Fig. 1. Visual change in an environment over the course of a day and in varying weather – (a) dawn, (b) morning, (c) rain and (d) dusk. As well as changing illumination other challenges are present such as motion blur from the jerky motion of the platform when travelling off-road.

To some degree these problems can be reduced by using more capable sensing equipment and implementing techniques such as high dynamic range [13]. However, high dynamic range techniques degrade in viability as the speed of the platform increases. Without active illumination of an environment, even long exposure images can look very different to an image obtained in sunlight during the day. Motion estimation from motion blurred images can be achieved by tracking edges, but is more difficult to incorporate into a mapping process [14]. More capable sensors and lenses are expensive, usually bulkier and heavier to accommodate larger imaging sensors and lenses, and require more power. While this approach is viable on large expensive platforms where the sensor cost is relatively small, there is an increasing interest in cheap robot platforms for large scale operations such as ecology monitoring. On these platforms size and cost considerations make such an approach unfeasible. Ultimately, even with sophisticated hardware, there are physical limits to optics which are unlikely to be solved in the near future.

In this paper, we describe research towards enabling any-time vision-based SLAM for outdoor robots in changing environments equipped with cheap consumer-grade cameras. The focus is on scenarios where, due to the combination of cost limitations, illumination changes and challenging terrain, the ability to reliably recognize traditional visual features is limited. We present a lightweight visual recognition algorithm based on patch normalization techniques that provides a high degree of invariance to changes in environment conditions such as lighting. A patch tracking algorithm provides visual odometry information, while the pose and visual filtering is provided by the biologically inspired RatSLAM system. We demonstrate the system working at real-time speed in a mixed off-road and urban environment at four different times of day with different environmental conditions – at dawn, during the morning, during a rain shower, and in fading light at dusk. The visual recognition algorithm is able to consistently recognize familiar places despite the changes in conditions. The visual odometry system is able to provide "good enough" motion information to perform reliable mapping and localization over all the datasets when combined with the visual loop closures. We also demonstrate the system is able to map and localize off all the datasets even when restricted to learning visual templates only at one time of day, showing that a single exposure to the environment is enough to enable navigation at the other times of day.

The work presented here builds on previous research including mapping of a suburban road network at different times of day [15, 16] and sequence-based localization on road networks [17, 18]. Unlike the highly constrained nature of a road network, this system is applied in a mixed urban and vegetated environment with off-road areas. The degree of perceptual change encountered in the datasets presented here is qualitatively larger than in [15]. We present a featureless approach to visual matching, rather than the feature and intensity profile-based techniques used in [15]. In contrast to [17, 18], which were localization only studies, we implement a full SLAM solution that calculates and uses motion information to build a map and localize within that map.

2 Approach

In this section we describe the visual recognition and visual odometry algorithms, and give a brief overview of the RatSLAM system.

2.1 RatSLAM System

Processing of the data output by the visual recognition and visual odometry algorithms is performed by the RatSLAM system. RatSLAM is a robot SLAM system based on models of the navigation processes thought to occur in the rodent brain, specifically the rodent hippocampus [19]. RatSLAM has been demonstrated mapping a large road network in static conditions [20] and a smaller road network with moderately varying illumination [15].

The RatSLAM system consists of three modules, shown in Figure 2. The local view cells encode visual scenes in the environment, with cells incrementally recruited to represent new distinct visual scenes as they are encountered. The pose cells are a network of highly interconnected neural units connected by both excitatory (positive or reinforcing) and inhibitory (negative) connections. They encode an internal representation of the robot's pose state, and filter both the place recognition and self-motion information provided by the visual recognition and visual odometry processes. Finally, the experience map is a graphical map made up of nodes called experiences that encode distinct places in the environment, and connected by transitions that encode odometry information. A graph relaxation algorithm [20] is run continuously on the experience map, resulting in the continuous map evolution seen in the video accompanying the paper and also shown in Figures 9 and 11. Further information on the RatSLAM system can be found in [20, 21].

