Robust Appearance Based Visual Route Following in Large Scale Outdoor Environments

Alan M. Zhang and Lindsay Kleeman
Intelligent Robotics Research Centre
Department of Electrical and Computer Systems Engineering
Monash University, Clayton, Victoria 3800, Australia
alan.alanzhang@gmail.com
lindsay.kleeman@eng.monash.edu.au

Abstract

Route following is defined as the ability to repeat the same route after having traversed it once under the control of a human operator. This paper presents a fast and robust appearance based route following algorithm that utilises only monocular panoramic vision and odometry in large scale outdoor environments. Extensive experimental results have shown robustness against lighting variations and occlusion not yet demonstrated in the literature. The cumulative distance traveled by the robot in these experiments exceeds 20km and the longest repeated route is 732 meters.

1 Introduction

This paper presents a robust appearance based solution to the problem of route following in large scale outdoor environments. Route following is defined as the ability to repeat the same route after having traversed it once under external control. A potential area of application is robotic couriers where human operators drive the robot from one destination to another, subsequently with the robot able to repeat the route autonomously. The system targets urban environments during daylight hours. Passive vision has been chosen as the primary sensing modality due to it possessing a number of attractive properties. In comparison to laser range finders, which have finite range, vision is not constrained by distance to features. The appearance of the environment is also richer in information than the geometry of the environment. Vision sensors are small in size, low in power consumption, inexpensive, as well as being passive in nature. The challenge associated with vision however, is mainly two fold: loss of information when a 3D world is projected onto a 2D image plane, and the variation of appearance of visual features depending on lighting conditions. Possible solutions to vision based route following are discussed next.

Building a geometric map of the environment is a possibility. If the robot is capable of building a map of the environment during route teaching and subsequently localise in the map then its offset from the route can be compensated for, hence achieving route following. However, vision based SLAM (Simultaneous Localisation And Mapping) algorithms are still far from mature. A number of additional constrains make the route following problem a member of a superset of problems solved by SLAM. Construction of a globally consistent geometric map is the goal of SLAM. Whereas in route following, provided that a route could be repeated successfully, the form of representation in which information regarding the environment is encoded is irrelevant. Significant deviations from the route is not expected to occur during normal operation. Thus there is no need to predict the appearance of features from different viewing angles. A convergent feedback system only requires the correct *sign* of the lateral offset from the teaching route. Furthermore, ground surface can be assumed to be locally flat, resulting in the vertical orientation of the camera being the same when revisiting the same part of the route. Thus any differences in robot orientation while retracing the route introduce only a shift in the azimuth angle of the visual features. Under these constraints it is proposed that an algorithmically simpler approach would suffice for the route following problem. Many apparently simple algorithms, often biologically inspired, have been remarkably successful at solving navigation problems in robotics [Srinivasan *et al.*, 2004; Chahl and Srinivasan, 1999]. The appearance based approach presented in this paper is algorithmically simple and operates very fast, at the same time achieving robustness against occlusion and lighting variations that has yet to be demonstrated in the literature.

2 Hardware

Figure 1 shows the Pioneer P3-AT outdoor robot that was used as the mobile platform. A webcam directed towards a panoramic mirror of the profile given in [Chahl

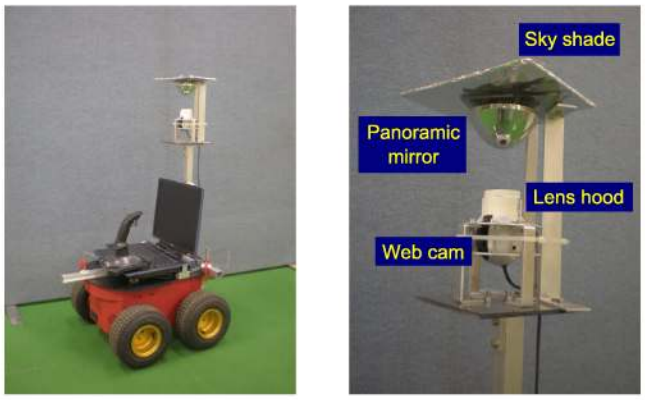


Figure 1: Hardware setup consisting of a Pioneer P3-AT outdoor robot, camera/mirror assembly for panoramic vision, a laptop and a joystick for manual route teaching.

and Srinivasan, 2000] provides panoramic views with a vertical Field Of View (FOV) ranging from $-90°$ to $48°$ in elevation and $360°$ in azimuth (the mirror support beams cause some occlusion). A sky shade blocks out the sky to improve the camera automatic gain control. It is covered in aluminium foil such that its brightness changes with ambient lighting. A lens hood was used to mitigate the problem of lens flares. The camera/mirror assembly is 130cm above the ground.

