Localisation using Automatically Selected Landmarks from Panoramic Images

Simon Thompson[†], **Toshihiro Matsui**[‡] and **Alexander Zelinsky**[†]

Intelligent Systems Division[‡]
Electro-Technical Laboratory
1-1-4 Umezono, Tsukuba, Ibaraki 305 Japan

Robotic Systems Laboratory[†]
Department of Systems Engineering, RSISE
The Australian National University
Canberra, ACT 0200 Australia
[simon|alex]@syseng.anu.edu.au, matsui@etl.go.jp

Abstract

The use of visual landmarks for robot localisation is a promising field. It is apparent that the success of localisation by visual landmarks depends on the landmarks chosen. Good landmarks are those which remain reliable over time and through changes in position and orientation. This paper describes a system which learns places by automatically selecting landmarks from panoramic images and uses them for localisation tasks. An adaption of the biologically inspired Turn Back and Look behaviour is used to evaluate potential landmarks. Normalised correlation is used to overcome the affects of changes in illumination in the environment. Results from real robot experiments are reported, showing successful localisation from up to one meter away from the learnt position.

1 Introduction

Visual localisation is one of the key problems in making successful autonomous robots. Vision as a sensor is the richest source of information about a mobile agent's environment and as such contains information vital to solving localisation problems. One limit to visual localisation is the narrow field of view offered by normal monocular camera systems. Obviously the greater the field of view of the camera the more information about the environment can be extracted through visual processing.

Figure 1: A panoramic image

Another problem is that vision provides us with too much information. In fact one of the main tasks of vision science is to work out which parts of the information is necessary for the given job and to discard or put aside that which is not.

1.1 Panoramic Imaging for Localisation

Panoramic imaging can solve the first problem mentioned above [Yagi *et al.*, 1994]. By the use of convex mirror shapes greater fields of view can be achieved. Panoramic mirrors give the full 360 degrees of horizontal visual field and can get over 140 degrees in the vertical field as well as shown in figure 1. This increase in the visual field comes at the cost of resolution in some parts of the picture however, as it must be captured with a camera with a normal field of view.

With the full 360 degree horizontal field of view there is more information from which to achieve the navigation tasks of localisation, path planning and path traversal. [Matsumoto *et al.*, 1997] use panoramic pictures to navigate a robot down corridors using a view sequence approach. The robot memorises a sequence of panoramic images along a route, acquiring new images once the current view drops below a set correlation threshold with the last image. This works well in the corridor environment, with the environment only changing sufficiently to warrant a new view being memorised every 1-2 meters. In a non-regular environment however the rate of acquisition would increase dramatically as the panoramic image would change dramatically over small distances.

1.2 Natural Landmarks for Navigation

An attractive solution to this problem is the use of landmarks for localisation tasks. Using landmarks in localisation requires only parts of an environment to be stored and can contribute to accurate localisation with the aid of triangulation techniques. The use of landmarks would overcome the acquisition rate problem by not requiring the entire visual field to remain constant, therefore maximising the area in which one set of landmarks can be used for localisation.

Obviously the success of localisation using landmarks depends on the choice of landmarks [Thrun, 1996]. [Bianco and Zelinsky, 1999] describe a monocular system where landmarks are chosen on the basis of their reliability. To be selected, landmarks must display uniqueness in the immediate surroundings and the ability to remain reliable as the robot moves through the environment. To this end the system monitors the reliability of landmarks through a series of small arcs. This 'turn back and look' behaviour is inspired from observations of honey bee flight patterns [Collet and Zeil, 1996] [Lehrer, 1993].

The applicability of panoramic camera systems to visual localisation using landmarks has been investigated [Thompson *et al.*, 1999]. In light of the increased field of view available to the robot, both the landmark selection strategy and the localisation procedures need to be adjusted. Like Bianco and Zelinsky, Thompson *et al.* describe systems which have the goals of:

1. Learning places in an environment associating areas with sets of appropriate visual landmarks.
2. Localising mobile robots by locating sets of landmarks which denote a particular place.
3. To ultimately use these landmarks to home to places, and to navigate between places.

This paper details the implementation of such a system and reports on some initial results from localisation experiments on a real robot.

2 The System

2.1 System Setup

This system is implemented on a Nomad200 mobile robot platform from Nomadic Technologies Inc. A spherical mirror is mounted on an upwards pointing camera on top of the robot to achieve panoramic vision. Images are captures via a ACVC capture card and unwarped in software to produce a 320×120 greyscale image. All image processing is done in software, taking advantage of Intel's MMX Technology to speed correlation computations.

