Performance Comparison of Landmark Recognition Systems for Navigating Mobile Robots

Tom Duckett
Dept. of Technology
University of Örebro
S-70182 Örebro
Sweden
Tom.Duckett@aass.oru.se

Ulrich Nehmzow
Dept. of Computer Science
University of Manchester
Manchester M13 9PL
England
ulrich@cs.man.ac.uk

Abstract

Self-localisation is an essential competence for mobile robot navigation. Due to the fundamental unreliability of dead reckoning, a robot must depend on its perception of external environmental features or landmarks to localise itself. A key question is how to evaluate landmark recognition systems for mobile robots. This paper answers this question by means of quantitative performance measures. An empirical study is presented in which a number of algorithms are compared in four environments. The results of this analysis are then applied to the development of a novel landmark recognition system for a Nomad 200 robot. Subsequent experiments demonstrate that the new system obtains a similar level of performance to the best alternative method, but at a much lower computational cost.

Introduction

The most important requirement for robot navigation — other than staying operational and avoiding collisions — is that of establishing one's own position (*self-localisation*). One possible self-localisation method is by dead reckoning using the robot's odometry. However, major problems with odometry include drift errors caused by wheel slippage and the need for *a priori* knowledge of the robot's position. A solution to both these problems is to use perception of external environmental features (*landmarks*).

The primary motivation for the work presented here was to develop a robust landmark recognition system for a mobile robot navigating over large, real world environments. This goal was achieved by conducting an experimental comparison of existing systems for landmark recognition, based on quantitative performance measures, and then using the results of this analysis to develop a novel landmark recognition system.

The performance criterion applied is *localisation quality versus computational cost.* Experiments are presented in which the self-localisation performance of a Nomad 200 robot equipped with ultrasonic range-finder sensors and a compass (see Fig. 1) is assessed while traversing a series of environments, using each of the different landmark recognition systems under investigation. The new landmark recognition system is shown to obtain a similar level of localisation quality to the best alternative method, but at a significantly reduced computational cost.

Copyright © 2000, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

Related Work

So far, relatively few attempts have been made to quantify robot-environment interactions or to conduct experimental comparisons of navigating robots. Exceptions include (Schöner & Dose 1992; Smithers 1995), where fundamental sensor-motor behaviors were analysed in terms of dynamical systems theory; (Lee & Reece 1994), where exploration strategies for mapping unfamiliar environments were evaluated; and (Gutmann *et al.* 1998; Thrun 1998), where various algorithms for self-localisation were compared.

Performance Measurement

The experiments were conducted using played-back sensor data recorded by the robot using wall-following. In each environment, the data from the robot's first lap of the environment was used for landmark learning ("map building"), and the data from the subsequent laps was used for testing ("localisation").

Localisation quality was measured using a statistic known as the *uncertainty coefficient, U, of L given R*, which measures the extent to which the robot's response, R (the response of the particular landmark recognition system under investigation to a perceptual stimulus) predicts the robot's true location, L, and is defined by

$$U(L \mid R) = \frac{H(L) - H(L \mid R)}{H(L)},$$

$$H(L) = -\sum_j p_{\bullet j} \ln p_{\bullet j},$$

$$H(L \mid R) = -\sum_{i,j} p_{ij} \ln \frac{p_{ij}}{p_{i\bullet}},$$

where $p_{\bullet j} = \sum_i p_{ij}$, $p_{i\bullet} = \sum_j p_{ij}$, and p_{ij} refers to the probability that the response is i and the true location

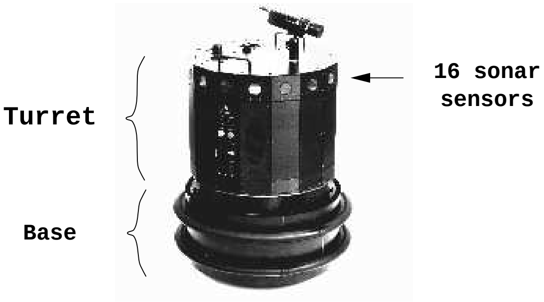

Figure 1: Nomad 200 mobile robot. A flux gate compass was used to keep the turret, and therefore the sensors, at a constant orientation during data collection.