3 Related Work

As stated earlier, SLAM solves a set of problems of which route following is a subset. A number of pure visual SLAM implementations in the literature could perform route following but most work only indoors. That number reduces drastically when operation in outdoor environments is required. As this paper does not use SLAM techniques, only a few prominent vision only SLAM systems are discussed here to highlight their shortcomings as solutions to route following. SLAM using monocular vision alone has been achieved only recently and some emphasis has been placed on real-time map generation [Davison, 2003]. Since route following does not require online map building Royer [Royer *et al.*, 2005] used a time consuming but more accurate method to build maps offline. However, image feature tracking and association still needs to be done in real-time for localisation. The systems of Davison [Davison, 2003; Royer *et al.*, 2005] and most other geometric mapping based systems use points/corners as visual features. For example, the Harris corner detector was used in [Royer *et al.*, 2005] to detect interest points. Feature association was done using normalised cross-correlation of a neighbourhood of pixels around the interest point. These features are sensitive to the choice of scale at which they are detected. Ideally, the features should be detected for all scales by progressive image subsampling. Point features offer the advantage of being easy to detect and localise. However, features such as the outlines of buildings and vegetation are significant visual features. But these features, especially the outlines of vegetation, are difficult to model geometrically because they change quickly with viewing angle. While visual SLAM using straight lines has been achieved [Smith *et al.*, 2006], the use of free-form curves that are typical of the outlines of trees and bushes has yet to be demonstrated. The detection of more complex features is also more computationally expensive.

The approach presented in this paper is best classified as an appearance based approach to route following. A related area of research is visual homing where a robot is to return to a reference position when it is initialised at a nearby position. This can be achieved with a simple feedback mechanism without geometric reconstruction of the environment. One of the first to address this problem is Collett [Collett, 1992] and subsequently by many others [Weber *et al.*, 1998; Bianco *et al.*, 1997; Argyros *et al.*, 2001]. Foraging insects such as ants and bees appear to use this strategy [Srinivasan, 1998; Judd and Collett, 1998; Trullier *et al.*, 1997]. The basic idea is to take visual snapshots at the home/reference position, find the displacement of features in the current visual scene with the reference snapshot and then move in a direction that reduces the average feature displacement. This basic algorithm and its variants can be proved to be globally convergent provided that the feature associations are correct. The route following problem is then decomposed into a series of reference positions such that visual homing from one reference position to the next is effectively following the route. This is a valid solution but it is less than ideal. The algorithm in [Argyros *et al.*, 2001] is representative of most visual homing methods where image features need to be associated between current and reference images. However, feature association is difficult when occlusions and lighting variations are considered. A more fundamental issue is that the direction of movement at any instant in time can not be proven to be directly towards the reference position, even when the final position is provably convergent. It is therefore difficult to ensure a smooth trajectory while servoing.

The proposed approach is similar to that of Matsumoto [Matsumoto *et al.*, 1999] in the sense that the basic premise for convergence is the same, both use panoramic vision systems and cross-correlation for image matching. However, the approach presented here is robust against occlusion and lighting changes in outdoor environments. In contrast to most existing works in the appearance based route following literature that are of-

ten not accompanied by adequate experimental results, this paper presents a comprehensive set of experiments.

4 Appearance Based Route Following

4.1 System Overview

Route following consists of two phases, route teaching and autonomous route following. The core concept is to capture a sequence of reference panoramic images during route teaching. During the autonomous phase, measurement image is compared against the closest reference image to recover a relative orientation. Only the frontal 180 degrees field of view is used to ensure a convergent behaviour. The robot then corrects its heading to zero the relative orientation. Both lateral offset and orientation error can be corrected with this behaviour. Relative orientation is recovered by first unwarping the image into azimuth-elevation coordinates, followed by a cross-correlation in the azimuth axis and detecting for peaks in the correlation coefficients. Figure 2 shows the overall system architecture. Following sections concentrate on the main components in Figure 2: image pre-processing, image cross-correlation, along route localisation, and relative orientation tracking. Reference image selection is discussed in Section 4.4.

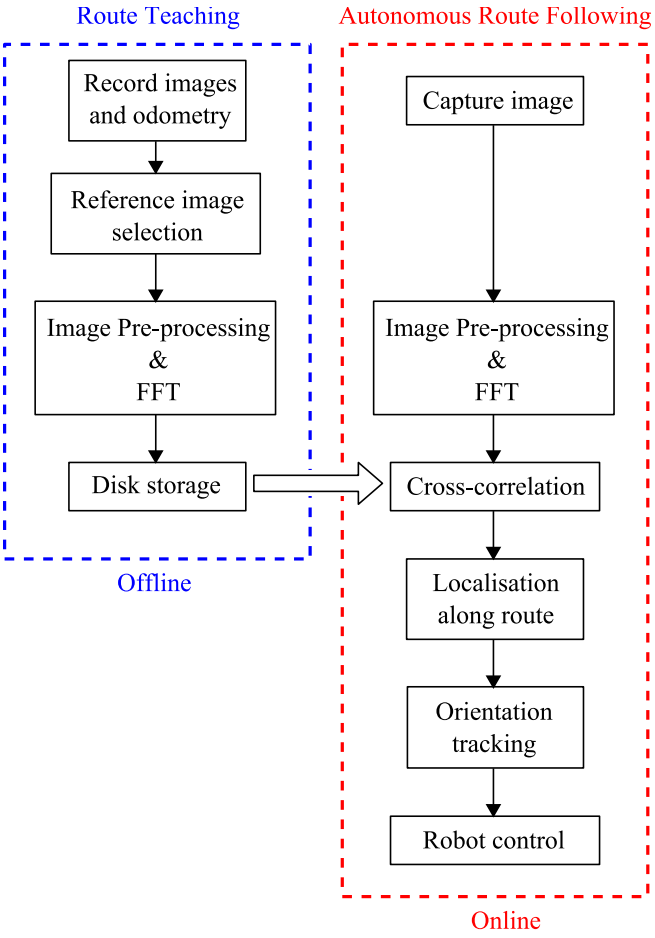


Figure 2: System Overview

4.2 Image Pre-processing

Identical image pre-processing steps are applied to both reference and measurement images. Input colour image is first converted into greyscale (colour information is unstable under changing lighting conditions) then "unwarped" (i.e. remapped) onto azimuth-elevation coordinates. An example of the original colour image and its unwarped greyscale image is shown in Figures 3a and 3b respectively, where horizontal axis is azimuth and vertical axis is elevation. Vertical field of view is restricted to $[-50°, 20°]$. Patch normalisation is then applied to compensate for changes in lighting condition. It transforms the pixel values as follows:

