Quantitative Analysis of Mobile Robot Localisation Systems

Tom Duckett and Ulrich Nehmzow
Department of Computer Science
University of Manchester
Manchester M13 9PL
United Kingdom
duckettt@cs.man.ac.uk, u.nehmzow@cs.man.ac.uk

5.9.97

Abstract

This paper addresses the question of how to measure the performance of mobile robot localisation systems. A general performance metric and a standard experimental procedure are introduced, where no semantic interpretation of the robot's internal world model is required. A novel mechanism for tracking the "true" location of the robot is also presented.

Together, these methods facilitate the comparison of disparate localisation systems on the same robot in the same environment, and also the replication of such experiments by other researchers. To demonstrate the utility of the approach taken, three different localisation systems are compared using the same data collected by a real robot travelling over a route of $175m$.

1 Introduction

At the present time, many self-localisation systems have been proposed for navigating mobile robots. However, quantitative comparison of the different approaches is largely impossible, because researchers use different robotic platforms, apply different quality measures, and conduct different experiments to validate their implementations. In addition, the lack of experimental detail in many robotics publications prevents replication of experiments by other researchers.

Mobile robot navigation systems are also rarely tested outside of small-scale, laboratory environments (the use of "toy worlds" is a common trait in many areas of AI), and the question of scaleability is avoided. To be really useful, we believe that navigating robots should be tested in middle-scale, "real world" environments. The definition of middle-scale navigation is taken from biology (see e.g., [10]), referring to journeys of up to 10^6 times the animal's body length. For our Nomad 200, this means navigating in environments of several hundred m^2 in size.

In this paper, we therefore propose a general performance measure (see section 2) and a standard experimental procedure (see section 4) for evaluating localisation quality. Together, these methods facilitate the comparison of different localisation systems on the same robot in the same (middle-scale) environment. A data recording mechanism is described, including a novel mechanism for autonomously tracking the "true" location of the robot (section 4.4). To demonstrate the utility of the approach taken, three different localisation systems are compared in section 5.

2 The Performance Metric

Various different measures have been used by mobile robotics researchers to assess localisation quality. Absolute error, i.e., the distance between the robot's predicted and actual position in a Cartesian reference frame, is one approach (see e.g., [6]). However, such a metric does not facilitate comparison with systems

Proc. TIMR UK 97, British Conference on Autonomous Mobile Robotics and "Scientific Methods in Mobile Robotics", Manchester 1997. Technical Report Series, Department of Computer Science, Manchester University, ISSN 1361 - 6161. Report number UMCS-97-9-1. http://www.cs.man.ac.uk/csonly/cstechrep/titles97.html

Location (L)

Response (R)						
	0	2	15	0	1	18
	10	10	0	0	0	20
	0	2	1	0	19	22
	5	7	3	1	1	17
	0	0	0	23	0	23
	15	21	19	24	21	100

Figure 1: *Example Contingency Table.* The rows correspond to the response produced by the particular system under investigation, and the columns to the "true" location of the robot as measured by an independent observer. This table represents 100 data points, and also shows the totals for each row and column.

which produce a non-Cartesian response to their perceived location, e.g., Kohonen feature map [5], and is also difficult to measure over large environments.

Another approach is to evaluate a percentage correct figure for the system under investigation (see e.g., [11]). However, this calculation requires that the human observer is able to interpret the meaning of the robot's response with respect to the robot's environment. Such an interpretation would not be readily available in many cases, e.g., hippocampal [1] and self-organising systems [5], where the robot forms its own internal representation of the world.

In our approach, we therefore adopt a "black box" model (or stimulus-response model) of the robot's interactions with its environment. A performance metric based on the concept of mutual information [8] is detailed in the following section. This effectively measures the extent to which the robot's response predicts its true location, without requiring any semantic interpretation of the robot's internal world model by the observer.

2.1 Contingency Analysis

The basis of this method of quantitative analysis is a data structure known as a contingency table (see figure 1). In this example, a sample consisting of 100 data points has been collected. Each data point has two attributes; one corresponding to the location predicted by the robot (the robot's *response*, R), and the other to the actual location of the robot as measured by an observer (the robot's *location*, L). By convention, the rows of the table are used to represent the response, and the columns to represent the location. For example, figure 1 shows one cell containing 19 data points where the robot's response was measured as row 3 and the location as column 5.

