2D Visual Localization for Robot Vacuum Cleaners at Night

James Mount, Venkateswara Rao Rallabandi, Michael Milford

Abstract— **Vacuum cleaning robots are by a significant margin the most populous consumer robots in existence today. While early versions were essentially dumb random exploration cleaners, recent systems fielded by most of the major manufacturers have attempted to improve the intelligence and efficiency of these systems. Both range-based and visual sensors have been used to enable these robots to map and localize within their environments, with active range sensing and active visual solutions having the disadvantage of being more intrusive or only sensing a 1D scan of the environment. Passive visual approaches such as those used by the Dyson robot vacuum cleaner have been shown to work well in ideal lighting conditions; their performance in darkness is unknown. In this paper we present research working towards a passive and potentially very cheap vision-based solution to vacuum cleaner robot localization that utilizes low resolution contrast-normalized image matching algorithms, image sequence-based matching in two-dimensions, and place match interpolation. In a range of experiments in a domestic home and office environment during the day and night, we demonstrate that the approach enables accurate localization regardless of the lighting conditions and lighting changes experienced in the environment.**

I. INTRODUCTION

Personal robots have been on the market since the early 1950s, but it has only been with the advent of vacuum cleaner robots that they have achieved widespread market penetration. Vacuum cleaning may seem to be an easy task, but for robots it is a challenging problem that has required many years of research. Navigation and localization in particular are required capabilities that are challenging to develop, with current "solutions" using inexpensive lasers, high quality cameras during the day-time, or avoiding the problem completely by instead implementing random movement behaviours.

Due to rapid increases in camera capabilities and computer processing power, vision has become an increasingly popular sensor for robotic navigation and object classification. Vision provides a variety of cues about the environment, such as motion, colour, and shape, all with a single sensor, and has advantages over other sensors including low cost, small form factor and low power consumption [6], all relevant characteristics in the context of vacuum cleaner robots. However, visual sensors are highly sensitive to a robot's viewpoint and environmental lighting conditions, and current passive vision-based autonomous vacuum cleaning systems have not yet been demonstrated to work robustly under challenging illumination conditions. Vision-based navigation solutions are typically troubled by both low light conditions and scenes with highly varied illumination, conditions which are common in the domestic home.

Figure 1: A robot vacuum cleaner base equipped with the Ricoh Theta camera used in this work, with a sample 360 × 180 degree image from the camera.

This paper presents a new localization system based on low resolution, contrast-enhanced image comparison, sequence-based image comparison in two dimensions, and place match interpolation, in order to enable accurate localization in the home by a vacuum cleaning robot. We demonstrate the effectiveness of the system in both a domestic home and an office environment, during both day-time and night-time. The aim of our work is to develop a set of vision-based localization algorithms that could be employed on robotic platforms operating in human environments, like the domestic home, using inexpensive visual and computational hardware.

The paper proceeds as follows. In Section II we provide a short literature review on autonomous vacuum cleaners, place recognition approaches for robots, and discuss the nature of the vision invariance problem. In section III we provide an overview of the approach taken, while Section IV summarises the experimental setup. Section V presents the results, with discussion in Section VI.

II. BACKGROUND

Robot vacuum cleaners have been in development since the 1980s. Only in the past decade however have they become a household name, with iRobot selling an estimated 6 million "Roomba" robots between the years 2002-2010 [Vaussard et al., 2014]. Later models have implemented more sophisticated technologies, such as navigation and path planning methods, compared to their earlier siblings, which used random path methods and simple behaviours such as "edge-following" and "spiral" [Vaussard et al., 2014]. These improvements have further increased the efficiency of autonomous cleaners and

JM, VR and MM are with the Australian Centre for Robotic Vision and the School of Electrical Engineering and Computer Science at the Queensland University of Technology, Brisbane, Australia, michael.milford@qut.edu.au.

This work was supported by an Australian Research Council Future Fellowship FT140101229 to MM.

consequently their consumer appeal, by lowering energy consumption, time to job completion, and improving the robot's floor coverage.

To achieve these improvements, robot vacuum cleaner systems have used IR, 2D laser scanners, and/or ceiling facing cameras to help quantify, map, and navigate an area [Vaussard et al., 2014], in conjunction with Simultaneous Localization And Mapping (SLAM) algorithms. Simultaneous Localization and Mapping is the process of learning, through mapping, an unknown environment while simultaneously localizing a robot's position. While theoretically the problem can be thought of as solved, practically there are still several challenges including dealing with varying illumination and the challenge of building contextually rich maps for use in SLAM algorithms [Durrant-Whyte & Bailey, 2006]. Vacuum cleaning robots require precise area coverage, as well as accurate and robust localization.