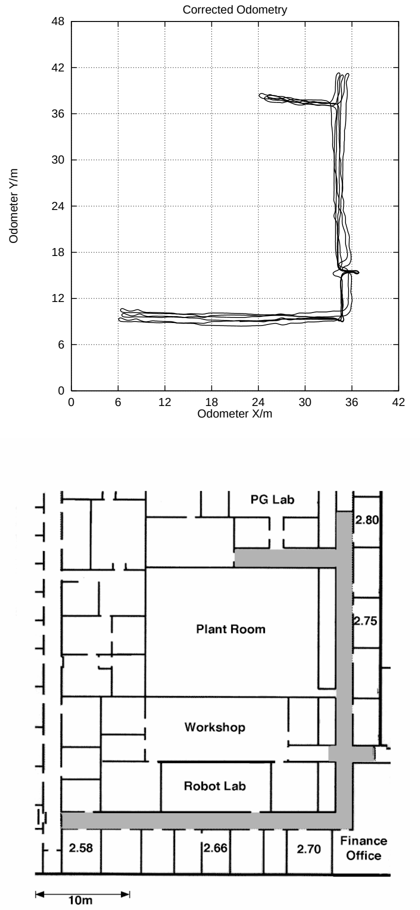

Figure 2: Top: location binning mechanism. Bottom: corresponding floor plan (environment C in Table 1). The dotted grid was used to coarse-code the corrected location data into bins of size 6 m $\times$ 6 m — for performance measurement, the current bin was taken as the robot's "true" location L.

is j (Press *et al.* 1992). A value of $U = 1$ implies that the robot's response always predicts its true location (perfect localisation), while $U = 0$ implies that the response never predicts the true location. The higher the value of U, the better the performance of the system.

To calculate U, some means of tracking the true location L of the robot is required. In initial experiments, the robot's position was measured by hand, but this process was costly and prone to human error. A mechanism for location tracking was therefore developed which is based on retrospectively corrected odometer data. Here, a flux gate compass was used to remove the rotational error affecting the robot's on-line dead reckoning, then the remaining translational drift error was removed by hand (through manual identification of prominent features in the trajectory). The corrected odometer data was coarse-coded into bins, as shown in Fig. 2; see also (Duckett & Nehmzow 1998).

The compass was also used to keep the robot's sensors at a constant orientation, so that the appearance of locations depended only on the robot's position, not the direction of travel.

Landmark Recognition Systems

Robot navigation is possible using artificial landmarks such as beacons, markers or induction loops. However, modifying the robot's environment is costly and inflexible. It is therefore desirable to use "natural" landmarks, i.e., the sensory perceptions a robot obtains in an unmodified environment.

One possible approach is to provide the robot with *a priori* designer-determined landmarks such as doors or ceiling lights. However, this approach can be brittle, due to the different perception of an environment by the designer from that of the robot, and can only be used in environments which contain these features. Instead, we concentrate on landmark recognition systems in which the robot is able to represent its own, *arbitrary* sensor patterns and to exploit whatever features are naturally present in an environment.

RCE Classifier

The first approach considered was the simple classifier mechanism used by (Kurz 1996), in which the robot's sensor patterns are classified according to the nearest neighbour among a set of stored prototypes (see Fig. 3). Each pattern consists of a normalised vector of sonar readings, and the dot product is used to compare vectors. During training, a new pattern is created if the input pattern fails to lie within a fixed sphere of any existing pattern.

ART2 Classifier

Several authors (Racz & Dubrawski 1995; Balkenius & Kopp 1996) have used neural networks based on Adaptive Resonance Theory (Carpenter & Grossberg 1987) for self-localisation. The principal difference between ART and feedforward classifiers such as RCE is the addition of a feedback phase, in which the best matching

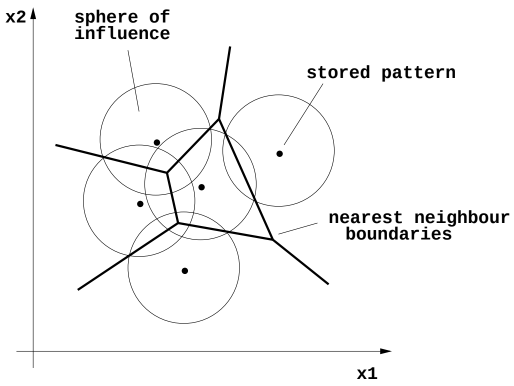

Figure 3: RCE classifier (example in 2 dimensions). The sensory input is classified according to the nearest stored pattern — for performance measurement, the winning pattern was taken as the robot's response R.

unit in the feedforward phase may be rejected and the system searches for another prototype to match the input pattern. This mechanism is used to implement the *self-scaling property*, which prevents any pattern that is a subset of another from being classified in the same category. During training, a new pattern is created if none of the stored patterns is similar enough to the input, otherwise the winning unit is modified to be more similar to the input pattern.