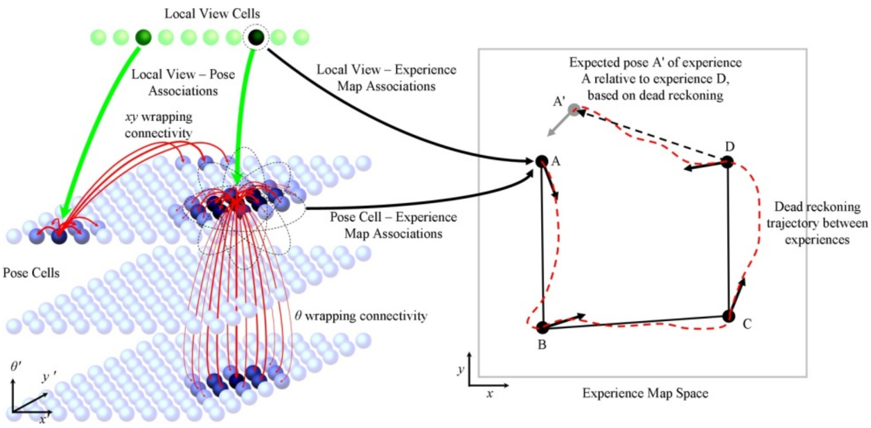

Fig. 2. The RatSLAM system. The local view cells encode distinct visual scenes, while the pose cells encode an internal representation of the robot's pose and perform filtering of place recognition estimates and self-motion information. The experience map is a graphical map formed by the combination of the output from the local view cells, pose cells and self-motion information.

2.2 Patch-Based Visual Odometry

The visual odometry system is a modified version of the system deployed on a quad rotor in [22]. The system tracks movement of two image patches to calculate translational speed and yaw of the platform, as shown in Figure 3a. The primary assumptions are that of a non-holonomic platform at a consistent height above the ground surface. Frame to frame motion of the top patch provides the yaw information and bottom patch motion provides the translational speed. The odometry gain was calibrated by running the car along a known length of ground and calculating the required gain constant, given in Table 2. Patch comparisons were performed by calculating the mean of the intensity difference between each pixel in the patch compared to the corresponding pixel in the previous image. Further implementation details are provided in [22].

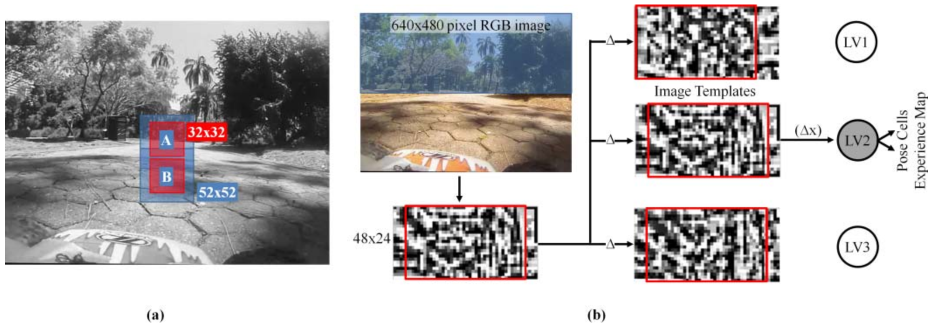

Fig. 3. (a) Patch-based visual odometry and (b) patch-normalized template matching.

2.3 Patch-Normalized Visual Template Learning and Recognition

The visual place recognition process is illustrated in Figure 3b. Camera images are captured and the bottom half removed. While the ground is useful for patch-based visual odometry, its proximity means that its appearance, when using a "whole of image" based recognition process, is sensitive to slight changes in vehicle pose when closing the loop, which tends to make place recognition brittle.

Once cropped, the image is resolution reduced to 48×24 pixels. Patch normalization is applied to the image in discrete square patches (rather than continuously over the image). Patch normalized pixel intensities, I', are given by:

$$I'_{xy} = \frac{I_{xy} - \mu_{xy}}{\sigma_{xy}} \tag{1}$$

where μ_{xy} and σ_{xy} are the mean and standard deviation of pixel values in the patch of size P_{size} that (x, y) is located within. Mean image differences between the current visual scene and all the learnt visual templates are calculated using a normalized sum of intensity differences, performed over a range of horizontal offsets:

$$D_j = \min_{\Delta x \in [-\sigma, \sigma]} g(\Delta x, i, j) \tag{2}$$

where σ is the template offset range, and $g(\,)$ is given by:

$$g(\Delta x, i, j) = \frac{1}{s} \sum_{x=0} \sum_{y=0} \left(p^{i}_{x+\Delta x, y} - p^{j}_{x,y} \right) \tag{3}$$

where s is the area in pixels of the template sub frame. If the minimum difference across all existing templates and relative offsets is larger than a threshold D_t, a new template is learned. Otherwise an existing template is matched, leading to activation of pose cells associated with that visual scene and a possible loop closure event. The range of horizontal offsets provides (assuming the majority of objects in the image are relatively distal) some invariance to camera pose. This invariance enables loop closure even when routes are repeated at slightly different lateral offsets or at different orientations. This capability is important for off-road motion (in contrast to movement along a road network) where repeated paths vary due to environmental change or variation in the path executed by the human or autonomous navigation system.

3 Experimental Setup

This section describes the testing platform, camera, environment and datasets used for this work.

3.1 Testing Platform and Camera

The testing platform was a Team Losi Mini-LST2 remote control car with a Contour+ camera mounted facing forwards. The camera has a fisheye wide-angle lens (2.8mm focal length, approximately 170º field of view) and logged GPS data. Figure 4a shows the platform, while Figure 4b shows an autonomous version under development. Due to the risk of water damage during the rain dataset and extreme nature of some of the off-road terrain (small logs, deep leaf litter) the non-autonomous platform was used. The video feed and GPS coordinates were logged onboard and processed offline. To reduce the effect of vibration and jerkiness due to the rough terrain and small size of the vehicle, videos were run through a stabilizing filter (VirtualDub Deshaker filter, available at [23], default values used). The use of a stabilizer introduces a one frame lag between image capture and the image being available to the localization and odometry routines, equivalent to 33 milliseconds at real-time speed.

3.2 Testing Environment and Datasets

Experiments were run over a one-week period in an area including the Queensland University of Technology campus and the City Botanic Gardens in Brisbane, Australia (Figure 4c). The testing area measures approximately 200m × 200m and contains a mixture of open grass, pathways, gravel baths, shrubbery, garden beds and buildings. The car was remotely driven by an operator following the vehicle.

A set of four datasets were gathered under a range of environmental conditions and at different times of the day (Table 1). Each dataset repeated the same route, although minor deviations were inevitable due to pedestrian traffic, construction work and the difficulty of the terrain in sections. A single traverse of the route was approximately 1310 meters in length (calculated by tracing the route on an aerial map) and took an average of approximately 15 minutes to complete. The car was jammed twice by branches and leaf litter and was stopped temporarily to remove the obstructing objects. These sections of video were cut, resulting in several discontinuous jumps in the footage. Frames were logged at 30 frames per second, with every frame processed by the visual odometry system but only every 5th frame processed by the visual template system, due to the high degree of overlap between neighboring frames. The 4 datasets are available online[1].