2.2 Learning a Place

The process of landmark selection is aimed at increasing the localisation and navigation abilities of a mobile robot in subsequent exposure to the environment. This means that landmarks must be reliable, strongly identifiable, they must be distributed throughout the image to minimise the error in navigation calculations. They also must be able to withstand distortions due to temporal and translational distortions. To this end Bianco and Zelinsky [1999] selects a number of landmarks from different sections of the static environment based on their reliability. This reliability is then tested in a dynamic environment by moving the robot about the original position.

Static Reliability of Landmarks

The static reliability of a landmark in Bianco and Zelinsky's [1999] model, is determined by the uniqueness of the landmark within its local region of the image. Local uniqueness is defined as the degree to which the landmark template differs from the area of the image immediately surrounding the landmark. This approach is based on the 'The Valley Method' proposed by Mori et al [Mori *et al.*, 1995] to generate attention tokens in a scene, which in turn appears to be an instance of a Moravec interest operator [Moravec, 1977] applied to feature tracking. Bianco and Zelinsky adapt this method for the present task of automatic selection of landmarks as shown in figure 2. The figure shows correlation results obtained by matching a 16×16 template on a 32×32 search window centered on the original template. The valley in the image is the template matched with itself, thus having a high correlation (low distortion). Comparing the minimum matching distortion resulting from the match of the landmark and the sixteen surrounding values, shown by the grey square, with the value at the bottom of the valley, the local uniqueness of the landmark is found. More formally:

$$r = 1 - g/g'$$

where r is the reliability of the landmark, g is the distortion of the landmark matched with itself, and g' is the minimum matching distortion from the surrounding circle of pixels. Given that g should only result in distortion due to noise, then the higher the distortion of the minimum of the surrounding templates, the steeper

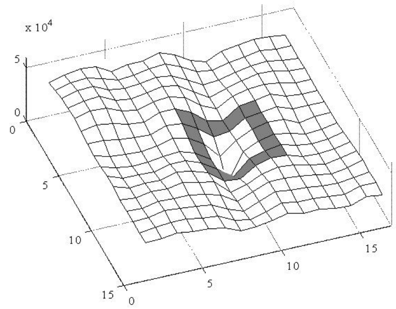

Figure 2: An example of the valley in the distortion matrix caused by a reliable landmark. Figure from Bianco and Zelinksy 1999

the valley in the distortion matrix and subsequently the more unique the local template should be (see figure 2).

The panoramic image is divided up into 4 sectors corresponding roughly to forward, back, left and right of the robot, and the best four landmarks in each sector are selected. By dividing up the visual field in such a way an equal distribution of landmarks can be ensured. The sixteen best landmarks are then evaluated on their dynamic reliability.

Dynamic Reliability of Landmarks

To be good navigational cues, landmarks need to be reliable in a changing environment. They need to remain strongly identifiable under small changes in lighting and shifts in perspective. To this end Bianco and Zelinsky test the reliability of the statically selected landmarks throughout a series of small arcs. This Turn Back and Look (TBL) phase is inspired by observations of wasps and bees on flights away from the hive [Collet and Zeil, 1996] [Lehrer, 1993]. Figure 3 shows the path taken by Bianco and Zelinsky's [1999] robot on the TBL phase. The robot moves further away from the original, or goal, position while keeping the camera oriented towards it. In a panoramic scene containing potential landmarks in any direction, a TBL phase is needed which tests the reliability of the landmarks throughout movements in all directions. At this stage a basic TBL movement is implemented in the shape of a cross as shown in figure 4. This covers movement in two planes and it is yet to be determined if a more elaborate movement would have an effect on the localisation performance.

Potential landmarks are tracked along this path and their reliability measures are evaluated at approximately 400 steps along the way. Each landmark's dynamic reliability is given by the average of the reliability measures taken over the entire path. The landmarks with the highest dynamic reliability measure from each sector are chosen to represent the place.

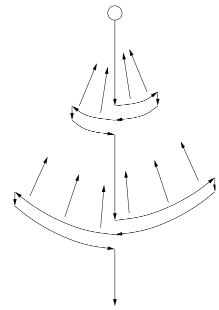

Figure 3: Bianco and Zelinsky's TBL phase for dynamic landmark selection

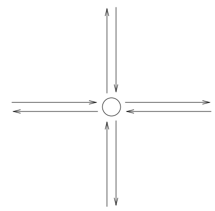

Figure 4: Modified TBL phase for dynamic landmark selection with panoramic vision

2.3 Localisation

Localisation is simply a matter of matching sets of landmarks, and their associated place, to the current visual scene. A brute force search of landmarks throughout the entire image is undertaken for each set of landmarks. The robot is assumed to be in the place associated with the set of landmarks which have the highest average correlation in the current scene.