The SLAM problem is based in probabilistic theory, and there are a variety of algorithms used to solve the problem. These methods include the Extended Kalman Filter (EKF), which utilizes a linearized state space model in conjunction with a Kalman Filter, and FastSLAM, which implements a Rao-Blackwellized particle filter [Durrant-Whyte & Bailey, 2006]. Other methods include MonoSLAM [Davison et al., 2007], FrameSLAM [Konolige & Agrawal, 2008], V-GPS [Burschka & Hager, 2004], Mini-SLAM [Andreasson et al., 2007] and [Andreasson et al., 2008; Cummins & Newman, 2009; Konolige et al., 2008; Milford & Wyeth, 2008, 2012; Paz et al., 2008; Royer et al., 2005; Zhang & Kleeman, 2009]. An important component of the SLAM problem is that of place recognition or loop closure; the key challenge that we address in this research.

Place recognition is the process of matching the current sensory snapshot to a previously learned sensory snapshot, and is a key component of mapping methods that form topological maps enabling robots to navigate [A et al., 2010]. Generally speaking, recognition algorithms can be split into two categories, global methods and local methods. Global methods operate over a large environment, while local methods work over a subset of the environment, but assume the adjacent neighbourhood is known. This means that local methods typically produce quantitative estimates, while global methods produce a more qualitative estimate [Dudek & Jugessur, 2000]. Since the appearance of the environment can change through human interference, such as moving a piece of furniture or turning off a light, place recognition algorithms that rely on an unchanging environment are likely to fail in human generated environments [Yamauchi & Langley, 1997].

III. Approach

In this section we provide a high level overview of the system architecture (Figure 2). Our approach in this work is based on the assumption that a robot vacuum cleaner would occasionally be run during the day in good lighting and with a source of motion information (from either wheel encoders or visual odometry), enabling the robot to gather a reference map of day-time images against which night-time localization can be performed. This approach is a reasonable one, as it is likely that current camera-based robot vacuum cleaners, such as the Dyson 360 Eye, are fully capable of generating a day-time map.

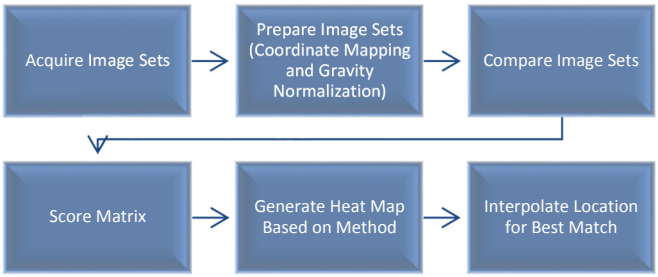

Figure 2: System architecture overview diagram.

A. Image Set Acquisition

The first step in the process is gathering a reference map of the environment during the day-time, consisting of a topological map and associated camera images at each of the map nodes. We designed a path through the environment with labelled markers for the purpose of ground truthing, and followed this path with a camera, taking images at equally spaced intervals (Figure 3). A second set of images was also acquired along a different, only partially overlapping path through the environment, which served as our query / test dataset. All acquired images were also manually mapped to a set of room co-ordinates for the purpose of later analysis.

B. Image Set Preparation

Image pre-processing involved stabilizing the panoramic images using the camera's inbuilt gravity sensor, to ensure that rotation variance in the third dimension would not affect the image comparison algorithms. Images were then histogram equalized, cropped slightly to remove the vacuum cleaner or camera mount base and resolution reduced. Finally, patch normalization was performed to reduce the effects of local variations in illumination, such as patches of sunlight on the floor which disappear at night. The patch normalized pixel intensities, I', are given by:

$$I'_{xy} = \frac{I_{xy} - \mu_{xy}}{\sigma_{xy}} \quad (1)$$

where μ_{xy} and σ_{xy} are the mean and standard deviation of pixel values in a patch of size P_{size} surrounding (x, y).

C. Image Set Comparison and Score Matrix

Images from each query/test dataset were compared to all images in the reference datasets using a rotation-invariant matching process. Each query image was compared using a sum of absolute differences to every image stored in the reference map, at all possible image rotations. The difference score for the k^{th} rotation, $C(k)$, is given by:

$$C(k) = \frac{1}{h \times w} \sum_{i=1}^{h} \sum_{j=1}^{w} \left| QS(i,j)_k - RS(i,j) \right|_{k=1}^{k=n} \quad (2)$$

where h and w are the size of the patch normalized image in the vertical and horizontal directions, respectively. $RS(i,j)$ is the cropped, down-sampled and patch normalized reference

set image, $QS(i,j)_k$ is the cropped, down-sampled and patch normalized query set image at the k^{th} rotation, and n is the number of pixel offset rotations. The difference scores between a query image and all the reference images were stored within an image difference matrix.

D. Heat Map Generation and Best Match Location

To exploit the two-dimensional nature of the environment and enable sequence-matching in two dimensions, a place match heat map was generated for each query image. To generate the heat map, the minimum rotated matching score between the given query image and all reference images was found:

$$MinScores(i) = \min(scores(j)_i) \quad (3)$$

where i represents the i^{th} reference image compared to the current query image, and where $scores(j)_i$ represents the scores for the current query image against the i^{th} reference set image for all relative image rotations. For visualization purposes, the values within this minimum score matrix were then inverted so that the maximum value "hot spot" corresponded to the best match.

