"Meaning" through Clustering by Self-Organisation of Spatial and Temporal Information

Ulrich Nehmzow
Department of Computer Science
Manchester University
Manchester M13 9PL
United Kingdom
ulrich@cs.man.ac.uk

Abstract

This paper presents an episodic mapping mechanism used for the self-localisation of autonomous mobile robots. A two layer self organising neural network classifies perceptual and episodic information to identify "perceptual landmarks" (and thus the robot's position in the world) uniquely. Through this process relevant information is obtained from the temporal flow of ambiguous and redundant sensory information, such that meaningful internal representations of the robot's environment emerge through an unsupervised process of self-organisation, constructing an analogy to the real world.

1 Introduction

1.1 Motivation: Transfer of redundant sensory perceptions to meaningful localisation information

The ability to move towards specific, identified places within the environment is one of the most important competences for a mobile robot.

While tasks that require only either random exploration or following canonical paths (such as paths marked by induction loops, beacons, or other markers) can be achieved by low-level behaviour based control (see e.g. [1]), more complex navigation tasks for mobile robots (such as for example delivery tasks) require the robot to determine its current position, the goal position, and the required motion between the two.

Apart from 'staying operational' (involving obstacle avoidance and staying within the machine's operational limits) navigation requires 1) the ability to construct a mapping between observed features and an internal representation ("mapbuilding"), and 2) the interpretation of this mapping ("map interpretation").

Following Webster's definition of "analogy" as a "resemblance in some particulars between things otherwise unlike" ([14]), such mapbuilding is the construction of an analogy between the real world and a robot's perception of it.

In order to use its representation of the world at all, the robot must know where it is in representation space: localisation is the most fundamental component of interpreting mappings. Unless the robot can identify its current position on the map, no path planning and hence no navigation can be performed.

To achieve self-localisation, the only information available to a mobile robot is the temporal flow of (redundant, noisy and ambiguous) sensory perception. The topic of this paper is how this temporal flow of perceptions can be used to generate stable and meaningful representations of space that can be used for mobile robot self-localisation. We present a localisation mechanism based on an "episodic mapping" paradigm that allows a mobile robot to identify its current position within the mapping of the environment, based on current and preceding perceptions of the world. Using an entropy-based quality metric we compare the episodic mapping with a static mapping paradigm ([9, 11, 10] (i.e. a paradigm that does not exploit temporal information), and demonstrate that episodic mapping produces measurably better localisation performance than static mapping.

[0]Published in C. Nehaniv (ed.), Computation for Metaphors, Analogy and Agents, Lecture Notes in Artificial Intelligence 1562, pp. 209-229, Springer Verlag, 1999.

1.2 Related Work

The mapping used here is similar in many ways to hippocampal mappings found in rats. In particular, place cells[1] in the rat's hippocampus can be likened to activity patterns observed in the self-organising feature maps used here. There have been a number of implementations of robot navigation systems that simulate such place cells, notably the work of Burgess, Recce and O'Keefe ([2, 3]), but also of others ([8, 9]).

Self-organising mechanisms for static sensor signal clustering have been used for robot localisation before ([9, 11, 6, 16]): the current sensory perception of a mobile robot is clustered through an unsupervised, self-organising artificial neural network, and the network's excitation pattern is then taken to indicate the robot's current position in perceptual space. If no perceptual aliasing (ambiguous sensory perceptions) were present, this would then also identify the robot's position in the world unambiguously. Contrary to work discussed in this paper, however, no information about perception over time was encoded in these cases.

Regarding the use of episodic information as input to a self-organising structure, some work has been done using such information in the input to a single layer Kohonen network ([10]). The work discussed here differs from that approach, in that here we use a second Kohonen network that clusters the already clustered sensory information encoded in the first layer network, rather than using sequences of raw sensory data.

There is also related work in the area of assessment of robot performance ([12]). Notably the work of Lee and Recce is relevant to the experiments reported here. In their case ([7]) however, mapbuilding performance was measured against a hand-crafted "optimal" performance (i.e. an absolute comparison). In contrast to this, we perform quantitative comparisons between two algorithms performing under identical conditions (i.e comparing against a relative standard).