Growing Cell Structures

A number of authors (Nehmzow & Smithers 1991; Kurz 1996; Janet & others 1995) have considered mobile robot localisation using the Self-Organising Map (Kohonen 1993), a neural network which preserves topological relations in its training data. However, both the structure and size of the Kohonen network have to be fixed in advance by the designer, which means that the system cannot be used to map environments whose size is not known *a priori*. To overcome this problem, Fritzke (1994) developed a growing self-organising network which can store an arbitrary number of patterns. Like the Kohonen network, the neighbouring patterns of the best matching unit are adapted during training. In addition, a new pattern is inserted at regular intervals into the most adapted region of the network; see (Fritzke 1994) for full details.

Nearest Neighbour Classifier

This is the landmark recognition mechanism used by (Duckett & Nehmzow 1998). The version used in these experiments was identical to the RCE classifier described above, except that it uses *a priori* position information (from the retrospectively corrected odometry in these experiments, as in Fig. 2) to decide when to add new patterns to the robot's map. In these experiments, new sensor patterns were added to the map at 1.5 m intervals.

Occupancy Grid Matching

Another method for landmark recognition is by matching local occupancy grids (Yamauchi & Langley 1997), where the robot's map consists of a set of stored occupancy grids, one for each place visited by the robot. During localisation, a recognition grid is constructed from the robot's immediate sensor readings and then matched against each of the stored grids. A hill climbing procedure is used to search the space of possible translations[1] between the recognition and stored grids, using an evaluation function to determine the quality of the match. The best matching grid pattern determines the location of the robot. In these experiments, grid patterns were added to the robot's map at 1.5 m intervals, as for the nearest neighbour classifier.

New System: Occupancy Histogram Matching

The main disadvantage of occupancy grid matching its high computational requirements. Here we describe a new landmark recognition system, which introduces a much faster method of matching local occupancy grids.

Again, the robot's map consists of a list of places added at 1.5 m intervals. Landmark information is attached to each of the places as follows. Firstly, the robot takes a detailed sonar scan at its current location and a local occupancy grid consisting of 64×64 cells is constructed, as in (Yamauchi & Langley 1997). However, in the new system, the occupancy grids themselves are not stored or matched. Instead, each grid is reduced to a pair of histograms (one in x direction, and one in y direction), which is then used as a stored signature for that place in the robot's map, as shown in Fig. 4. In the absence of a compass, we would also have to consider angle histograms, as in (Hinkel & Knieriemen 1988).

Each occupancy grid cell represents an area of 15 cm × 15 cm, and is considered as being in one of three possible states; occupied (O), empty (E) or unknown (U), depending on the corresponding probability of occupancy for that cell, i.e.,

$$State(c_{xy}) = \begin{cases} O & \text{if } p(c_{xy}) > 0.5 \\ U & \text{if } p(c_{xy}) = 0.5 \\ E & \text{if } p(c_{xy}) < 0.5 \end{cases}$$

where $p(c_{xy})$ refers to probability of occupancy for the cell at column x and row y. These probabilities were obtained using the standard method for updating occupancy grids (Moravec & Elfes 1985). One histogram is then derived by adding up the total number of occupied, empty and unknown cells in each of the 64 columns, and the other by adding up the totals for each of the 64 rows.

To begin landmark recognition, the robot takes a new sonar scan. Again, the resulting occupancy grid

[1] The self-orientation component of this system was disabled, using the compass instead, in order to make a fair comparison between systems.