To generate a smooth heat map even with discontinuous reference map image locations, image comparison scores were linearly interpolated across a regular overlaid grid to generate the heat map. Figure 8 shows an example of the resultant regular heat map, showing the response for comparison of a query image against all reference map images. The reference image locations are also plotted with green circles, with the circle size being directly proportional to the matching score for each location. Finally, the interpolated best match position P was found by finding the maximum matching score within the heat map:

$$P = \text{coordinate}(\max(\text{InvScores})) \quad (4)$$

The closest best match reference image was then also determined by finding the closest reference image location to the interpolated location.

E. Sequence Matching in 2D

Based on the success of sequence-based matching in varying lighting conditions in 1-dimension [Milford, 2013; Milford & Wyeth, 2012], we developed a two-dimensional sequence matching approach utilizing the heat map. Sequence-based heat maps were generated based not only on the current matching scores, but also on the n previous matching scores, depending on the number of frames used in the sequence.

To generate the sequential heat map, the previous interpolated heat map is taken and translated the same distance as the shift in the query image location from the previous query location and then summed with the current query image heat map. For these experiments we used simulated odometry; for a live robot implementation this data would need to come directly from either the robot's wheel encoders or a visual odometry system, or both. The best match position and closest reference image match were then found using the same process as for the single frame matching method.

IV. Experimental Setup

This section describes the experimental setup, dataset acquisition and pre-processing, ground truth creation and key parameter values. All processing was performed on a Windows 7 64-Bit machine running Matlab 2014 and the Ricoh Theta computer software.

A. Camera Equipment

A Ricoh Theta M15 camera was utilized for the majority of the experiments, with a Nikon D5100 Digital DSLR used for a single ultra-low-light experiment. The Ricoh Theta is a spherical camera that consists of two back-to-back fish eye lenses mounted on a slim body. The proprietary nature of the camera means the exact specifications of the Ricoh Theta camera are unknown, however it is estimated that the camera's approximately 5 mega-pixels images come from a small (compact-camera-like) sensor.

B. The Dataset

Experiments were performed using five datasets taken within a lounge/living room within a Brisbane townhouse, as well as within an internal office with ceiling lights but without windows. The datasets were taken using both the Ricoh Theta camera and the DSLR camera.

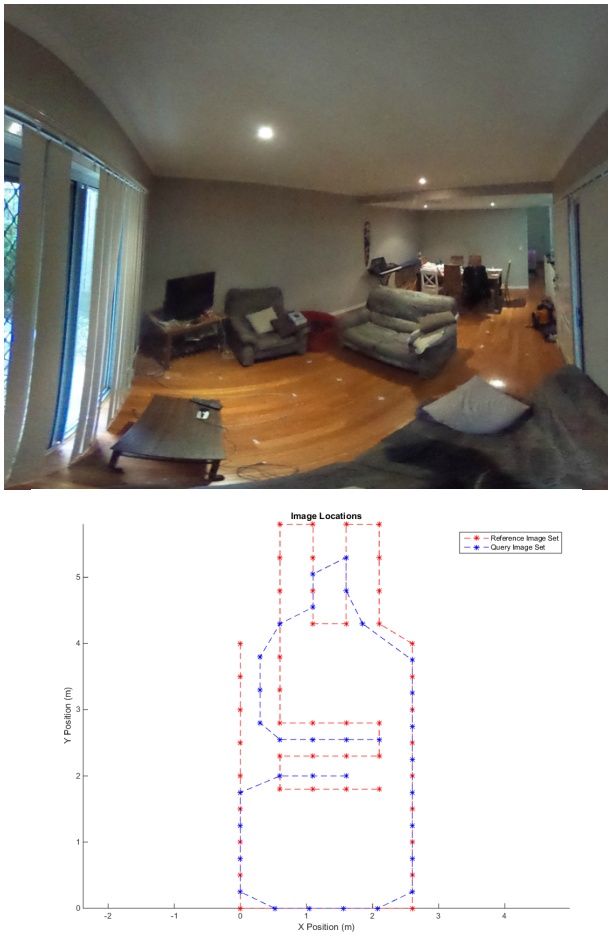

Figure 3: shows the path and image locations for the reference set (red) and query sets (blue) in the lounge room environment, as well as a photo taken with the Ricoh Theta camera of the entire area.

The first dataset, the reference set, consisted of 52 images in total, over the 6 by 3 metre lounge area. The second set, the daytime query set, was taken again during the day but at "random" locations throughout the referenced area, and consisted of 32 images. The final lounge image set, the night time query set, was taken at low light levels following the same path as the daytime query set, the second image set. The second and third (query) image sets traced the same path, which covered a majority of the area. Figure 3 shows the locations, and path, at which the reference set and query sets were taken. Query images did not necessarily overlap precisely with the reference images, creating a viewpoint invariance problem in addition to the condition-invariance problem.

The fourth and fifth datasets were taken within an internal (no windows) office space. The fourth image set was taken at 8 locations with the lights on and the door open. The fifth and final image set was taken at the same locations, but with the door closed and lights off to create an ultra-low-light environment. A summary of the datasets is shown in the following table.