3 Localisation Experiments

Localisation experiments were carried out to determine the performance of the system under a variety of conditions described below. All experiments took place in the Electro-Technical Laboratory, Intelligent Systems Division. The environment is a semi-structured corridor about 2.5 meters in width and over 15 meters in length. All experiments consisted of an initial phase of learning a place by guiding the robot to the desired place and initiating the automatic landmark selection software. A subsequent phase of guiding the robot to several positions and attempting to locate the learnt landmarks gave the system's localisation performance results. The results are correlation values from 0 to 1, with 1 representing perfect correlation. The desired localisation performance is to achieve high correlation results over the largest possible area, while still being able to discriminate between seperate places. By having landmarks which can cover a wide area, the system can cut back on the amount of storage and processing time needed for localisation.

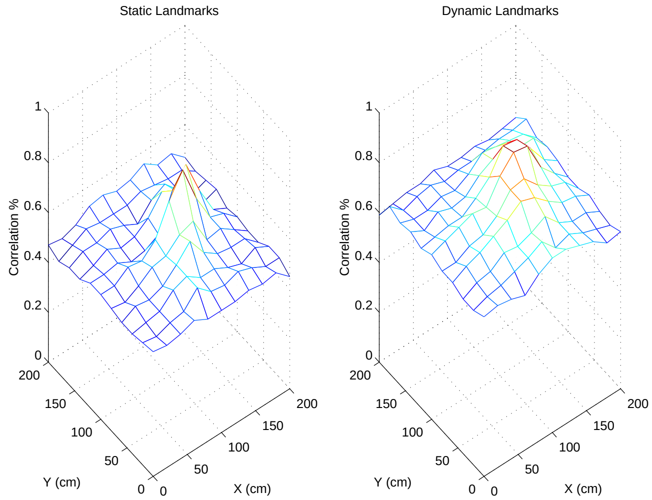

Figure 5: Localisation performance of Statically vs Dynamically selected landmarks

High correlation within that area can allow more accurate positioning by landmark triangualtion and can lead to navigation behaviours such as homing and moving between places.

3.1 Static Landmarks vs Dynamic Landmarks

First to establish the usefulness of the Turn Back and Look phase in automatic landmark selection, a comparision of localisation performance for a given place was made using landmarks selected by either soley static landmark selection or by both static and dynamic selection. In the place learning phase the most reliable static landmarks are stored in addition to the best dynamic landmarks. Both were used to localise within a 2×2 meter section of the corridor environment, centered on the place that was learnt. Measurements were taken at 20cm intervals.

The results of using static landmarks for localisation are shown in the first plot of figure 5. A sharp peak is evident near the center, peaking at 0.79, but falling to around 0.50, just 40cm from the peak, and maintaining this to the edges of the graph. The second plot of figure 5 shows the results when using dynamic landmarks. Again there is a peak near the center (0.86) but it is not nearly so sharp and drops less rapidly. At about 60cm from the peak with correlation values around 0.70, the slope of the graph decreases further and eventually falls to approximately 0.63 at the edges of the graph.

Comparing the two graphs shows that the use of dynamic landmarks for localisation results in higher correlation measures over a greater area around the learnt place than when using static landmarks. The two different slopes observed in the dynamic graph can be attributed to the higher distortion of landmarks located on the sides of the corridor (closer to robot), when compared to those at the end of the corridor (further away).

Figure 6: Sample panoramic image captured at 15:00

Figure 7: Sample panoramic image captured at 20:00

Obviously the distances measured were not enough to say for certain when the graph of the dynamic landmark localisation would approach the 0.50 value. This should be observable in the results from the place discrimination experiments.

3.2 Localisation under Illumination Changes

Next the affect of changes of illumination on the system's performance was investigated. A place was learnt at 15:00, and the selected landmarks stored. The localisation phase was carried out immediately after learning and again at 20:00 that evening, using the same set of learnt landmarks. Again the results given are for a 2×2 meter section of the corridor centered on the learnt place, and the measurements taken at 20cm intervals. Sample images from the robot during the 15:00 run and the 20:00 run are shown in figures 6 and 7 respectively to demonstrate the variance in illumination between trials.

Figure 8 shows the results from the localisation experiments conducted with differing levels of illumination. Both peak at the same position in the graph with similar values (0.86, 0.87 respectively) and both follow the same two step slope described in the previous experiment, with values of around 0.63 at the edges of the graph.