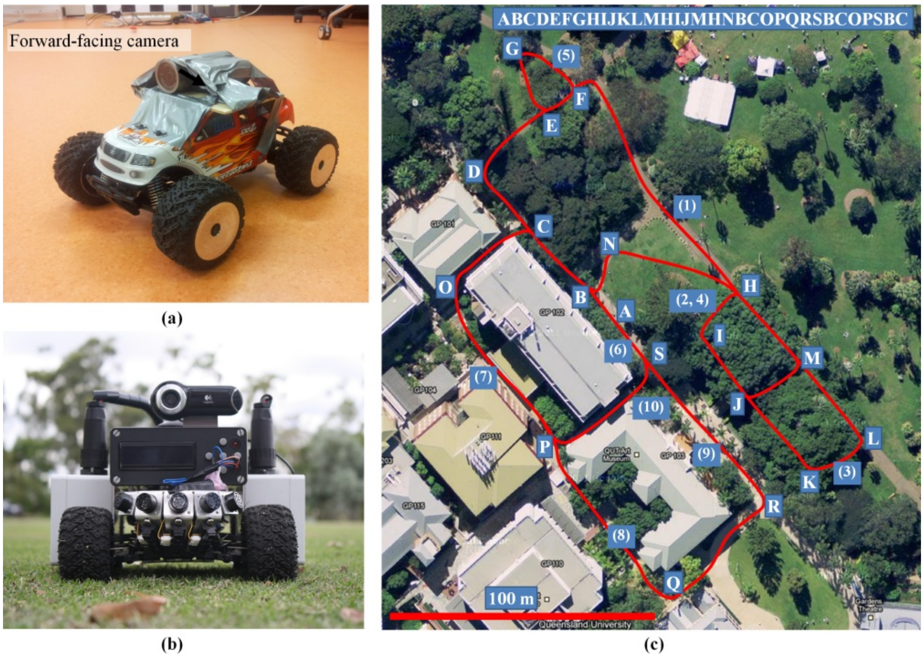

Fig. 4. (a) Testing platform, a small but capable off-road enthusiast hobby car with mounted consumer camera, and (b) an autonomous version under development. (c) The vehicle path, with order indicated by the letter sequence. The numbers show the sample frame match locations from Figure 8. Aerial photos from Google Maps.

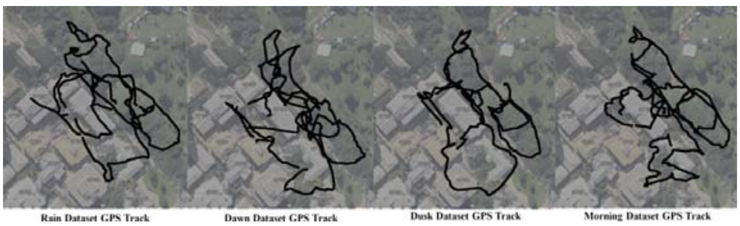

Fig. 5. GPS was unreliable especially under tree cover and around buildings.

An attempt was made to use GPS tracking (CEP 10 m) as a ground truth measure. However, due to the heavily vegetated and urban canyon nature of much of the environment, the quality of the GPS tracking was too poor to be useful (far worse than specifications), as shown in Figure 5.

[1] https://wiki.qut.edu.au/display/cyphy/Michael+Milford+Datasets+and+Downloads

Table 1. Dataset descriptions. Times in Australian Eastern Standard Time (AEST).

Dataset Name	Time and Comments
Dawn	5:45 am. Sun just above local horizon, most areas in shade, excessive sun flare in sections.
Morning	10:00 am. Sun high up in sky, large ground areas in bright sunlight.
Rain	10:30 am. Rain drops on lens, wet ground, overcast and dark.
Dusk	6:45 pm. Sun setting, extremely dark in heavily vegetated areas, significant motion blur and lack of ground texture.

Table 2. Parameters.

Parameter	Value	Description
r	32 pixels	Odometry patch size
ς	0.375 °/pixel	Yaw gain constant
v	0.0125 m/pixel	Translational speed constant
ρ	10 pixels	Odometry patch offset range
s	48×24 pixels	Template sub frame size
D_t	0.06	Template learning threshold
σ	4 pixels	Template offset range

4 Results

In this section we present the visual odometry, place recognition and mapping results as well as computational statistics.

4.1 Visual Odometry

Figure 6 shows the trajectory output by the patch-based visual odometry system for all four datasets, for the common starting pose of (x,y,θ) = (*0 m, 0 m, 0 degrees*). Although the trajectories clearly do not match on a global scale, subsections of the route are similar for all four datasets, such as the small loop (sequence *EFGEF*) in Figure 4. The differences in the odometry-only trajectories were primarily caused by underestimation of yaw angles and translational speeds in the rain dataset, probably due to reflections in the water lying on the ground, and underestimation of the translational speed in the dusk dataset, due to the poor illumination and consequent lack of ground textures. The differences in translational speed calculations are most easily seen by looking at the length of the first section of each trajectory starting at (0,0) leading up to the first right turn.