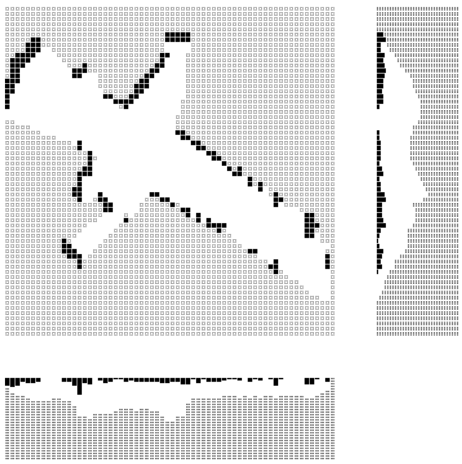

Figure 4: Example occupancy grid and histograms. Occupied cells are shown in black, empty cells in white and unknown cells in grey. A separate pair of histograms is used to represent each individual place in the robot's map.

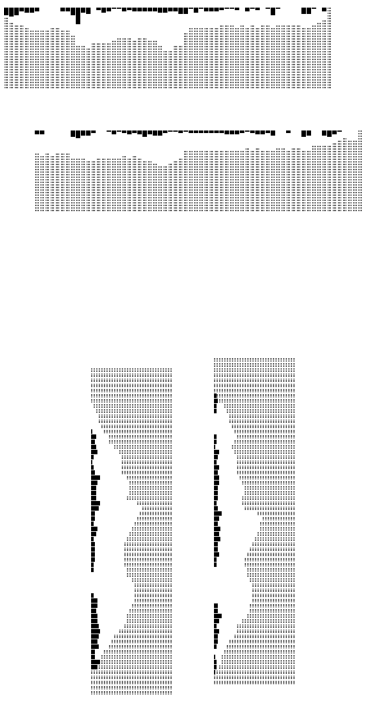

Figure 5: Matching the x and y histograms. The new histograms are convolved with the stored histograms to find the best match.

is processed to produce a pair of histograms. These histograms are then convolved with the corresponding stored histograms for all of the places in the robot's map, as illustrated in Fig. 5. The strength of the match between two histograms T^a and T^b is calculated using the following evaluation function:

$$Match(T^a, T^b) = \sum_j \left[min(O_j^a, O_j^b) + min(E_j^a, E_j^b) + min(U_j^a, U_j^b) \right],$$

where O_j, E_j and U_j refer to the number of occupied, empty and unknown cells contained in the jth element of histogram T. In the convolution, the stored histogram is kept stationary and the recognition histogram is translated against it, using the above function to calculate the best match over the 64 elements of the stored histogram. Any non-overlapping elements in the recognition histogram due to the translation are assumed to consist entirely of unknown cells.

A combined match score M_i for each stored place i is then calculated from best matching alignment for both the x and y histograms, i.e., the translations producing the highest match scores, as

$$M_i = Match(T_x^S, T_x^i) \times Match(T_y^S, T_y^i)$$

where the T^S refer to the x and y histograms for the new sonar scan S, and the T^i to the best matching histograms for place i. During testing, the place with the best match score M_i is taken as the robot's response.

Experiments

Robot sensor data were collected in four real world environments chosen to test the different systems under a variety of conditions, including high levels of perceptual aliasing[2], specular reflection and cross-talk (see Table 1). Environment D is an extreme case, consisting of a very long corridor with few distinctive features. All environments were subject to unpredictable variations in the sensor data, for example, due to people walking past the robot or doors being opened and closed.

The sonar readings were recorded by stopping the wall-following robot every 0.50 m, and rotating its turret to obtain a detailed scan consisting of 144 sonar readings, i.e., nine sets of 16 sonar readings taken at 2.5° intervals.

Performance measurement was carried out for all six landmark recognition mechanisms, in all four environments. The uncertainty coefficient U in each experiment was determined, as was the computational cost of landmark recognition. The latter was determined as the mean processor time required to match a landmark. To enable a fair comparison between systems, the parameters for each of the mechanisms were configured as closely as possible so that each system produced the

[2]where several places are perceptually similar enough to be confused by the robot

Environ.	**Description**	**Approx. Size**	**Route Length**	**Data Points**	N_L	N_R
A	T-shaped hallway	16 m × 13 m	54 m	623	5	16
B	Conference room	16 m × 11 m	49 m	668	6	25
C	L-shaped corridor	34 m × 33 m	147 m	854	12	42
D	Long corridor	53 m × 3 m	111 m	645	9	32

Table 1: Characterisation of environments. N_L denotes the number of location bins, and N_R the average number of responses used in the calculation of the uncertainty coefficient $U(L \mid R)$. The number of data points used for performance evaluation is also indicated.