TABLE I

DATASET SUMMARY

Name	Size	Frames	Location	Description
Reference Image Set	3x6 m	52	Townhouse	Reference set was taken during the daytime, with all house lights on, and taken at ground truthed points.
Daytime Query Set	3x6 m	32	Townhouse	Daytime query set was taken during the daytime, with all house light on, and taken at random locations throughout the reference set area.
Night-Time Query Set	3x6 m	32	Townhouse	Night-time query set was taken during the evening, with two small lamps and oven light on, and were taken at the same random locations as the daytime query set.
Office Reference Image Set	2x3 m	8	Internal Office	The internal office image set was taken in a small internal office with the lights on. It was taken both with the Ricoh Theta camera and the DSLR camera.
Office Image Query Set	2x3 m	8	Internal Office	The internal office query set was taken with the door closed and the lights off, which created a near pitch black environment. It was taken both with the Ricoh Theta camera and the DSLR camera.

C. Ground Truth

The first dataset was taken at ground truth points marked with masking tape throughout the lounge room (Figure 4). Each point was measured and marked out by hand using a tape measure and masking tape, from an arbitrarily placed origin point. Points were generally placed, within the constraints of the lounge room furniture, at points on a grid of square size 500mm.

Figure 4: The experimental robot vacuum cleaner with the Ricoh Theta camera attached in the home test environment.

D. Parameter Values

Parameter values are given in Table I. These parameters were heuristically determined over a range of development datasets and then applied to all the experimental datasets.

TABLE 2

PARAMETER LIST

Parameter	Value	Description
R_x,R_y	48,24	Whole image matching resolution
P_{size}	4	Patch-normalization radius
Interpolation Size	400,400	The size of the interpolated heat map.

V. RESULTS

In this section we present the results of the place recognition experiments. This section is split into 3 parts;

- The daytime results – which show the image matching results between the daytime reference set and the daytime query set.
- The night-time results – which show the results of the image matching between the reference set and the night-time query set.
- The DSLR results – which show matching performance when used with an alternative vision sensor in an ultra-low-light situation.

There is also a video accompanying the paper illustrating the results, also available at https://youtu.be/w2rz9c4zpZI.

A. Matching Performance – Daytime Results

The results of the daytime image matching and place recognition results can be found in following figures. Figure 6 shows a daytime query set image and its best reference set matched image, while Figure 7 shows the equalized, cropped, down-sampled and patch normalized images for a sample correct image matching pair.

Figure 8 shows the heat map results of a single frame image match. As can be seen, the reference image locations closest to the query image location (the red-cross) have the maximal matching scores, as indicated by the size of the green circles. The interpolated location for where the query image was taken is correct to within 0.2 metres 31% of the time, and within 0.4 metres 100% of the time, for the single frame matching method between the daytime query set and the reference set, see Figure 5 for error plots.

Figure 9 shows the results when using sequence SLAM methods. As can be seen, the reference image positions closest

to the location of the query image become more prominent, while false areas that were "hot" in the single frame matching method have become "cooler". However, since the dataset is not perceptually challenging, the difference between the single frame and sequence-based method is not as apparent as in the later night-time experiments. The following table summarizes the results of the comparison between the reference set and daytime query set in terms of the error in estimating/interpolating the location of where the query image was taken.

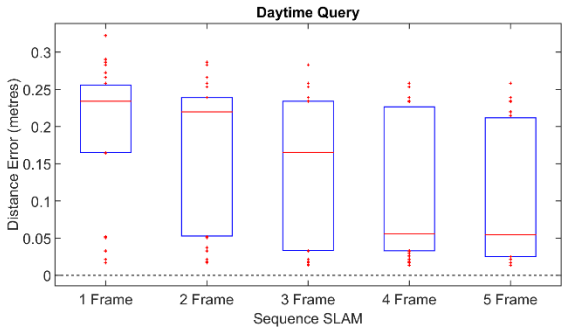

Figure 5: Shows error plots for each of the different size frames for sequence SLAM for the Daytime query set.

Figure 6: The 28th query image (top) and the best match reference image for the daytime query set (bottom).

Figure 7: The cropped, down-sampled and patch normalized images for the 28th daytime query image (top) and the best match reference image for the daytime query set (bottom). The query image has been rotated to the rotation at which the best match was found.

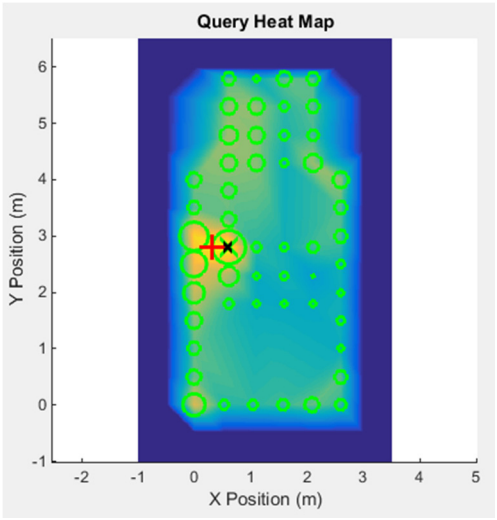

Figure 8: The heat map for the 28th query image in the daytime reference set. The red-cross shows the ground truth of the query image, while the green cross shows the best matched reference image, and the black cross indicates the best interpolated position (the "hot spot" in the heat map). The green circles are at the coordinates of the reference set image locations, and their size are indicative of how well the current query set image matches to each reference image.