4.2 Visual Place Recognition

Figure 7 displays a graph of the active (recognized or just learnt) visual template versus frame number over all four datasets in the order they were processed, starting with the dawn dataset. The area of the graph below the dashed line is the area in which visual templates learned during the first dawn traverse of the environment were recognized during the subsequent datasets. The system was able to recognize places from the dawn dataset at regular intervals throughout the other three datasets. However, the graph also shows additional templates representative of the subsequent datasets being learnt in parallel and bound to those locations in the map. Learning of new templates was due to the zigzag nature of much of the robot's movement through the environment, resulting in different image sequences each time a section was traversed.

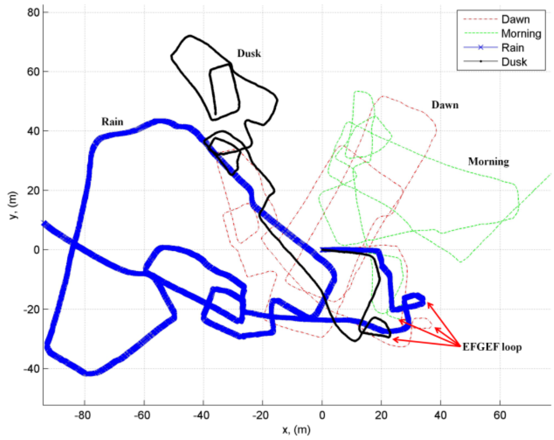

Fig. 6. Vehicle trajectories calculated by the patch-based visual odometry system for the four datasets.

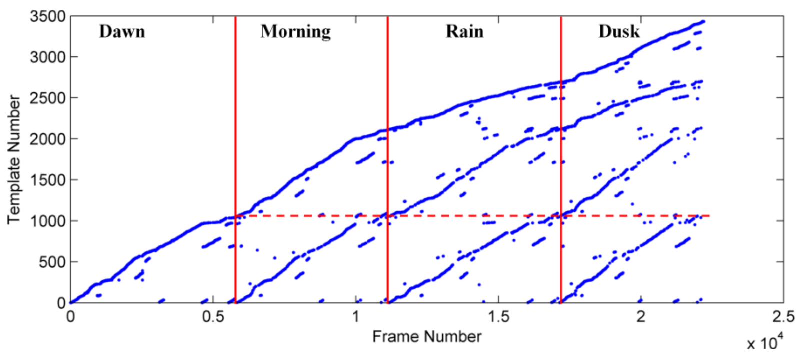

Fig. 7. Visual template learning and recognition over the four datasets.

4.3 Matched Frames

Figure 8 shows a selection of ten pairs of frames that were matched by the visual template system for locations throughout the entire route. The original video frames are shown for clarity purposes, although the actual processed images were 48×24 pixel patch-normalized images. The corresponding locations are shown in Figure 4.

The visual system was able to match frames with significantly varying appearance due to (1, 3) sun flare, (2) obscuring leaf litter, (4) motion blur, (5-7) major shadow change, (3, 6, 9-10) large overall illumination change and (10) water on the camera lens. The frames also show the challenge faced by the visual odometry system due to jerky vehicle motion (4) and lack of ground texture in low light (1, 3, 6-10).

Fig. 8. Matched visual templates over the four datasets. Corresponding locations are shown in Figure 4.

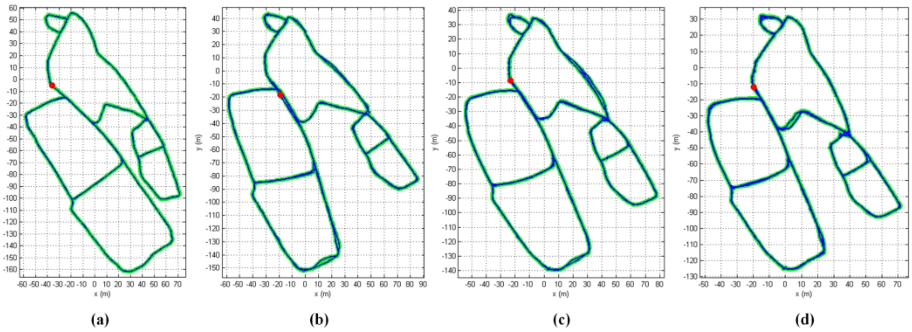

Fig. 9. Experience map evolution over time. Experience maps are from after the (a) dawn, (b) morning, (c) rain and (d) dusk datasets.