$$I'(x,y) = \left(\left(I(x,y) - \mu(x,y)\right)/\sigma(x,y)\right) \qquad (1)$$

where $I(x,y)$ and $I'(x,y)$ are the original and normalised pixels respectively, $\mu(x,y)$ and $\sigma(x,y)$ are the mean and standard deviation of pixel values in a neighbourhood centred around $I(x,y)$. Figure 3c shows the result of applying patch normalisation to Figure 3b. A neighbourhood size of 17 by 17 pixels has worked well in the experiments.

4.3 Image Cross Correlation

This section addresses the problem of measuring an orientation difference between a measurement image and a reference image. Ground surface along the route is assumed to be locally flat such that the radial axis of the panoramic vision system is perpendicular to the ground plane. Orientation difference between reference and measurement image is therefore only a shift along the azimuth axis. This shift is recovered using Image Cross-Correlation (ICC) performed efficiently in the Fourier domain. Let θ denote azimuth and ϕ elevation. The frontal 180° field of view of the reference image serves as the template, i.e. $\theta \in [-90°, 90°]$. Let the search range be $\pm\theta_{srch}$ such that the measurement image is limited to the angular range $[-90° - \theta_{srch}, 90° + \theta_{srch}]$. Because only a 1D cross-correlation along the azimuth axis is performed, each row in the image is transformed into Fourier domain separately. Reference image is padded with zeros to the same size as the measurement image. If the measurement image is N_ϕ by N_θ pixels, then the Fourier domain image consists of N_ϕ sets of 1D Fourier coefficients, each of a single row. Algorithmic complexity for a single image is $O(N_\phi N_\theta \log N_\theta)$. Convolution in spatial domain is equivalent to multiplication in Fourier

(a) Original image

(b) Unwarped greyscale image

(c) Patched normalised

Figure 3: (a) Original colour image. (b) Converted to greyscale and mapped into azimuth-elevation coordinates, where the azimuth-axis is horizontal. (c) Patch normalised to remove lighting variations, using a neighbourhood of 17 by 17 pixels.

domain:

$$\begin{aligned}
\lambda &= \sum_{i=1}^{N_\phi} R^i * M^i \\
\Rightarrow \mathcal{F}\{\lambda\} &= \mathcal{F}\left\{\sum_{i=1}^{N_\phi} R^i * M^i\right\} \\
&= \sum_{i=1}^{N_\phi} \mathcal{F}\{R^i\} \cdot \mathcal{F}\{M^i\} \\
\Rightarrow \lambda &= \mathcal{F}^{-1}\left\{\sum_{i=0}^{N_\phi - 1} \mathcal{F}\{R^i\} \cdot \mathcal{F}\{M^i\}\right\} \qquad (2)
\end{aligned}$$

where λ is the ICC coefficients, R^i and M^i are the i'th row in the reference and measurement image respectively, $*$ is the convolution operator and $\mathcal{F}\{\bullet\}$ is the Fourier transform operator. Equation 2 states that each corresponding row of the measurement and reference images are multiplied in Fourier domain. The results are then summed followed by an inverse Fourier transform to obtain the spatial domain cross-correlation coefficients. Complexity for the multiplication in Fourier domain is $O(N_\phi N_\theta)$ and for inverse Fourier transform is $O(N_\theta \log N_\theta)$. Fourier transforms for the reference images are calculated offline after the teaching run and stored. The complexity of a complete ICC is thus $O(N_\phi N_\theta \log N_\theta) + O(mN_\phi N_\theta) + O(mN_\theta \log N_\theta)$ where m is the number reference images to compare against. This is significantly better than the complexity of ICC performed in spatial domain which is $O(mN_\phi N_\theta \theta_{srch})$. Comparing against 11 reference images only takes 2.3ms on a 2.4GHz Mobile Pentium 4 per measurement image. More timing results are presented in Section 5.5.

4.4 Along Route Localisation

Existing works [Matsumoto *et al.*, 1999; Jones *et al.*, 1997; Matsumoto *et al.*, 2000] localise the robot to the nearest reference image, where each measurement image is compared to the current nearest reference image and the next one along the route. If the next reference is a better match then the robot is relocalised to this new reference. This method is far too sensitive to errors in image matching. The proposed approach uses a Markov localisation filter [Fox *et al.*, 1999] to track robot position along the route, where the state variable is the distance from the starting point. States are discretised at a uniform resolution of 7cm per state.

Distance between reference images are longer than the distance between localisation states, and also at a variable resolution. Selecting reference images is non-trivial. They should be allocated densely at turns or when visual features are close to the robot. One possible selection method is presented in [Matsumoto *et al.*, 2000]. This paper concentrates on the online algorithm used to repeat the route. Hence a simple selection method was employed where reference images are allocated with a maximum separation of 35cm and 5° rotation according to odometry.