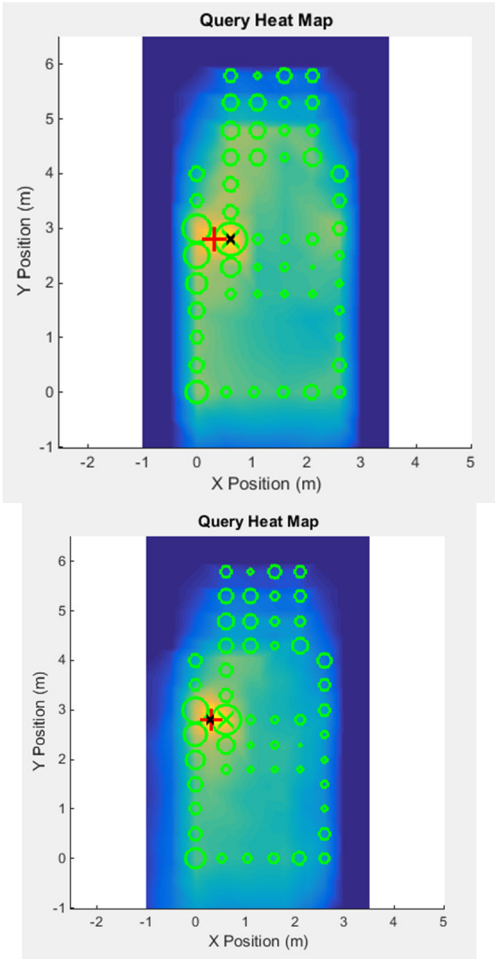

Figure 9: The sequence-based heat maps for the 28th query image of the daytime query set with different sequence lengths. The top heat map is for 3 point sequence SLAM, while the bottom heat map is for 5 point sequence SLAM.

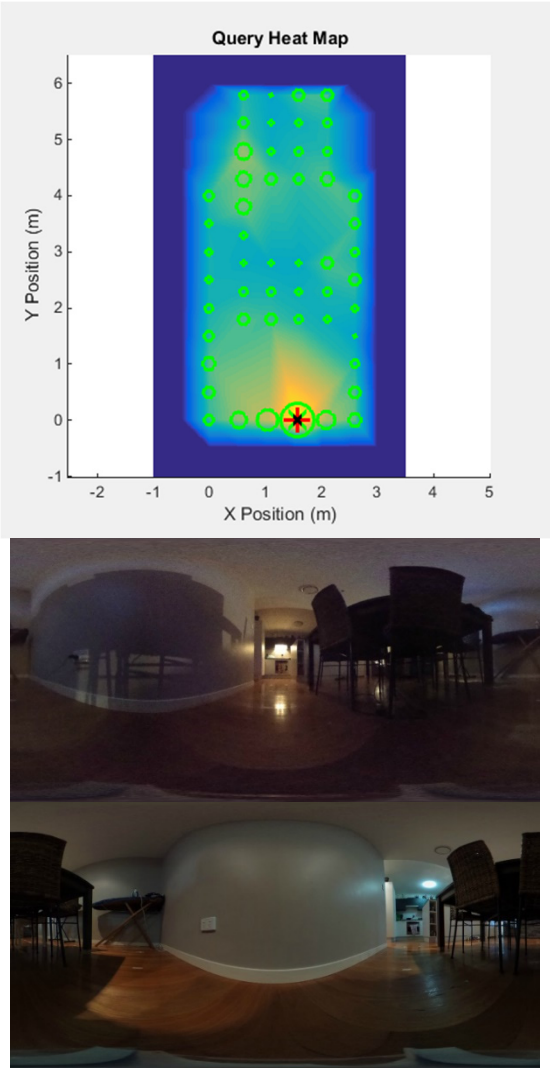

Figure 10: A true positive match using single image matching for the night-time query set. It shows the heat map and the image for the 10th night-time query set image (top image), as well as the best matched image from the reference set (bottom image).

B. *Night-Time Results*

The results of matching from night-time query images to daytime reference images can be seen in the following figures. Figure 10 and Figure 12 show the results of a positive and a false positive match, respectively, for the single frame matching method. As can be seen even in the false positive case, the reference cell green circle nearest the query set location (red-cross) is still significantly larger (stronger match) than a portion of the reference image locations. The best match image is within 0.2 metres 28% of the time, and within 0.4 metres 78% of the time. Clearly reliable single image matching is challenging under these conditions.

Figure 13 and Figure 14 show the results when using sequence SLAM methods on the night-time query set. As is shown by Figure 13, sequence SLAM greatly improves the performance. For example, for the 28th query image, the heat map resolves to the correct location in the 5 point sequence SLAM, in contrast to the near homogenous heat map for the single frame match with no clear match.