4.4 Experience Maps

The final test of the system was to create a map of all four datasets. Figure 9 shows the evolution of the experience map after running through each dataset in order. The map is topologically correct after the dawn and morning datasets, although globally it is warped. The map shrinks slightly, primarily due to the underreporting of translational velocity in the dusk dataset and to a lesser extent the rain dataset. However, the constant loop closure within and across datasets ensures the map topology remains correct. The final map layout, although not metrically precise, has the correct topology. A video of the experience map and frame matching processes is available online[2].

4.5 SLAM with Only Visual Templates from a Single Time

To test the ability of the system to map and localize with only the visual templates learned at one particular time of day, we conducted an additional experiment where template learning was disabled after the first dawn dataset. From that point onwards the visual template system either recognized a familiar template or reported no match, but did not learn any additional templates (Figure 10). Figure 11 shows the evolution of the experience map under these conditions. There are three locations where place recognition failed briefly, all at places where the vehicle was turning corners and actual physical paths varied significantly. Although successful loop closures were achieved surrounding these points, the variation in visual odometry meant that the graph relaxation process was not able to draw these trajectories together to correctly overlap. The local topology in these areas is incomplete but correct, meaning navigation could still be achieved but might be suboptimal.

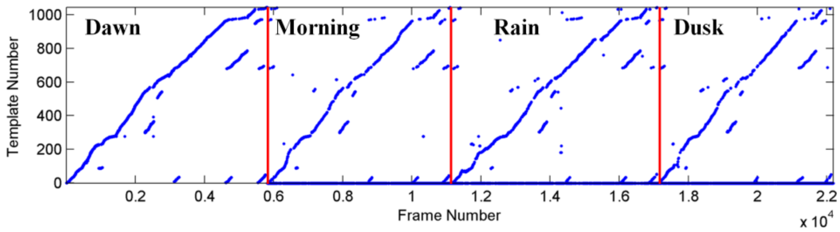

Fig. 10. Visual template recognition performance with learning only enabled for the dawn dataset. Non-matches where a template would normally be learned appear as number zero templates.

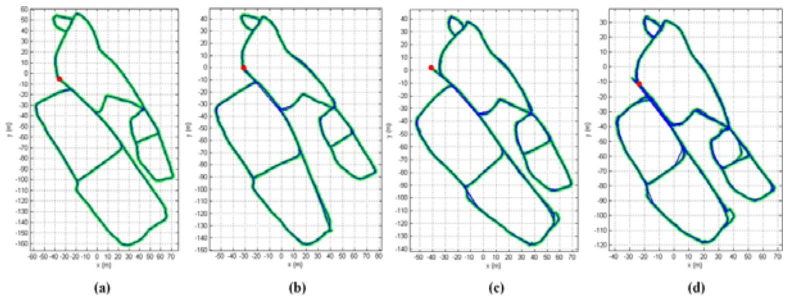

Fig. 11. Experience map evolution with template learning disabled after the first dataset. Map shown after the (a) dawn, (b) morning, (c) rain and (d) dusk datasets.

[2] https://wiki.qut.edu.au/display/cyphy/Michael+Milford+Datasets+and+Downloads

4.6 Compute and Storage

To demonstrate the feasibility of real-time performance on low cost hardware, we present some pertinent computational statistics. The primary storage requirements come from the visual template library. Over all four datasets, a total of 3353 templates were learned, taking up 5.8 MB of storage. Compute wise, the system performs all computation on a fixed time basis, except for visual template comparison and experience map graph relaxation which are both order O(N) (experience map graph relaxation approximates to order O(N) in a typical sparsely interconnected map). Each of these two processes was run on a separate CPU on a standard desktop PC. At the end of the dusk dataset when the system was at maximum load the visual template system was performing 104 million pixel to pixel comparisons per second of data, which ran at real-time speed in unoptimized Matlab code. Experience map graph relaxation is performed with leftover compute cycles. At the end of the experiment, an average of 156 global graph relaxation iterations were performed per second of real-time. This figure can be compared with the 8 iterations per second performed at the end of a previous indoor mapping experiment [24], which was still sufficient to maintain a map that was used for robot navigation. A low power onboard CPU (such as the 1 GHz processor on the robot shown in Fig. 4b) should be capable of running the entire system in real-time for an environment of this size. The RatSLAM system used as the mapping backend has had lightweight versions implemented on a Lego Mindstorms NXT [25] and a small mobile robot called the *iRat* [26], demonstrating the feasibility of running the system on a cheap platform.