Figure 9 shows a good example of the robustness of the normalised template matching routines. Landmarks are tracked successfully despite rapidly diminishing illumination.

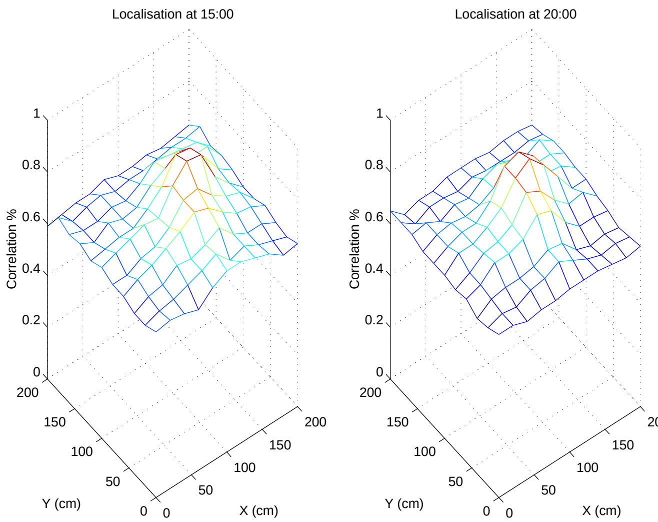

Figure 8: Localisation performance with differing light levels: tests performed at 13:00 and 20:00

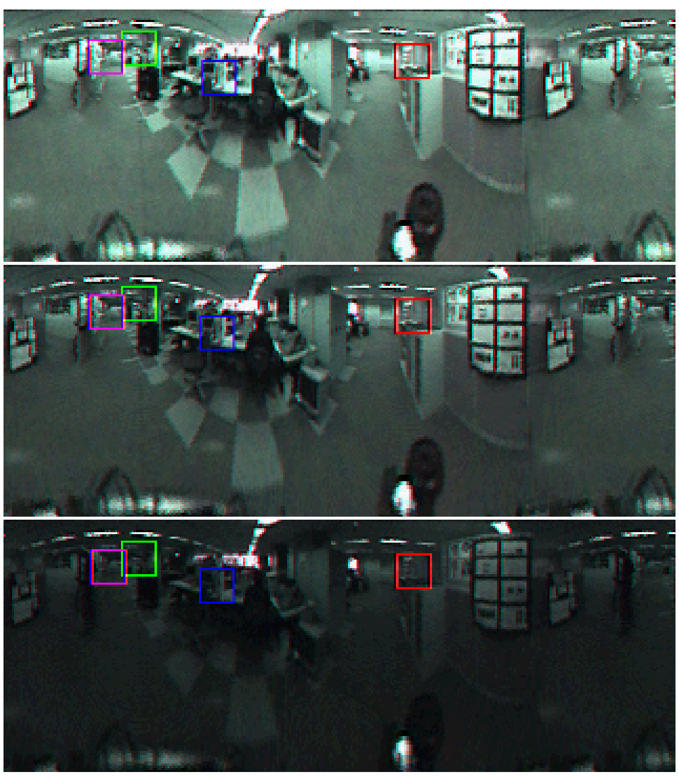

Figure 9: Landmark tracking under changing illumination

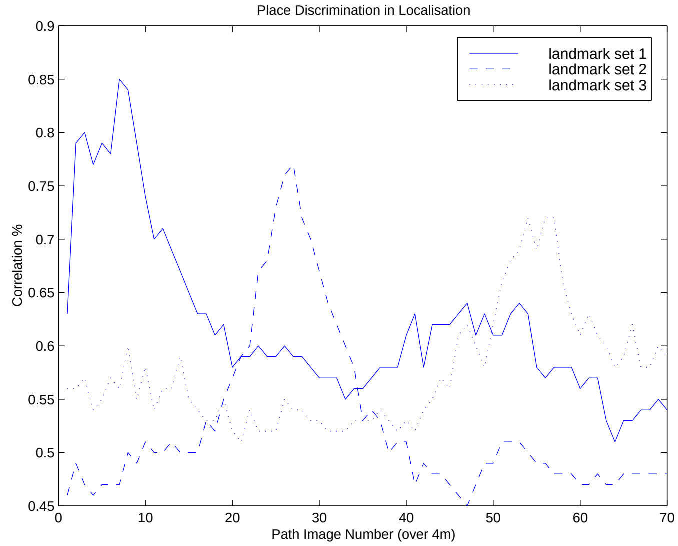

Figure 10: Place discrimination using learnt landmarks. 3 sets of landmarks associated with 3 learnt places are used to localise a robot along a 4 meter path. Places are learnt at the positions associated with image numbers 8 (landmark set 1) , 27 (set 2) and 58 (set 3).