The question we wish to ask with the metric is *"How much information does R provide about L?"*. For the metric to be meaningful, it is therefore important to make sure that both response and location are measured at the same degree of resolution, i.e., that the number of rows equals the number of columns. Procedures for coarse-coding R and L to the appropriate level of resolution are provided later in sections 4.4 and 4.5.

For a contingency table, the **Row Totals** for each response i (equation 1), **Column Totals** for each location j (equation 2) and the **Table Total** (equation 3) are calculated as follows, where N_{ij} refers to the number of data points contained in the cell at row i and column j.

$$N_{i\bullet} = \sum_j N_{ij} \tag{1}$$

$$N_{\bullet j} = \sum_i N_{ij} \tag{2}$$

$$N = \sum_{i,j} N_{ij} \tag{3}$$

For the example table, the total of row 3 ($N_{3\bullet}$) is 22, the total of column 5 ($N_{\bullet 5}$) is 21, and the table total is 100.

The **Row, Column and Cell Probabilities** can then be calculated using the following equations.

$$p_{i\bullet} = \frac{N_{i\bullet}}{N} \tag{4}$$

$$p_{\bullet j} = \frac{N_{\bullet j}}{N} \tag{5}$$

$$p_{ij} = \frac{N_{ij}}{N} \tag{6}$$

For the example table, the probability of a data point lying in row 3 is 0.22, the probability of a data point lying in column 5 is 0.21, and the probability of a data point lying in cell $(3, 5)$ is 0.19.

The next set of equations are used to evaluate the entropy of the variables under consideration, i.e., the amount of information required to remove any uncertainty in these quantities. Here, we define the **Entropy of L** (equation 7), the **Entropy of R** (equation 8) and the **Mutual Entropy of L and R** (equation 9).

$$H(L) = -\sum_j p_{\bullet j} \ln p_{\bullet j} \tag{7}$$

$$H(R) = -\sum_i p_{i\bullet} \ln p_{i\bullet} \tag{8}$$

$$H(L, R) = -\sum_{i,j} p_{ij} \ln p_{ij} \tag{9}$$

In particular, we are interested in measuring the useful information provided by R in predicting the value of L. (We are not concerned with the reverse relationship; for example, if two responses both predict the same location this should not have a negative impact on the metric.) Therefore, the **Entropy of L given R** is obtained as follows.

$$H(L \mid R) = H(L, R) - H(R) \tag{10}$$

where

$$0 \leq H(L \mid R) \leq H(L) \tag{11}$$

This last property (equation 11) means that we can scale the metric to lie between 0 and 1. The following statistic, known as the **Uncertainty Coefficient of L given R**, provides a natural measure of the association or "dependency" of L on R.

$$U(L \mid R) \equiv \frac{H(L) - H(L \mid R)}{H(L)} \tag{12}$$

A value of $U(L \mid R) = 0$ means that R provides no useful information about L, and implies that the robot's response never predicts its true location. A value of $U(L \mid R) = 1$ means that R provides all the

Figure 2: *The Manchester 'FortyTwo'.* Exploration was carried out using the translational and rotational motors located in the base of the robot. A separate behaviour was used to rotate the turret at small speeds towards "North" as indicated by a compass.

information required about L, and implies that the response always predicts the true location. It should also be noted that the ordering of the rows and columns in the contingency table makes no difference to the outcome of this calculation.

3 The Robot

The experiments presented here were conducted using the robot shown in figure 2. The robot has 16 sonar and 16 infrared sensors mounted around its turret, which can rotate independently relative to the base of the robot (the camera shown was not used here). Two other motors located in the base of the robot are used to control the translational and rotational movement of the robot.

Two reactive behaviours were used to control the robot's motor actions during these experiments. A wall-following behaviour was used to control the translational and rotational motors (see section 4.2). The turret was controlled by a separate behaviour, rotating it at small speeds in the direction of "North" as indicated by the robot's compass. The effect of this second behaviour was to smooth out local variations in the magnetic field of the environment, providing the robot with a global compass sense.

Thus, the robot always explored its environment with its sensors facing in the same global direction rather than the direction of travel. The appearance of locations to the robot therefore depended on the robot's position alone, not its orientation. This compass sense also made it possible to apply the location tracking mechanism described in section 4.4.

4 Standard Experiment

In this section, we describe an experimental procedure for evaluating localisation systems which has the following characteristics. Each aspect is then described in further detail.