Since the "kidnapped robot problem" is not one of the operation scenarios and that the robot is always initialised at the starting point, there is no need for global localisation. Thus at any instant in time, only a local neighbourhood of localisation states centred at the most likely robot location is considered to reduce computational complexity.

Prediction update involves shifting the probability distribution along the route according to odometry measurement. During observation update, observation likelihood is calculated from ICC coefficients. ICC is performed on the measurement image against 11 reference images centred around the estimated current robot location. Local maxima in the ICC coefficients are detected. For each reference image, the *score* of a single local maxima with an relative orientation value closest to that of the current estimated robot relative orientation determines the matching likelihood of that reference image. The robot's relative orientation and its estimation is discussed in the next section. Because the localisation states are denser than the reference images, each state

is then given a score via linear interpolation of scores of the reference images in front and behind that state. The actual observation likelihood is obtained by first normalising the scores, followed by addition of a constant, and renormalisation. A larger additive constant has the effect of reducing the confidence placed in the observations. This constant was experimentally determined.

4.5 Relative Orientation Tracking

The robot's "relative orientation" refers to the difference between the robot's current orientation and that of the reference image. It is designated θ_δ. Note that a multimodal tracker for this quantity is not necessary since control decisions must be made based on the current best estimate. Route following will diverge if the most probable mode does not track the true orientation.

A Kalman filter is used to track θ_δ. Prediction and correction updates are presented next. Prediction update uses readings from odometry as follows:

$$\theta_\delta(d+\Delta d) = \theta_\delta(d) + \Delta\theta_{msur}(d+\Delta d) - \Delta\theta_{ref}(d+\Delta d) \quad (3)$$

where d is the distance from the start of the route, Δd is the distance traveled since the last update, $\Delta\theta_{msur}(d+\Delta d)$ and $\Delta\theta_{ref}(d+\Delta d)$ are the changes in orientation measured by odometry in the distance interval $[d,\, d+\Delta d]$ during the autonomous and teaching runs respectively. Change in robot orientation in a differentially steered robot can be modeled as:

$$\Delta\theta = \frac{d_R - d_L}{W} \quad (4)$$

where d_L and d_R are the distances traveled by the left and right wheels, W is the effective wheel separation and the variance is:

$$Var(\Delta\theta) = \frac{Var(d_R) + Var(d_L)}{W^2} \quad (5)$$

The variance of a distance measurement is directly proportional to the distance measured:

$$Var(d_R) = \alpha \cdot |d_R| \quad (6)$$

$$Var(d_L) = \alpha \cdot |d_L| \quad (7)$$

where α controls the size of the odometry error.

Observation update uses ICC results from only two reference images. One is immediately in front and the other behind the current robot location. One local maxima from each reference that are nearest to the current predicted θ_δ are selected. A linear interpolation of the two maxima positions controlled by the location of the robot in between the references provide the observation of θ_δ. The observation needs to pass a validation gate set at 95% confidence before being accepted for state update.

4.6 Robot Control

The control algorithm aims to zero the robot's estimated relative orientation using a proportional controller:

$$\omega = -\kappa \cdot \theta_\delta \quad (8)$$

where ω is the robot's rate of rotation and κ is the experimentally determined system gain that depends on the system's processing speed and the robot dynamics.

5 Experimental Results

As a consequence of the proposed approach being behaviour based, any improvement to the algorithm needs to be tested online. The development process consists of conducting experiments which expose failure scenarios, improving the algorithm, experimental validation, and more experiments that expose further problems. Only the latest version of the algorithm has been described in Section 4. The effect of any improvement to the algorithm can only truly be verified with online experiments, under as many different lighting conditions as possible. It is clearly infeasible to repeat all the experiments after any improvement was made. So the procedure was to first apply the improved algorithm to the recorded experiments that exhibited a specific failure scenario to verify that a better behaviour is achieved. The improved algorithm is then applied to recordings of other successful experiments to ensure that the behaviour does not deviate drastically from the old algorithm. This offline verification process significantly increases the confidence in the improvement not introducing other failure modes. Although some experiments presented in this section used earlier versions of the algorithm, they nevertheless validate various aspects of the final algorithm. Differences in the algorithms will be explicitly stated for each set of experiments. The cumulative length of the experiments conducted is over 20km.

5.1 Robustness of Relative Orientation Measurement

This set of experiments demonstrates the robustness of the image cross-correlation method for relative orientation measurement against occlusion and lighting variations in outdoor environments. In order to isolate the causes of failures, prediction update based on odometry information in orientation tracking was ignored, making the system more sensitive to errors in relative orientation measurement between images.

Figure 4 shows the teaching route in our university campus. The green circle and red square are the start and end points. The route is 416 meters in length. A total of eight successful runs were made under differing lighting conditions. Cumulative odometry for these experiments is shown together in Figure 5. Odometry

of the teaching run is shown in black. The routes' end points are indicated with asterisks. Errors of the end points are in orders of tens of meters.