5 Discussion

This paper presents a study into the feasibility of using a lightweight, “whole of image” approach to vision-based SLAM on small, low cost vehicles expected to operate in a wide range of environments and in highly varied conditions. The visual processing techniques require no prior training[3], and are demonstrated to enable topological mapping in a varied vegetated and urban environment. Furthermore, the results demonstrate the viability of the approach in a wide range of conditions such as varying time of day and weather. Lastly, the techniques are able to create and consistently localize within a topological map even when it is not possible to obtain high quality visual odometry, such as during the rain and dusk datasets, and when traditional visual features are not available in blurred or very dark images. Here we discuss the limitations of the presented approach and areas for future work.

We used a forward facing camera only, and hence had no ability to close the loop when retracing a route in the opposite direction. However, past work has demonstrated that such a forward facing system can be adapted to utilize omnidirectional imagery [24, 27]. The ability of the system to function with low resolution imagery would also be likely to enable the combination of cheap and compact panoramic im-

[3] No training is required to generate a topological map. To obtain a map with absolute scale, a short calibration of the translational gain constant is required.

aging rigs with a low cost camera (the mirror could be mass produced with loose specifications). In contrast, much current robot research makes use of high end panoramic imaging setups such as the Point Grey Ladybug 2 (~10000 USD). Alternatively, two perspective cameras mounted in opposite directions along the primary vehicle axis would provide forward-backwards recognition capability.

The visual template system is not suited to open-field operation in large open environments where movement is unrestricted and paths are not necessarily repeated. However, this restriction is also present in many vision-based SLAM systems developed to date. One common approach to overcoming this limitation is to combine a SLAM system with absolute positioning information provided by GPS, when available. It is interesting to note that GPS availability and visual SLAM viability tend to be complementary, at least in the system presented in this paper. In the mixed urban and vegetated environment, when GPS was unavailable the vehicle was usually travelling along urban canyons or off-road paths where paths are constrained, situations in which the presented approach works well.

Future work will pursue a number of research directions beyond those mentioned above. The first will be to pursue optimization of the template matching algorithm, which is predicted to be the computationally limiting factor as environments get larger. Secondly, we will investigate how best to add a feature-based mapping technique such as FAB-MAP; FAB-MAP will provide a higher degree of pose invariance when features are detectable, while the visual template method will bind together map locations where features are not reliably detected. Lastly, the quality of the maps exceeds that of those used successfully for robot navigation previously [24], suggesting navigation using these maps is feasible. We will investigate combining state of the art local obstacle avoidance techniques with RatSLAM navigation algorithms [24] in order to enable navigation under challenging and changing environmental conditions.

6 References

1. M. Cummins and P. Newman, "Highly scalable appearance-only SLAM - FAB-MAP 2.0," in *Robotics: Science and Systems*, Seattle, United States, 2009.
2. A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, "MonoSLAM: Real-Time Single Camera SLAM," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 29, pp. 1052-1067, June 2007 2007.
3. K. Konolige and M. Agrawal, "FrameSLAM: From Bundle Adjustment to Real-Time Visual Mapping," *IEEE Transactions on Robotics,* vol. 24, pp. 1066-1077, 2008.
4. D. Burschka and G. D. Hager, "V-GPS (SLAM): Vision-based inertial system for mobile robots," 2004, pp. 409-415 Vol. 1.
5. H. Andreasson, T. Duckett, and A. Lilienthal, "Mini-SLAM: Minimalistic Visual SLAM in Large-Scale Environments Based on a New Interpretation of Image Similarity," in *International Conference on Robotics and Automation*, Rome, Italy, 2007, pp. 4096-4101.
6. H. Andreasson, T. Duckett, and A. Lilienthal, "A Minimalistic Approach to Appearance-Based Visual SLAM," *IEEE Transactions on Robotics,* vol. 24, pp. 1-11, 2008.
7. L. M. Paz, P. Pinies, J. D. Tardos, and J. Neira, "Large-Scale 6-DOF SLAM With Stereo-in-Hand," *IEEE Transactions on Robotics,* vol. 24, pp. 946-957, 2008.