- Middle-Scale Environment
- Exploration using Wall-Following
- Trace-Replay Mechanism

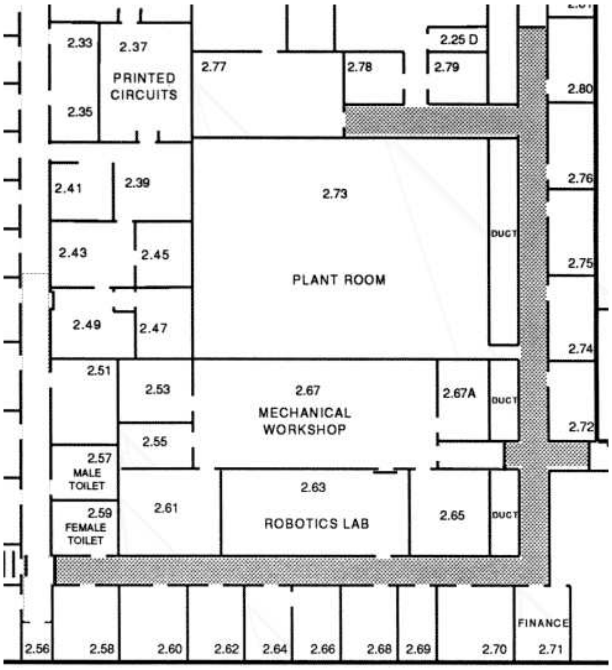

Figure 3: *Middle-Scale Environment.* The shaded area shows the enclosure explored by the robot in these experiments at Manchester University's Computer Building.

- Location Binning Mechanism
- Response Binning Procedure
- Uncertainty Coefficient over Time

4.1 Middle-Scale Environment

The experiment requires a large environment, as we are interested in the robot's ability to recognise distinct places in the "real world", rather than precise positioning over a small area. Some authors have drawn a distinction between "global" and "local" positioning (see e.g., Piasecki [7]); whilst this separation seems somewhat artificial, we are certainly more interested in the former of the two.

In the experiments presented here, the data used was collected by a robot travelling repeatedly over a circular route of $175m$ in Manchester University's Computer Science building (see figure 3).

4.2 Exploration Using Wall-Following

In our experiments, a reactive wall-following strategy was used for exploration with an average speed of $0.260ms^{-1}$. Wall-following was used because it is easily reproducible, and also because the purely reactive nature of this strategy means that sensor data can be recorded and played back for later experiments, whilst preserving the full complexity of robot-environment interaction.

4.3 Trace-Replay Mechanism

As the robot explored its environment, a trace-replay mechanism (as in Lee [4]) was used to record all of the sensory information available to the robot into a datafile at one second intervals. The collected data could then be played back to assess the performance of different localisation systems as required. Using the same recorded data throughout ensured that all experiments were conducted under identical conditions.