The experiments were performed during summer months with the presence of very strong sun light. All images in Figure 6 were captured from the same position. Figure 6a is the reference image, Figures 6b, 6c, 6d, 6e and 6f were at different times of sunny days, and 6g was on an overcast day. These images illustrate significant changes in appearance typically associated with changing lighting conditions in outdoor environments. The patch-normalised unwarped images of Figure 6 together with their ICC results against the reference image are shown in Figure 7. Units of x-axis are in degrees where positive values correspond to the reference image shifting to the right relative to the measurement image. Global maxima are clearly visible near 0° in every set of ICC coefficients.

Figure 8 shows sections of the autonomous runs containing large amounts of occlusion by pedestrians. The figures show the reference images (top image, showing only the frontal 180°), measurement images subjected occlusion and the ICC coefficients. Lighting changes are also quite significant as can be seen in Figure 8d, where the outline of the building in the right half of the image is obscured by saturation. Yet cross-correlation is robust enough to recover the orientation in each case.

Figure 4: Route of experiments in Section 5.1

5.2 Experiments Under Various Lighting Conditions

Algorithm robustness against lighting changes is tested more exhaustively here than the previous set of experiments. The 322 meters long teaching route is shown in Figure 9. Figures 10a and 10b show changes in lighting conditions at two locations along the route. Table 1 summarises the 23 successful experiments conducted over a period of 9 days with each at a different time of day or weather condition. Sunny days were given greater sampling density because they generate a much greater variety of visual appearances as the position of the sun changes with time.

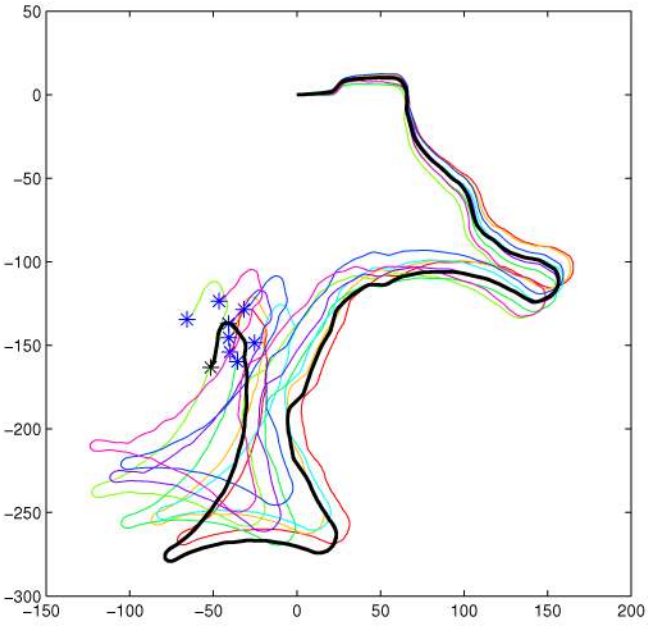


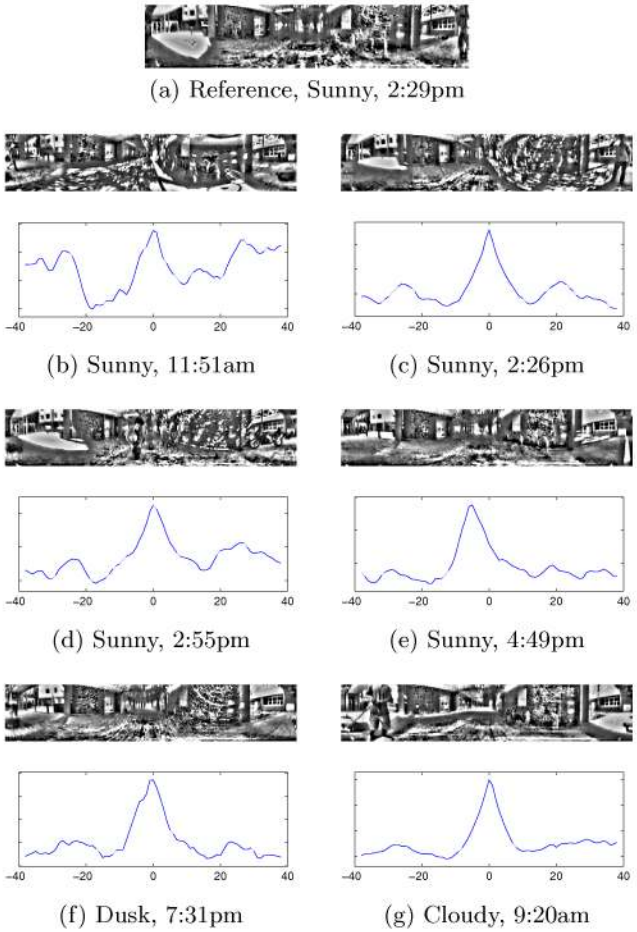
Figure 5: Cumulative odometry of experiments in Section 5.1. Units are in meters. Thicker black line is from the teaching run.

(a) Reference, Sunny, 2:29pm

(b) Sunny, 11:51am (c) Sunny, 2:26pm (d) Sunny, 2:55pm

(e) Sunny, 4:49pm (f) Dusk, 7:31pm (g) Cloudy, 9:20am

Figure 6: Variation in lighting conditions. All images were captured at the same position.