Environment	**RCE**	**ART2**	**NstNbr**	**GCS**	**OccGrd**	**OHM**
A	0.554	0.573	0.650	0.719	0.732	0.806
B	0.552	0.669	0.715	0.770	0.879	0.850
C	0.462	0.538	0.551	0.502	0.644	0.632
D	0.220	0.265	0.350	0.340	0.487	0.439
$\overline{U}$	0.447	0.509	0.567	0.582	0.686	0.682
T	t	$218\,t$	t	t	$13051\,t$	$31\,t$

Table 2: Localisation quality (uncertainty coefficient) $U(L \mid R)$ and mean uncertainty coefficient $\overline{U}$ for the RCE classifier, ART2 classifier, Nearest Neighbour classifier (NstNbr), Growing Cell Structures (GCS), occupancy grid classifier (OccGrd) and the new landmark recognition system (OHM) in environments A to D. The computational cost T for each algorithm is given in time per landmark match, where $t = 1.8 \times 10^{-5} s$ as measured on a Sparcstation 20.

same number of responses N_R in each environment (see Table 1). Results are given in Table 2 and Fig. 6.

As can be seen from Fig. 6, of the three landmark recognition mechanisms with the lowest computational cost (the RCE, Nearest Neighbour and GCS classifiers), the RCE classifier always comes out worst, with the other two being almost equal in performance.

The remaining classifiers all incur a computational cost increased by one to four orders of magnitude. ART2, despite being two orders of magnitude more computationally expensive, actually performs worse than the Nearest Neighbour and GCS classifiers.

The best landmark recognition mechanism in terms of localisation performance alone is occupancy grid matching ($\overline{U} = 0.686$). Marginally lower in performance ($\overline{U} = 0.682$), but three orders of magnitude cheaper in computation, is the histogram matching classifier. Our conclusion from this data is that for real time, autonomous operation it is optimal to use occupancy histogram matching.

A statistical test was also performed to evaluate the significance of these results. This consisted of a pair-wise comparison of the systems, using Student's t-test for paired samples (Press *et al.* 1992) to test the null hypothesis that their performance U over the four environments is really the same. The results in Table 3 indicate significant differences between all of the systems ($p \preceq 0.05$), except in the comparisons between the Nearest Neighbour and GCS classifiers ($p = 0.60$), and the new histogram matching and occupancy grid classifiers ($p = 0.90$). There is a slight anomaly in the com-

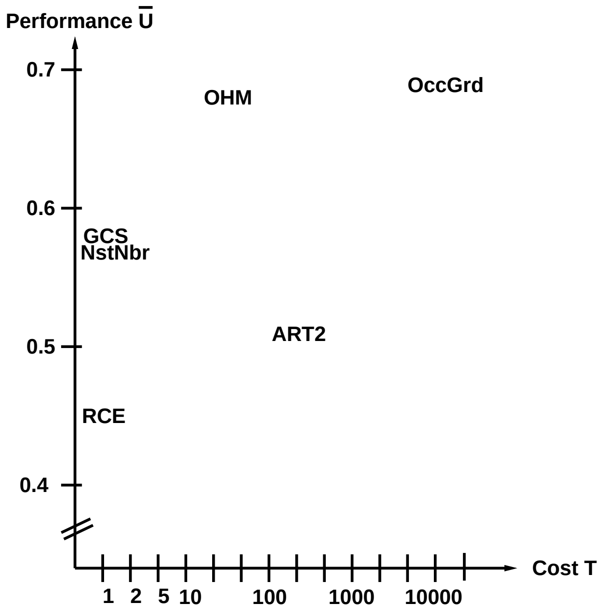

Figure 6: Localisation quality versus computational cost for the systems investigated (see also Table 2). Abbreviations as in Table 2.

p	**ART2**	**NstNbr**	**GCS**	**OccGrd**	**OHM**
RCE	0.06	0.01	0.04	0.01	0.01
ART2		0.04	0.16	0.01	0.01
NstNbr			0.60	0.01	0.01
GCS				0.05	0.01
OccGrid					0.90

Table 3: Paired Student's t-test results for the comparative study. Each pair of systems in table 2 was compared in turn, computing the probability p of obtaining these results assuming the null hypothesis that their performance U is really the same.

parison of ART2 and GCS, though we should expect some variations given the size of the samples.