8. E. Royer, J. Bom, M. Dhome, B. Thuilot, M. Lhuillier, and F. Marmoiton, "Outdoor autonomous navigation using monocular vision," in *IEEE International Conference on Intelligent Robots and Systems*, 2005, pp. 1253-1258.
9. A. M. Zhang and L. Kleeman, "Robust Appearance Based Visual Route Following for Navigation in Large-scale Outdoor Environments," *The International Journal of Robotics Research,* vol. 28, pp. 331-356, March 1, 2009 2009.
10. K. Konolige, M. Agrawal, R. Bolles, C. Cowan, M. Fischler, and B. Gerkey, "Outdoor mapping and navigation using stereo vision," 2008, pp. 179-190.
11. D. G. Lowe, "Object recognition from local scale-invariant features," in *International Conference on Computer Vision*, Kerkyra, Greence, 1999.
12. H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded Up Robust Features," in *Computer Vision – ECCV 2006*, ed, 2006, pp. 404-417.
13. I. Kiyoshi, Y. Tomoaki, and T. Masahiro, "A High Dynamic Range Vision Approach to Outdoor Localization," in *International Conference on Robotics and Automation*, Shanghai, China, 2011.
14. G. Klein and D. Murray, "Improving the agility of keyframe-based SLAM," *European Conference on Computer Vision,* pp. 802-815, 2008.
15. A. J. Glover, W. P. Maddern, M. J. Milford, and G. F. Wyeth, "FAB-MAP + RatSLAM: Appearance-based SLAM for Multiple Times of Day," in *International Conference on Robotics and Automation*, Anchorage, United States, 2010.
16. M. Milford and G. Wyeth, "Improving Recall in Appearance-Based Visual SLAM using Visual Expectation," in *Australasian Conference on Robotics and Automation*, Brisbane, Australia, 2010.
17. M. Milford and G. Wyeth, "SeqSLAM: Visual Route-Based Navigation for Sunny Summer Days and Stormy Winter Nights," in *IEEE International Conference on Robotics and Automation*, St Paul, United States, 2012.
18. M. Milford, "Towards Condition-Invariant Sequence-Based Route Recognition," in *Australasian Conference on Robotics and Automation*, Melbourne, Australia, 2011.
19. M. J. Milford, *Robot Navigation from Nature: Simultaneous Localisation, Mapping, and Path Planning Based on Hippocampal Models* vol. 41. Berlin-Heidelberg: Springer-Verlag, 2008.
20. M. Milford and G. Wyeth, "Mapping a Suburb with a Single Camera using a Biologically Inspired SLAM System," *IEEE Transactions on Robotics,* vol. 24, pp. 1038-1053, 2008.
21. M. Milford and G. Wyeth, "Persistent Navigation and Mapping using a Biologically Inspired SLAM System," *The International Journal of Robotics Research,* 2009.
22. M. Milford, F. Schill, P. Corke, R. Mahony, and G. Wyeth, "Aerial SLAM with a Single Camera Using Visual Expectation," in *International Conference on Robotics and Automation*, Shanghai, China, 2011.
23. G. Thalin, "Deshaker - video stabilizer," 2.5 ed, 2010.
24. M. Milford and G. Wyeth, "Persistent Navigation and Mapping using a Biologically Inspired SLAM System," *International Journal of Robotics Research,* vol. 29, pp. 1131-1153, 2010.
25. "RatSLAM NXT - http://code.google.com/p/rsnxt08/," ed: The University of Queensland, 2008.
26. D. Ball, S. Heath, M. Milford, G. Wyeth, and J. Wiles, "A navigating rat animat," ed: MIT Press, pp. 804-811.
27. D. Prasser, M. Milford, and G. Wyeth, "Outdoor simultaneous localisation and mapping using RatSLAM," in *International Conference on Field and Service Robotics*, Port Douglas, Australia, 2005.