3.3 Place Discrimination

Finally the ability of the system to correctly discriminate between places was tested. Three places were learnt at sites roughly 1.5 meters apart down the center of the corridor. Localisation measures were then taken continuously for each of the learnt places as the robot moved in a straight path of 4 meters in length down the middle of the corridor. The correlation results for each set of landmarks (and their associated place) are given in figure 10. As the robot progresses down the path the landmark set associated with the nearest learnt place has the highest correlation. The deterioration in the performance from landmark set 1 to set 3 can be attributed to the robot wandering from the center of the corridor.

For the majority of the four meter traverse the robot knows approximately where it is. It achieves this from purely visual clues, not relying on any past observations or odometry. Clearly the process could benefit from knowledge of previous positioning and a probabilistic model could be added to the system to enhance performance and robustness. Also an expectation of encountering a place due to past observations could significantly reduce the search time for evaluating potential places and add to the scalelability of the system.

4 Conclusions and Further Work

This system described in this paper demonstrates the use of automatically selected landmarks from panoramic images in mobile robot localisation. The Turn Back and

Look phase of landmark selection results in landmarks which are more reliable and can be recognised from a greater area in the robot's environment. These landmarks can also be located despite dramatic changes in illumination, due to the normalised correlation routines implemented. Finally the system was demonstrated by the robot localising between a number of learnt places along a four meter trajectory.

An obvious extension of this work, as mentioned above, is to introduce a probabilistic model for localisation, with past observations influencing the outcome of the next observation. The association of landmarks with measures of reliability and correlation to the observed scene should lend well to this approach.

The number and positioning of landmarks that are desirable to represent a place is also an interesting question. While in these experiments a minimal number of landmarks were used, today's hardware allows a much greater number of landmarks to be tracked at frame rate. The physical position of landmarks in relation to the robot also poses a problem. For gross localisation, landmarks which are far away and will not change under translation, or even move in the visual field, are desirable. For accurate localisation and navigation tasks, however, landmarks that do not move, or move very slightly in the visual field in response to ego motion, cannot provide the robot reliable information about its relative position. So landmarks which are very far away or lie along the direction of motion of the robot are not as useful.

The desirable number and positioning of landmarks for localisation and navigation is currently being investigated. Navigation algorithms for homing to places and navigation between places are also being developed.

References

[Bianco and Zelinsky, 1999] G. Bianco and A. Zelinsky. Biologically inspired visual landmark learning and navigation for mobile robots. In *Proceedings of the 1999 IEEE/RSJ International Conference on Intelligent Robotic Systems (IROS '99)*, volume 2, pages 671–676, 1999.

[Collet and Zeil, 1996] T.S. Collet and J. Zeil. Flights of learning. *Journal of the American Psychological Society*, pages 149–155, 1996.

[Lehrer, 1993] Lehrer. Why do bees turn back and look? *Journal of Comparative Physiology*, 172(A):549–563, 1993.

[Matsumoto *et al.*, 1997] Y. Matsumoto, M. Inaba, and H. Inoue. Memory-based navigation using omni-view sequence. In *Int. Conf. of Field and Service Robotics*, pages 184–191, 1997.

[Moravec, 1977] H. Moravec. Towards automatic visual obstacle avoidance. In *Proceedings of the 5th International Joint Conference on Artificial Intelligence)*, volume Vision-1, page 584, 1977.

[Mori *et al.*, 1995] T. Mori, Y.Matsumoto, T. Shibata, M. Inaba, and H. Inoue. Trackable attention point generate based on classification of correlation value distribution. In *JSME Annual Conf. on Robotics and Mechatronics (ROBOMEC 95).*, pages 1076–1079, 1995.

[Thompson *et al.*, 1999] S. Thompson, A. Zelinsky, and M.V.Srinivasan. Automatic landmark selection for navigation with panoramic vision. In *Proceedings of Australian Conference on Robotics and Automation (ACRA'99)*, 1999.

[Thrun, 1996] S. Thrun. A bayesian approach to landmark discovery and active perception in mobile robot navigation. Technical report, School of Computer Science Carnegie Mellon University, 1996.

[Yagi *et al.*, 1994] Y. Yagi, S. Kawato, and S. Tsuji. Real-time omnidirectional image sensor (copis) for vision-guided navigation. *IEEE Transactions on Robotics and Automation*, 10(1):11–21, 1994.