The following figure, Figure 11, summarizes the results of the place recognition experiments at night-time in terms of the error in estimating/interpolating the location of where the query image was taken. Sequences of 4 images and above achieve 100% matching accuracy within 0.4 metres of the correct location.

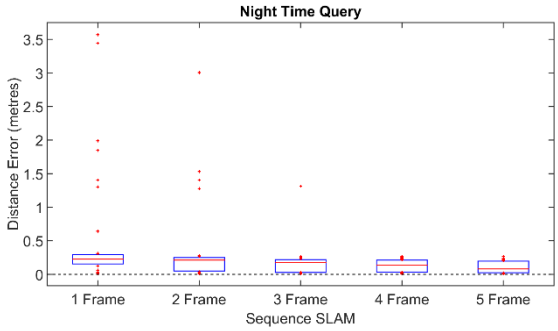

Figure 11: Shows error plots for each of the different size frames for sequence SLAM for the night time query set.

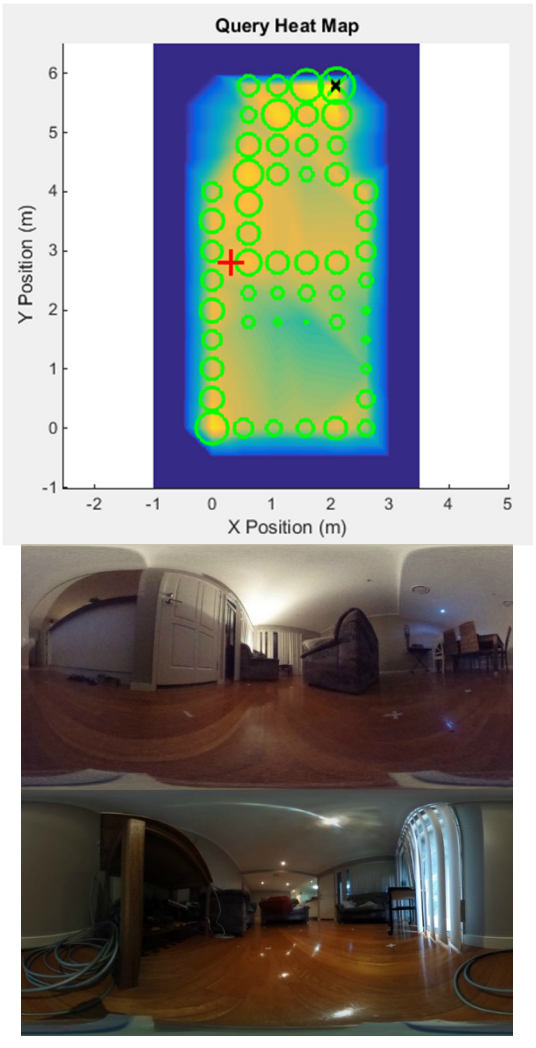

Figure 12: A false positive match using single image matching for the night time query set. It shows the heat map and the image for the 28th night-time query set image (top image), as well as the best matched image (incorrect) from the reference set (bottom image).

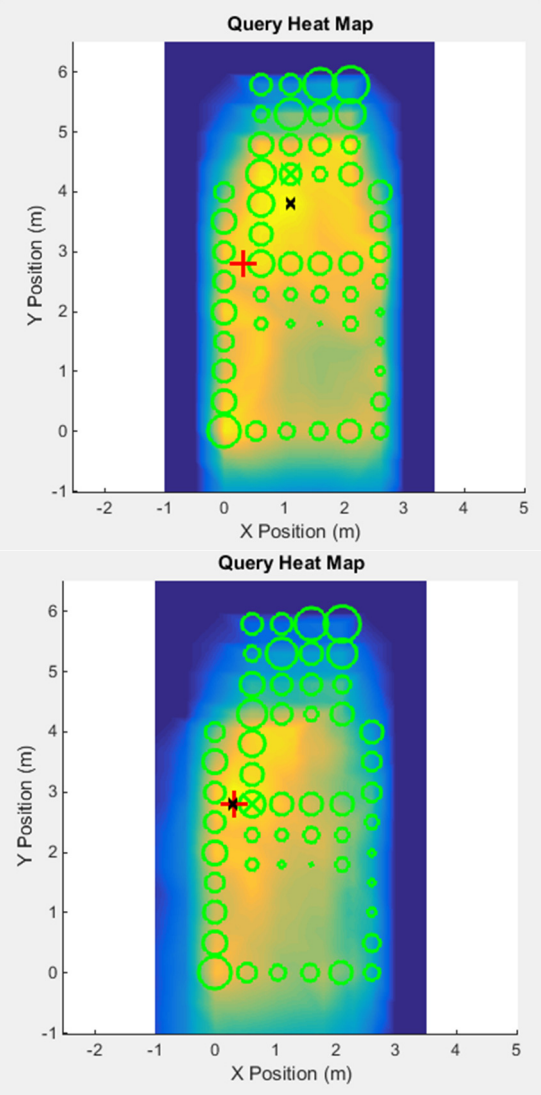

Figure 13: The sequence SLAM heat maps for the 28th night-time query set image. The top heat map is for 3 point sequence SLAM, while the bottom is for 5 point sequence SLAM.