Conclusion

There have been many proposals in the AI literature for navigating mobile robots. However, only recently have there been any attempts to make objective comparisons between different approaches. In this paper, we decided to investigate various methods for performing landmark recognition. The results showed that good localisation can be obtained by matching local occupancy grids, as in (Yamauchi & Langley 1997). Unfortunately, this performance was obtained only at a computational cost four orders of magnitude higher than that of the "cheaper" systems investigated. The new occupancy histogram matching method presented in this paper offers a viable alternative. It has a localisation performance similar to the occupancy grid matching method, at a cost that is only one order of magnitude higher than the "cheaper" methods. Our conclusion is that occupancy histogram matching is a strong candidate for landmark recognition by a navigating mobile robot, especially in situations where computational cost matters. In ongoing work, we have successfully applied this new technique in a complete navigation system which uses previous location information to further improve self-localisation performance (Duckett & Nehmzow 1999).

Acknowledgements

This work was carried out as part of the first named author's PhD thesis at the University of Manchester, who gratefully acknowledges a studentship provided by the Department of Computer Science.

References

Balkenius, C., and Kopp, L. 1996. The XT-1 vision architecture. In *Proc. Symposium on Image Analysis, Lund University, Sweden.*

Carpenter, G., and Grossberg, S. 1987. ART2 : Self–organization of stable category recognition codes for analog input patterns. *Applied Optics* 26(23).

Duckett, T., and Nehmzow, U. 1998. Mobile robot self-localisation and measurement of performance in middle scale environments. *Robotics and Autonomous Systems* 24(1–2). Online at http://www.cs.man.ac.uk/~duckettt.

Duckett, T., and Nehmzow, U. 1999. Knowing your place in real world environments. In *Proc. EUROBOT '99*, 135–142. IEEE Computer Press. Online at http://www.cs.man.ac.uk/~duckettt.

Fritzke, B. 1994. Growing cell structures - a self-organizing network for unsupervised and supervised learning. *Neural Networks* 7(9).

Gutmann, J.-S.; Burgard, W.; Fox, D.; and Konolige, K. 1998. An experimental comparison of localization methods. In *Proc. IROS'98.*

Hinkel, R., and Knieriemen, T. 1988. Environment perception with a laser radar in a fast moving robot. In *Symposium on Robot Control (SYROCO'88).*

Janet, J., et al. 1995. Global self-localization for autonomous mobile robots using self-organizing Kohonen neural networks. In *Proc. IROS'95.*

Kohonen, T. 1993. *Self-Organization and Associative Memory, 3rd ed.* Springer.

Kurz, A. 1996. Constructing maps for mobile robot navigation based on ultrasonic range data. *IEEE Trans. Systems, Man and Cybernetics B* 26(2).

Lee, D., and Reece, M. 1994. Quantitative evaluation of the exploration strategies of a mobile robot. In *Proc. AAAI'94.*

Moravec, H., and Elfes, A. 1985. High resolution maps from wide angle sonar. In *Proc. ICRA'85.*

Nehmzow, U., and Smithers, T. 1991. Mapbuilding using self-organising networks in really useful robots. In *Proc. SAB '91.*

Press, W.; Teukolsky, S.; Vetterling, W.; and Flannery, B. 1992. *Numerical Recipes in C, 2nd. edition.* Cambridge University Press.

Racz, J., and Dubrawski, A. 1995. Artificial neural network for mobile robot topological localization. *Robotics and Autonomous Systems* 16.

Schöner, G., and Dose, M. 1992. A dynamical systems approach to task-level system integration used to plan and control autonomous vehicle motion. *Robotics and Autonomous Systems* 10:253–267.

Smithers, T. 1995. On quantitative performance measures of robot behaviour. *Robotics and Autonomous Systems* 15(1-2):107–133.

Thrun, S. 1998. Bayesian landmark learning for mobile robot localisation. *Machine Learning* 33(1).

Yamauchi, B., and Langley, P. 1997. Place recognition in dynamic environments. *J. Robotic Systems* 14(2).