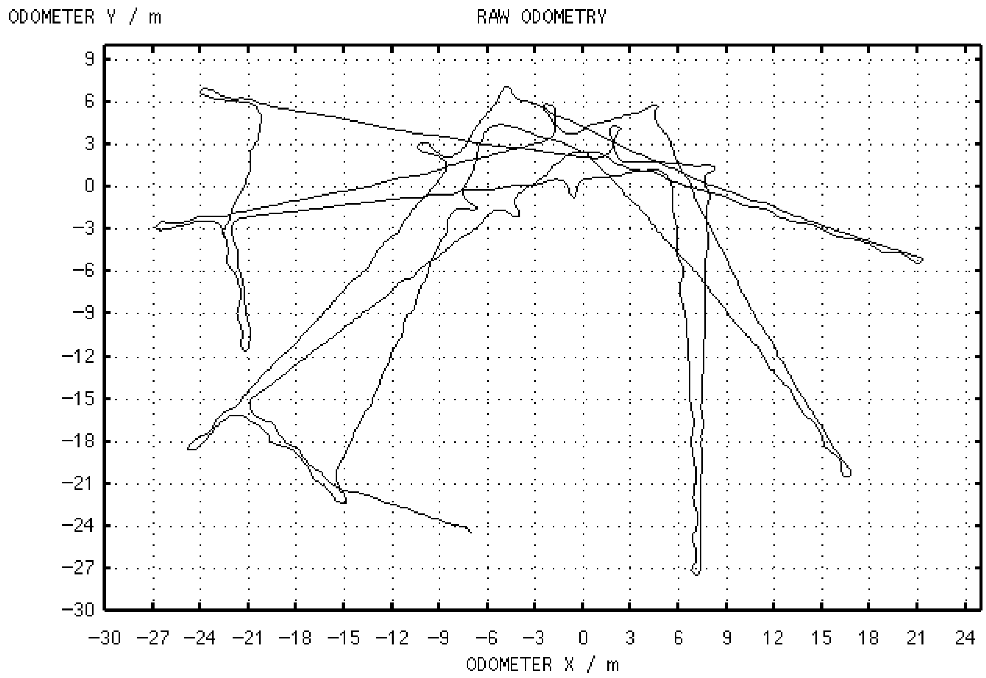

Figure 4: *Raw Odometry Data.* The recorded odometry trace shows accumulated rotational and translational drift errors due to wheel slippage; the rotational error in particular is affected by a strong clockwise bias.

4.4 Location Binning Mechanism

The problem addressed here was how to find some independent means of recording the "true" location of the robot over time for later comparison with the reponses of a particular localisation system. In earlier experiments (see [3]), this was done manually by drawing a grid on the floor and recording the time whenever the robot moved into a new grid cell. However, this process was time-consuming and prone to human error, so we decided to look for an automatic location recording mechanism. (A camera mounted on the ceiling would be possible in a laboratory environment, but not suitable here.)

The robot's own odometry data is unsuitable for location tracking because of the familiar problems of accumulated rotational and translational drift errors due to wheel slippage (see figure 4). However, using the robot's compass sense we were able to overcome the problem of rotational drift, and then retrospectively remove the remaining translational error as follows.

Firstly, instead of using the robot's rotational odometry for the on-line dead reckoning, we used the relative angular displacement of the turret against the base of the robot (see figure 5). This had the effect of removing the accumulated angular drift, because the robot's turret orientation was anchored to a global "North" by the compass sense described in section 3.

Secondly, the translational error was removed off-line by retrospectively correcting each data point by an amount proportional to the distance travelled along the route. This was done separately for each lap of the environment, and then an offset was applied to each of the successive laps to cancel out the remaining translational error between them. In these experiments, the recorded data was divided into laps manually by finding an obvious landmark (e.g., corner) in the odometer trace, although this could easily be done automatically by using an external sensor to detect the completion of another lap by the robot (as in Smithers [9]).

Finally, having obtained a reasonably accurate method of tracking the robot's location, the data was coarse-coded into equally sized bins ($3m \times 3m$ in these experiments) as illustrated in figure 5. The positioning of the dotted grid shown in figure 5 over the corrected odometer trace was determined using a hill-climbing procedure to minimise the number of bins occupied, although this could also be achieved manually by visual inspection of the data.