Figure 14: The 28th night-time query set image is correctly matched to the 41st image within the daytime reference when using sequences.

C. *Alternative Low-Light Sensor – DSLR Results*

The Ricoh sensor is a commodity sensor not specifically designed for low light, and hence performance breaks down if a room is nearly pitch black. To provide an indicator of what could be done with a dedicated sensor that trades pixel resolution for larger pixel pitches (wells/receiving area), we provide some illustrative results with a cheap 4 year old Nikon D5100 DSLR camera.

The figures below compare the results of using the Ricoh Theta camera to a digital SLR camera in a completely darkened environment using the same image matching techniques. As can be seen, even though the room is completely dark, except for the light through the air vent, the DSLR is still able to expose most of the environment and perform a successful image match (Figure 15), while the Ricoh Theta produces a nearly pitch black image (Figure 16).

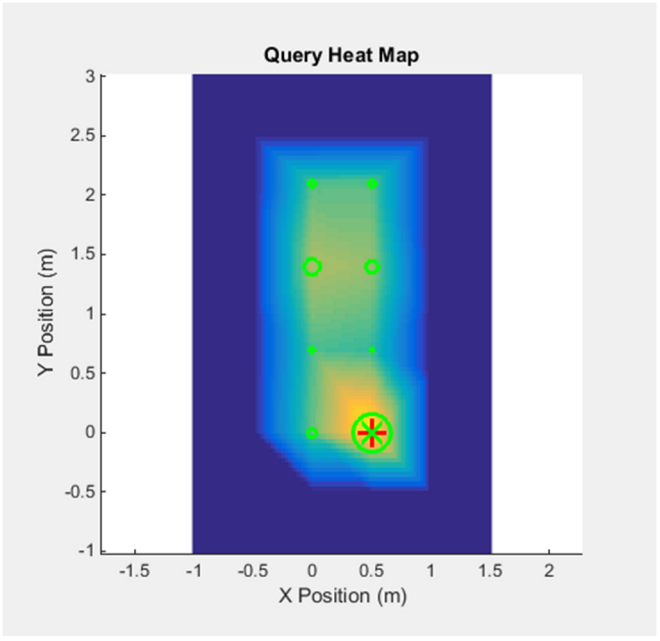

Figure 15: The images taken by the digital SLR camera with the lights on (top image) and the completely dark room (bottom image), as well as the successful image comparison via the single frame heat map.

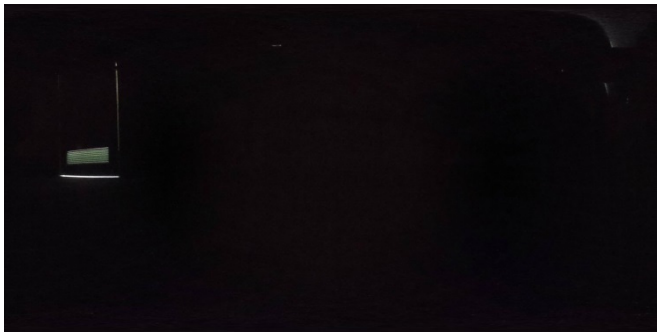

Figure 16: The same location as in the previous figure, captured using the Ricoh Theta camera. As can be seen, little information is contained within the image.

D. Computational Efficiency

The current algorithms are implemented as unoptimized Matlab code. For the datasets presented here, the primary computational overhead is the image comparison process. When comparing a query image to 52 reference images, at a resolution of 48 × 24 pixels at every rotation (48 rotations), we are performing 2,875,392 pixel comparisons for every query image. A CPU can perform approximately 1 billion single byte pixel comparisons per second, while a GPU can do approximately 80 billion per second using optimized C code, hence the techniques presented here could likely be performed in real-time on a robotic platform when optimized, even on lightweight computation hardware.

VI. Discussion and Future Work

In this paper we have investigated the potential of low resolution, sequence-based image matching algorithms for performing localization on domestic robots such as a robot vacuum cleaner in challenging or low light conditions. While single-matching image performance is poor, using short sequences of a few images enables 100% matching accuracy to a reasonable degree of accuracy (0.4 metres). In our current research we are working towards increasing this accuracy by another order of magnitude in order to enable autonomous and accurate robot navigation in the home at any time of day or night.

Future work will pursue this aim in a number of ways. Firstly, extracting some estimate of depth from the image, such as estimating depth from single images using deep learning techniques [Milford et al., 2015] or optical flow, will enable the generation of synthetic images at novel viewpoints, potentially enabling a higher degree of metric localization accuracy. Understanding scene depth will also enable an investigation of the required environmental sampling; how sparse can the reference day-time map be without adversely affecting night-time localization? Finally, from a practical perspective, the next step will be to deploy the system in online, real-time experiments using embedded hardware on a vacuum cleaner robot platform.