(a) Reference, Sunny, 2:29pm

(b) Sunny, 11:51am (c) Sunny, 2:26pm

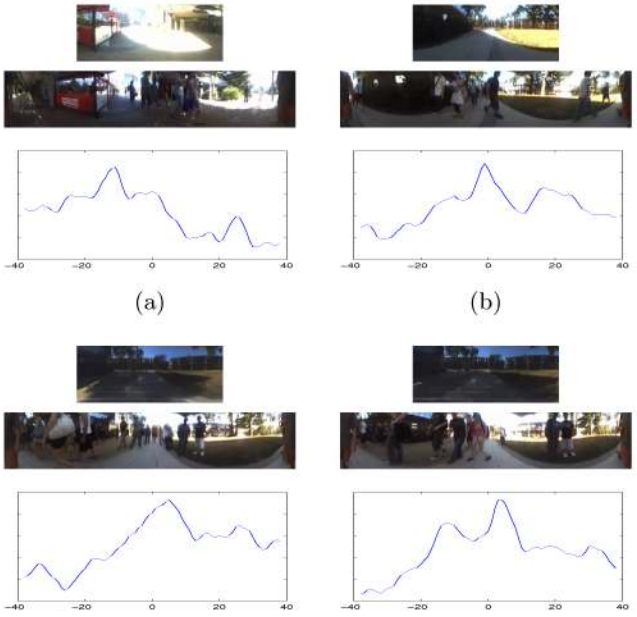
(d) Sunny, 2:55pm (e) Sunny, 4:49pm

(f) Dusk, 7:31pm (g) Cloudy, 9:20am

Figure 7: Demonstrating the robustness of image cross-correlation against lighting variations. Top most image is the reference. Global maxima are clearly present in the ICC coefficients despite the dramatic changes in lighting conditions. Units of x-axis are in degrees. Positive values in x-axis correspond to the reference image shifting to the right relative to the measurement image.

Exp.	Time	Weather	Exp.	Time	Weather
1	9:14	Sunny	13	13:13	Sunny
2	9:24	Sunny	14	13:21	Partly Sunny
3	10:11	Sunny	15	13:44	Sunny
4	10:20	Partly Sunny	16	13:57	Overcast
5	10:38	Overcast	17	14:54	Sunny
6	10:43	Sunny	18	14:55	Sunny
7	11:17	Sunny	19	15:03	Sunny
8	11:47	Sunny	20	15:21	Sunny
9	12:13	Sunny	21	16:47	Dusk
10	12:40	Sunny	22	17:34	Dusk
11	12:49	Overcast	23	17:57	Dusk
12	13:11	Sunny			

Table 1: Weather conditions and time of day during which experiments were conducted for comprehensive testing of robustness against lighting variations.



(a) (b)

(c) (d)

Figure 8: Demonstrating the robustness of image cross-correlation against occlusion. Images from the top: frontal 180° of reference image; measurement image; ICC coefficients. Global maxima in ICC coefficients correctly recover the relative orientation between reference and measurement images. Units of x-axis are in degrees. Positive value in x-axis correspond to the reference image shifting to the right relative to the measurement image.

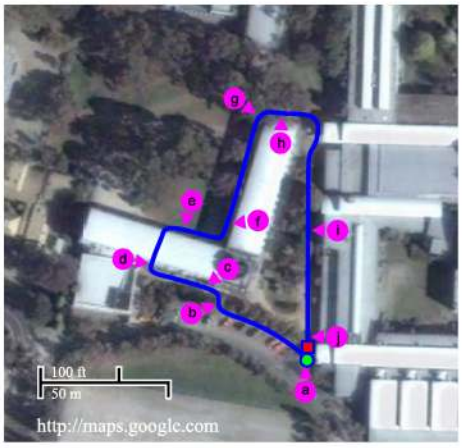


Figure 9: Route used in Section 5.2 for testing robustness against changes in lighting conditions.

(a) (b)

Figure 10: Reference images are in the top row, the rest are from autonomous runs carried out on sunny days at 9:14, 10:20, 11:17, 12:13, 13:13, 15:03, 16:47 and 17:34.

5.3 Ground Truth

The ground truth experiments provide quantitative measures on the accuracy of the algorithm. Ground truth measurements are difficult to obtain in this case because the experiments should be conducted in realistic environments. Real-time kinematic GPS [Royer *et al.*, 2005] provides enough resolution but is inapplicable due to buildings occluding GPS signals. Accurate laser range finders used in land surveying were also considered. But line-of-sight is difficult to maintain. The chosen method is to mark the robot position at waypoints along the route then manually measure the deviation at these way points during autonomous runs.

The route followed is the same as in the previous set of experiments. A new teaching run was made to allow for ground truth measurements during which the robot was made to stop at regular intervals along the route. Robot positions at these waypoints were marked on the ground. During the autonomous phase, robot stopped at these waypoints such that positional errors with respect to ground truth could be measured. A total of 22 waypoints were used. No waypoints were allocated in the middle of turns because stopping the robot during the autonomous phase changes the robot dynamics, hence interfering with the robot control algorithm and could potentially decrease the accuracy of route following. A total of 9 experiments were conducted. Table 2 summarises the results. The direction of the x-axis is perpendicular to the route. Because the localisation states are discretised at 7cm between states, this limits the accuracy along the y-axis. Lateral deviation (x-axis) is more representative of the algorithm's accuracy. Reference sequence was captured at 12:37pm on a sunny day. The fact that Table 2 shows no clear correlation between the amount of deviation and the weather conditions or time of day demonstrates robustness against changing lighting conditions.