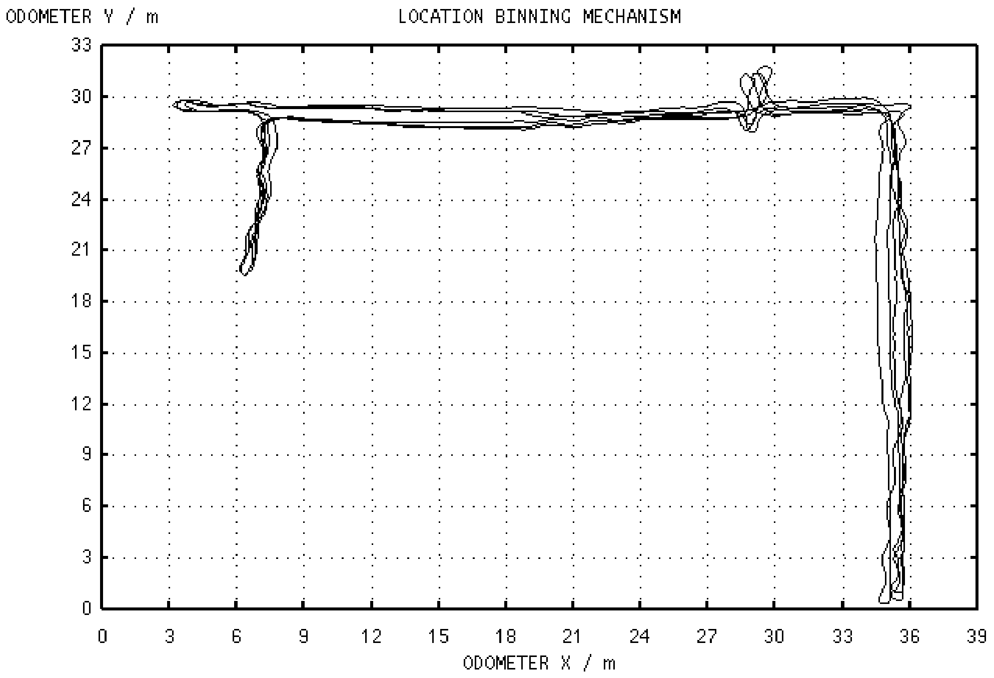

Figure 5: *Location Binning Mechanism.* The rotational error in the robot's odometry was removed on-line using a compass sense, and the remaining translational error was corrected off-line (see section 4.4). The dotted grid was used to coarse-code the location data into 24 bins.

4.5 Response Binning Procedure

This section addresses two important considerations in the treatment of the experimental data. Firstly, having determined the number of bins for measuring the robot's location L, the contingency analysis requires that the robot's response R also be coarse-coded into the same number of bins.

Also, since the size and shape of the location bins is decided arbitrarily, it is unlikely that there will be any direct correspondence with the size and shape of the locations recognised by the robot. This situation is undesirable, as it means the analysis could be adversely affected by quantisation errors (e.g., unusual overlaps between the two different place representations).

The solution taken to the quantisation problem is firstly to deliberately choose a much higher level of resolution for the response bins than the location bins, so that a typical location bin will contain a high number of different responses. (In one of the experiments, the performance of a Kohonen feature map with 169 possible responses was measured over 24 location bins.) However, this means of course that the corresponding contingency table will have a much higher number of rows than columns. Therefore, a special procedure is applied, which repeatedly combines rows of the table until the number of rows is equal to the number of columns. This has the effect of "squashing" the contingency table to the desired dimensions.

The method of choosing the rows to be combined here is very important, as a haphazard treatment could easily destroy any statistical association present in the data. A standard procedure is therefore provided (see figure 6), which attempts to maintain the best possible value of $U(L \mid R)$ by merging only the most similar rows together. This is a purely syntactic operation on the table; again, no external interpretation of the data is required.

4.6 Uncertainty Coefficient over Time

During the experiment, each of the localisation systems under investigation is tested the same number of times ("trials"), using the played-back sensori-motor experience recorded by the wall-following robot. Each trial begins from a different location along the route traversed by the robot, and lasts for a fixed begin

while (no. of R bins > no. of L bins)
begin

find row i_{min} with lowest frequency (i.e., smallest $N_{i\bullet}$)
find row $i_{nearest} \neq i_{min}$ which is most similar to row i_{min} (applying L2 norm and Euclidean distance between rows)
combine rows i_{min} and $i_{nearest}$ (adding together the corresponding cell totals)

end

end

Figure 6: *Response Binning Procedure.*

duration n. In the subsequent analysis, we assume that all trials begin nominally at time t_0 and end at time t_n.

After running a localisation system for the appropriate number of trials[1], the response binning procedure is applied to a contingency table constructed from all of the location-response data produced by that particular system. Finally, the resulting data is split into separate contingency tables, one for each second, t_i, from the nominal start of the experiment, t_0. Thus, the uncertainty coefficient can be plotted against time (see figure 7).

5 Experiments

To illustrate the workings of our approach, we compared the three following localisation methods using the experimental procedure described above. The first lap of the recorded robot data was used to train the systems (i.e., mapbuilding) if required, then the remaining two laps were used for testing (i.e., localisation).

- **Uncorrected Dead Reckoning.** At the start of each trial, the robot's uncorrected odometry (see figure 4) was reset to the "correct" position and orientation taken from the corrected odometer trace shown in figure 5. To obtain the response R, the (x, y) coordinates produced by the dead reckoning strategy over time were coarse coded into equally sized bins (using the same dotted grid as shown for the location binning mechanism), producing a total of 93 responses (i.e., the number of squares occupied in the dotted grid shown in figure 5) before the response binning procedure was applied.