References

[A, P., B, C., P, J., & H, C., 2010] A, P., B, C., P, J., & H, C. A realistic benchmark for visual indoor place recognition. *Robotic and Autonomous System, 58*(1), 81-96. (2010)

[Andreasson, H., Duckett, T., & Lilienthal, A., Year] Andreasson, H., Duckett, T., & Lilienthal, A. *Mini-SLAM: Minimalistic Visual SLAM in Large-Scale Environments Based on a New Interpretation of Image Similarity.* Paper presented at the International Conference on Robotics and Automation, Rome, Italy Year.

[Andreasson, H., Duckett, T., & Lilienthal, A., 2008] Andreasson, H., Duckett, T., & Lilienthal, A. A Minimalistic Approach to Appearance-Based Visual SLAM. *IEEE Transactions on Robotics, 24*(5), 1-11. (2008)

[Burschka, D., & Hager, G. D., Year] Burschka, D., & Hager, G. D. *V-GPS (SLAM): Vision-based inertial system for mobile robots* Year.

[Cummins, M., & Newman, P., Year] Cummins, M., & Newman, P. *Highly scalable appearance-only SLAM - FAB-MAP 2.0.* Paper presented at the Robotics: Science and Systems, Seattle, United States Year.

[Davison, A. J., Reid, I. D., Molton, N. D., & Stasse, O., 2007] Davison, A. J., Reid, I. D., Molton, N. D., & Stasse, O. MonoSLAM: Real-Time Single Camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 29*(6), 1052-1067. (2007)

[Dudek, G., & Jugessur, D., Year] Dudek, G., & Jugessur, D. *Robust place recognition using local appearance based methods.* Paper presented at the Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on Year.

[Durrant-Whyte, H., & Bailey, T., 2006] Durrant-Whyte, H., & Bailey, T. Simultaneous localization and mapping: part I. *Robotics & Automation Magazine, IEEE, 13*(2), 99-110. (2006)

[Konolige, K., & Agrawal, M., 2008] Konolige, K., & Agrawal, M. FrameSLAM: From Bundle Adjustment to Real-Time Visual Mapping. *IEEE Transactions on Robotics, 24*(5), 1066-1077. (2008)

[Konolige, K., Agrawal, M., Bolles, R., Cowan, C., Fischler, M., & Gerkey, B., Year] Konolige, K., Agrawal, M., Bolles, R., Cowan, C., Fischler, M., & Gerkey, B. *Outdoor mapping and navigation using stereo vision* Year.

[Milford, M., 2013] Milford, M. Vision-based place recognition: how low can you go? *International Journal of Robotics Research, 32*(7), 766-789. (2013)

[Milford, M., Shen, C., Lowry, S., Suenderhauf, N., Shirazi, S., Lin, G., . . . Upcroft, B., Year] Milford, M., Shen, C., Lowry, S., Suenderhauf, N., Shirazi, S., Lin, G., . . . Upcroft, B. *Sequence Searching With Deep-Learnt Depth for Condition-and Viewpoint-Invariant Route-Based Place Recognition.* Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops Year.

[Milford, M., & Wyeth, G., 2008] Milford, M., & Wyeth, G. Mapping a Suburb with a Single Camera using a Biologically Inspired SLAM System. *IEEE Transactions on Robotics, 24*(5), 1038-1053. (2008)

[Milford, M., & Wyeth, G., Year] Milford, M., & Wyeth, G. *SeqSLAM: Visual Route-Based Navigation for Sunny Summer Days and Stormy Winter Nights.* Paper presented at the IEEE International Conference on Robotics and Automation, St Paul, United States Year.

[Paz, L. M., Pinies, P., Tardos, J. D., & Neira, J., 2008] Paz, L. M., Pinies, P., Tardos, J. D., & Neira, J. Large-Scale 6-DOF SLAM With Stereo-in-Hand. *IEEE Transactions on Robotics, 24*(5), 946-957. (2008)

[Royer, E., Bom, J., Dhome, M., Thuilot, B., Lhuillier, M., & Marmoiton, F., Year] Royer, E., Bom, J., Dhome, M., Thuilot, B., Lhuillier, M., & Marmoiton, F. *Outdoor autonomous navigation using monocular vision.* Paper presented at the IEEE International Conference on Intelligent Robots and Systems Year.

[Vaussard, F., Fink, J., Bauwens, V., Retornaz, P., Hamel, D., Dillenbourg, P., & Mondada, F., Year] Vaussard, F., Fink, J., Bauwens, V., Retornaz, P., Hamel, D., Dillenbourg, P., & Mondada, F. *Lessons learned from robotic vacuum cleaners entering the home ecosystem*, P.O. Box 211, Amsterdam, 1000 AE, Netherlands Year.

[Yamauchi, B., & Langley, P., 1997] Yamauchi, B., & Langley, P. Place recognition in dynamic environments. *Journal of robotic systems, 14*(2), 107-120. (1997)

[Zhang, A. M., & Kleeman, L., 2009] Zhang, A. M., & Kleeman, L. Robust Appearance Based Visual Route Following for Navigation in Large-scale Outdoor Environments. *The International Journal of Robotics Research, 28*(3), 331-356. doi: 10.1177/0278364908098412 (2009)