Exp. No.	Condition	$\lvert x\rvert$ (cm) Max.	Avg.	σ	$\lvert y\rvert$ (cm) Max.	Avg.	σ
1	16:20 OC	10.0	4.6	4.8	24.0	8.4	10.4
2	16:46 OC	11.0	5.1	5.5	24.0	9.3	11.2
3	11:20 OC	7.0	4.5	4.0	22.0	7.3	9.0
4	12:46 S	6.0	3.4	3.0	20.5	8.5	10.1
5	14:28 S	14.0	5.8	5.0	26.0	8.2	9.9
6	15:22 S	13.0	4.9	4.8	25.5	7.9	9.4
7	15:37 S	11.0	4.6	3.7	23.5	8.4	10.1
8	9:38 S	21.5	8.4	7.4	25.5	9.3	10.9
9	10:27 S	20.0	8.1	6.7	16.0	9.1	10.1

OC - Overcast, S - Sunny, σ - Standard Deviation

Table 2: Error during autonomous route following compared with training route. The $|x|$ column is the magnitude of the lateral offset from the route at the waypoints and $|y|$ is localisation error along the route.

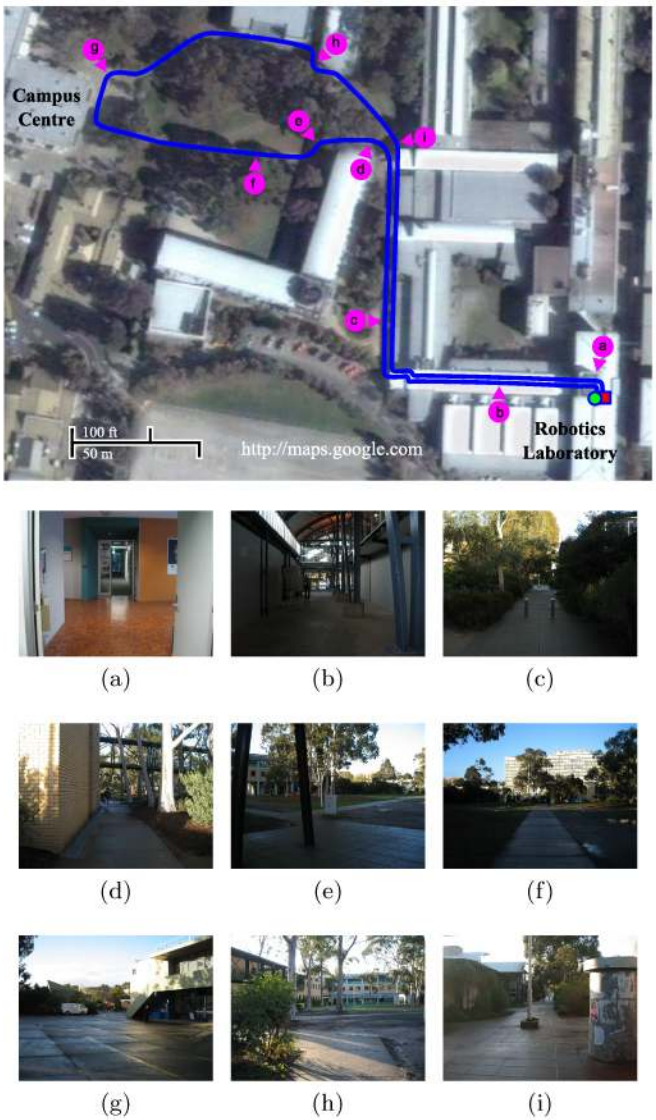
5.4 A Longer Experiment

The route is visualised in Figure 11. At 732 meters in length this is the longest route. The robot travels through a variety of environments including big open spaces (Figure 11f), footpaths predominantly surrounded by vegetation (Figure 11c), and non-uniformly distributed features (Figure 11d where a building facade is close by on one side and open space on the other). A total of 7 experiments were conducted, of which 5 were completely successful. The rest two experiments completed 730 meters but failed at 2 meters before the end of the route just before entering the robotics laboratory. One of these failures was caused by the slow response of the camera automatic gain control leading to the interior of the laboratory appearing completely saturated. The other failure is suspected to be caused by clutter in the laboratory. Again the experiments were conducted under various lighting conditions. Tolerance to occlusion is demonstrated in Figure 12 where a market was setup in the square in front of the campus centre during an autonomous run.

5.5 Timing and Storage Requirement

Below is a summary of system parameters and timing results:

- **Reference image size:** 70 x 180 pixels (= 70° elevation FOV, 180° azimuth FOV)
- **Measurement image size:** 70 x 256 pixels (i.e. ±38° azimuth search range)
- **No. of ref. images to compare against for along route localisation:** 11

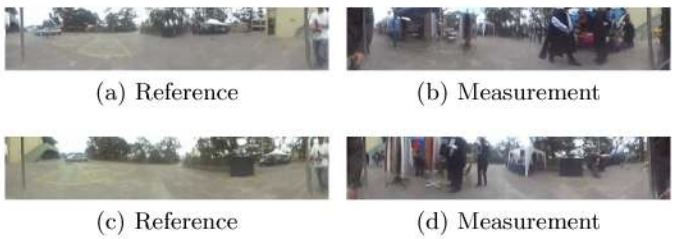


(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 11: Experiment set 2 where robot acts as courier going from the robotics laboratory to the university campus centre and back again.