- **Self-Organising Feature Map.** The robot's recorded sonar and infrared sensor data was combined in the input vector to a 13×13 self-organising feature map. During testing, the winning unit was taken as the robot's response R, producing a total of 169 possible responses before the response binning procedure was applied.

- **Evidence-Based Localisation.** This system (as described in [2] [3]) accumulates sensory evidence over time to choose between competing hypotheses of its perceived location. During mapbuilding, a map containing 163 place memories was constructed, each consisting of a feature vector obtained from the robot's sonar and infrared sensors associated with an (x, y) coordinate obtained by dead-reckoning (using the corrected compass-based odometry shown in figure 5). During localisation, the robot's sonar and infrared sensors are used to determine the place memory which most closely resembles the current sensory perception. Because of perceptual aliasing[2], typically a number of place memories emerge as likely candidates for the robot's current position. The localisation system then waits until the robot has moved by a certain distance, observing the relative displacement of

[1] In our experiments, we used 117 trials, each of duration $n = 180$ seconds, and starting at 10 second intervals along the route.

[2] identical sensory perceptions at different physical locations

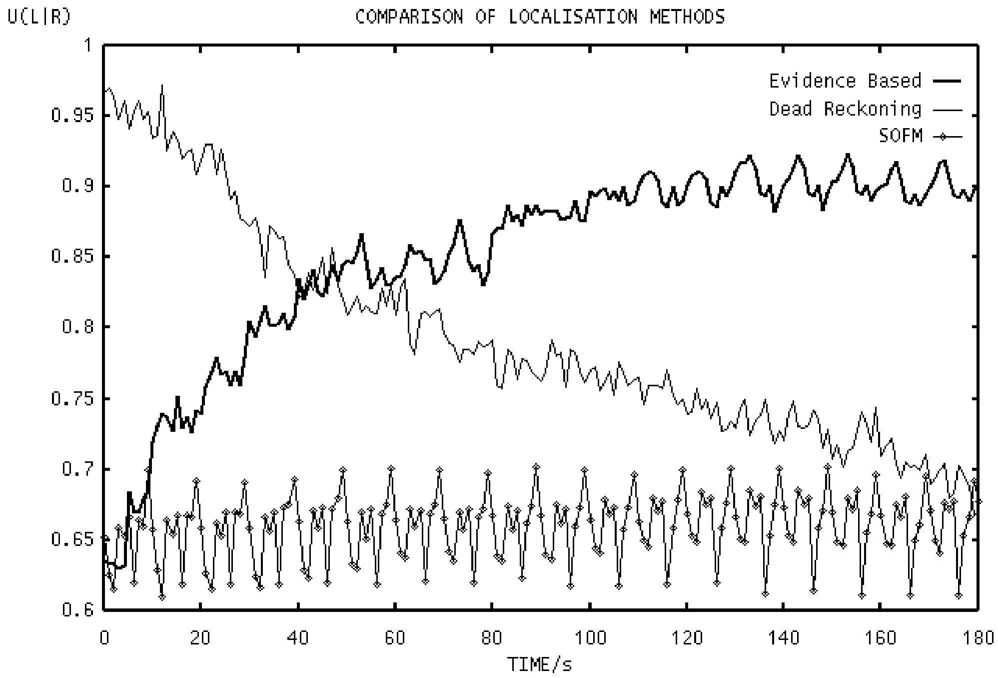

Figure 7: *Comparison of Localisation Methods.*

the robot in Cartesian space (using the robot's on-line compass-based odometry). The current perception, previous perception and displacement are then used to narrow down the list of candidates for the robot's current position, increasing or decreasing a confidence level associated with each of these hypotheses. During testing, the place unit with the highest confidence level was taken as the robot's response R, producing a total of 163 possible responses before the response binning procedure was applied.

The results are shown in figure 7. It can be seen that the performance of dead-reckoning falls over time, being overtaken by the evidence-based localiser at around $40s$ from the start of the experiment. The performance of the SOFM remains roughly constant over time. The relatively low value of the uncertainty coefficient for the SOFM here reflects the level of perceptual aliasing which occurred in this particular environment. We would expect the performance of dead-reckoning to fall below that of the SOFM over a longer experiment, and also to fall at a faster rate given a higher rate of odometer drift.

6 Summary

In this paper, we have presented a performance measure and an experimental procedure which enable different localisation systems to be compared on the same robot in the same environment. The proposed metric measures the information content of the responses of the particular system under investigation in predicting the observed location of the robot. A novel mechanism, based on a compass sense and retrospectively corrected odometry, was used for tracking the location of the robot.

Our experiments were conducted using logged sensor and odometry data recorded by a Nomad 200 following a circular path of 175 m by wall-following. Using recorded data, obtained from a real robot, ensured that all experiments were conducted under identical conditions, whilst retaining the complexity of robot-environment interaction.

Using the methods presented in this paper, we were thus able to compare three disparate localisation systems. These included a system using only proprioceptive sensory input (dead-reckoning), a system using only current exteroceptive sensory input (Kohonen feature map) and a system which combined both exteroceptive and proprioceptive sensory information over time to localise (evidence-based localiser). The results confirmed that the performance of the proprioceptive system worsens over time, while the exteroceptive system performs at a roughly constant level and the performance of the evidence-based system improves over time.

7 Conclusions and Future Work

The introduction of a quantitative performance measure and a standard experiment allowed us to compare different localisation algorithms directly on the same robot in the same environment.

In ongoing work, we intend to use these methods to compare the performance of localisation methods published in recent robotics literature, following reimplementation on our Nomad 200. These systems will be selected on the basis of their generality and known ability to perform on a real robot interacting with a real environment (e.g., [5] [11] [12]), rather than using simplified numerical models as in most robot simulations.

We believe that the quantitative analysis methods introduced here will serve as benchmark for future work on localisation, also allowing us to quantify the influence of different process parameters upon the performance of an entire system, rather than guessing "magic numbers" as in most current work in robotics.

To use quantitative performance measures is good scientific practice, and their use in mobile robotics not only facilitates comparison, but also replication of experiments by other workers in the field. We therefore believe that using quantitative methods more, while using mere existence proofs less, will move mobile robotics further towards an exact science.

References

[1] N. Burgess, J. G. Donnett and J. O'Keefe, "Robotic and Neuronal Simulation of Hippocampal Navigation", AISB-97 Workshop on Spatial Reasoning in Mobile Robots and Animals, University of Manchester, Technical Report Series UMCS-97-4-1, ISSN 1361-6161, 1996.

[2] T. Duckett and U. Nehmzow, "Experiments in Evidence-Based Localisation for a Mobile Robot", AISB-97 Workshop on Spatial Reasoning in Mobile Robots and Animals, University of Manchester, Technical Report Series UMCS-97-4-1, ISSN 1361-6161, 1996.

[3] T. Duckett and U. Nehmzow, "Knowing Your Place in the Real World", ECAL-97 Fourth European Conference on Artificial Life, University of Sussex, `http://www.cogs.susx.ac.uk/ecal97/`, 1997.

[4] D. C. Lee, "The Map–Building and Exploration Strategies of a Simple, Sonar- Equipped Mobile Robot; An Experimental, Quantitative Evaluation", Ph.D. thesis, University College London, England, 1995.

[5] U. Nehmzow, T. Smithers and J. Hallam, "Location Recognition in a Mobile Robot Using Self-Organising Feature Maps", in G. Schmidt (ed.), "Information Processing in Autonomous Mobile Robots", Springer Verlag, 1991.

[6] S. Oore, G. E. Hinton and G. Dudek, "A Mobile Robot That Learns Its Place", Neural Computation 9, pp. 683 - 699, 1997.

[7] M. Piasecki, "Global localization for mobile robots by multiple hypothesis tracking", Robotics and Autonomous Systems 16, pp. 93-104, 1995.

[8] C. E. Shannon and W. Weaver, "The Mathematical Theory of Communication", University of Illinois Press, 1949.

[9] T. Smithers, "On Quantitative Performance Measures of Robot Behaviour", Robotics and Autonomous Systems, Vol. 15, No. 1-2, pp. 107-133, 1995.

[10] R. Wehner, M. Lehrer and W. R. Harvey, editors, The Journal of Experimental Biology, 199. Special issue on Animal Navigation. 1996.

[11] B. Yamauchi and P. Langley, "Place Recognition in Dynamic Environments", Journal of Robotics Systems, Special Issue on Mobile Robots, Vol, 14, No. 2, pp. 107-120, 1997.

[12] U. R. Zimmer, "Self–Localization in Dynamic Environments", IEEE/SOFT International Workshop BIES'95, Tokyo, Japan, May 30-31 1995.