(a) Reference (b) Measurement

(c) Reference (d) Measurement

Figure 12: A market was set up in the square that was not present in the teaching run.

- **Computer configuration:** 2.4GHz Mobile Pentium 4, 500MB memory
- **FFT library:** FFTW[1] version 3.1.2
- **Image preprocessing:** 7.2ms per image, including conversion to greyscale, unwarping and patch normalisation
- **Image cross-correlation per measurement image:** 2.3ms, including FFT of measurement image and comparing against 11 reference images

Because the reference images are pre-processed and stored as Fourier coefficients, the overhead is only disk access. So pre-processing plus ICC takes only 9.5ms. There is still much optimisation possible with image preprocessing. Offline processing runs at 60fps. During the autonomous phase, with every measurement image encoded into a video, online processing runs at 4fps. At this processing rate the robot is able to travel at 60cm/s in straight sections and 45cm/s at sharp turns. All experiments were conducted at 4fps with logging of every measurement image.

With reference images being 70 x 180 pixels and allocated at 35cm intervals, a 1km route requires 36MB of storage space. This requirement is considered low by contemporary storage capacities.

6 Conclusion

An appearance based route following system has been presented in this paper. It uses a much simpler algorithm compared to existing visual SLAM methods, made possible by exploiting constraints specific to the route following problem. The resulting system has shown a high level of robustness against occlusion and lighting variations in an extensive set of experiments. The longest route was 732 meters in length and was successfully repeated in 7 experiments.

References

[Argyros *et al.*, 2001] Antonis Argyros, Kostas Bekris, and Stelios Orphanoudakis. Robot homing based on corner tracking in a sequence of panoramic images. In *2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001.

[Bianco *et al.*, 1997] G. Bianco, R. Cassinis, A. Rizzi, N. Adami, and P. Mosna. A bee-inspired robot visual homing method. In *Second EUROMICRO workshop on Advanced Mobile Robots*, pages 141 – 146, 22-24 Oct. 1997.

[Chahl and Srinivasan, 1999] J. S. Chahl and M. V. Srinivasan. Panoramic vision system for imaging,

[1]Distributed under the GNU General Public License, http://www.fftw.org

ranging and navigation in three dimensions. In *Proceedings of the International Conference on Field and Service Robotics FSA'99*, Pennsylvania, USA, 1999.

[Chahl and Srinivasan, 2000] J. S. Chahl and M.V. Srinivasan. A complete panoramic vision system, incorporating imaging, ranging, and three dimensional navigation. In *IEEE Workshop on Omnidirectional Vision*, 2000.

[Collett, 1992] T. S. Collett. Landmark learning and guidance in insects. *Royal Society of London Philosophical Transactions Series B*, 337:295–303, September 1992.

[Davison, 2003] Andrew J. Davison. Real-time simultaneous localisation and mapping with a single camera. In *Internation Conference on Computer Vision, Nice, October 13 - 16*, October 2003.

[Fox *et al.*, 1999] Dieter Fox, Wolfram Burgard, and Sebastian Thrun. Markov localization for mobile robots in dynamic environments. *Journal of Artificial Intelligence Research*, 11:391–427, 1999.

[Jones *et al.*, 1997] Stephen D. Jones, Claus Andresen, and James L. Crowley. Appearance based processes for visual navigation. In *1997 IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 2, pages 551–557, 1997.

[Judd and Collett, 1998] S. P. D. Judd and T. S. Collett. Multiple stored views and landmark guidance in ants. *NATURE*, 392:710–714, 1998.

[Matsumoto *et al.*, 1999] Yoshio Matsumoto, Kazunori Ikeda, Masayuki Inaba, and Hirochika Inoue. Visual navigation using omnidirectional view sequence. In *1999 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1999.

[Matsumoto *et al.*, 2000] Yoshio Matsumoto, Katsuhiro Sakai, Masayuki Inaba, and Hirochika Inoue. View-based approach to robot navigation. In *2000 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2000.

[Royer *et al.*, 2005] E. Royer, J. Bom, M. Dhome, B. Thuilot, M. Lhuillier, and F F. Marmoiton. Outdoor autonomous navigation using monocular vision. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, (IROS 2005)*, pages 1253– 1258, 2005.

[Smith *et al.*, 2006] Paul Smith, Ian Reid, and Andrew Davison. Real-time monocular slam with straight lines. In *British Machine Vision Conference*, volume 1, pages 17–26, September 2006.

[Srinivasan *et al.*, 2004] M. V. Srinivasan, S. W. Zhang, J. S. Chahl, G. Stange, and M. Garratt. An overview of insect-inspired guidance for application in ground and airborne platforms. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 218(6):375–388, 2004.

[Srinivasan, 1998] Mandyam V. Srinivasan. Animal navigation: Ants match as they march. *NATURE*, 392:660–661, apr 1998.

[Trullier *et al.*, 1997] Olivier Trullier, Sidney I. Wiener, Alain Berthoz, and Jean-Arcady Meyer. Biologically-based artificial navigation systems: Review and prospects. *Progress in Neurobiology*, 51(5):483–544, Apr 1997.

[Weber *et al.*, 1998] K. Weber, S. Venkatesh, and M. V. Srinivasan. An insect-based approach to robotic homing. In *Fourteenth International Conference on Pattern Recognition*, volume 1, pages 297 – 299, 16-20 Aug 1998.