1.3 "Anchoring" through exteroception

We contend that any navigation system that is to be used on a real robot, beyond the immediate vicinity of a "home location", and over extended periods of time, has to be anchored in exteroception[2], for the following reasons. Proprioceptive systems are subject to uncorrectable drift errors, which means that the anchor points of such navigation systems will change over time and introduce navigation errors that are not correctable through proprioception alone. Only through calibration using exteroception can this error be removed (a recent example of such a system is presented in [15]). Drift errors are inherent and not an engineering problem - more precise wheel encoders will simply mean that the proprioception-based navigation system will function correctly for a longer period of time. Eventually, however, it *will* fail, which is our main reason for investigating landmark-based robot self-localisation.

2 The episodic mapping algorithm

2.1 Introduction

The fundamental principle behind an episodic mapping mechanism is to take into account both the sensory perceptions (perceptual signatures) at the robot's current location, as well as a history of the robot's past perceptions. This allows the disambiguation of two locations with identical perceptual signatures, if the perceptions preceding those two locations differ. A localisation system based on this method has successfully been used at Manchester for localisation when the robot has got completely lost ([4]).

[1] Recordings from single units in and around the rat hippocampus show strong correlation between a freely moving animal's location and cell firing ([2]). Certain cells only fire when the rat is in a restricted portion of its environment.

[2] Sensory stimuli impinging on the robot from the outside, as opposed to proprioception (using internal sensory stimuli).

2.2 Why topological mappings?

There are two main shortcomings of any episodic mapping mechanism: firstly, it is dependent upon robot motion along a fixed path (or a few fixed paths), because a unique and repeatable sequence of perceptions is required to identify a location. Secondly, localisation is affected by "freak perceptions"[3] for a much longer time than in a navigation system based on the current perception only, because any erroneous (freak) perception is retained for n timesteps, where n is the number of past perceptions used for localisation. Such freak perceptions do not normally occur in computer simulations, but they occur frequently when a real robot interacts with the real world, because of sensor properties (e.g. specular reflection of sonar signals), sensor noise, or electronic noise.

The episodic mapping algorithm proposed here specifically addresses this question of how to cope with freak perceptions when using an episodic mapping mechanism. Because the mapping algorithm is based on topological (i.e. similarity preserving) mappings of both sensory perceptions and episodes of perceptions, freak perceptions will affect localisation performance to a far lesser degree than they would in systems that are dependent on uncorrupted perception of each one of n sensory images.

2.3 Topological clustering using self-organising feature maps

To cluster incoming sensory information, in both the static and the episodic mapbuilding paradigm, a self-organising feature map ([5]) was used.

The self-organising feature map (SOFM) is an example of an artificial neural network that performs a topological clustering of its input data using an unsupervised learning mechanism. The network consists of one layer of cells typically arranged as a two dimensional grid. Figure 1 shows the basic structure of such a network.

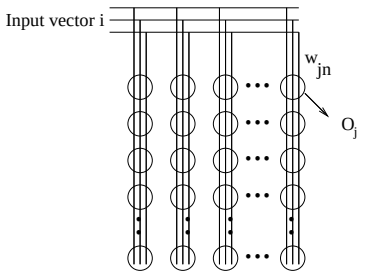

Figure 1: Structure of the SOFM

The input vector $\vec{i}$, containing sensory information, is presented to each output unit j of the network. Each unit j has a normalised weight vector $\vec{w}_j$.

The continuous-valued output o_j of unit j is found by calculating the weighted sum of its inputs, given by:

$$o_j = \sum_{k=1}^{n} w_{jk} i_k = \vec{w}_j \cdot \vec{i}, \tag{1}$$

with n being the number of elements in the input vector and the weight vectors.

The initial state of the network uses randomised values for the weights. Therefore, when a stimulus is presented to the network, one cell of the network will respond more strongly than the others to that particular input vector (see equation 1).

The weight vector of this "winning" unit as well as those of the eight neighbouring units are then changed according to equation 2:

$$\Delta w_{jk} = \alpha(i_k - w_{jk})$$

and

[3] Spurious sensory perceptions caused by intermittent processes such as specular reflections or sensor crosstalk.

$$w_{jk}(t+1) = w_{jk}(t) + \Delta w_{jk} \tag{2}$$

where α is the learning rate. Typical values for this parameter are in the range 0.2 - 0.5. A value of $\alpha = 0.25$ (constant over time) was used in the experiments presented here. Weight vectors are normalised again after being adjusted.

As this process continues the network organises into a state whereby dissimilar input vectors/patterns map onto different regions of the network, whilst similar patterns are clustered together in groups: a topological map of the input space develops.

When the network has settled, distinct physical locations will map onto distinct regions of the network[4], whilst similar perceptual patterns cluster together in a region. To achieve this, no symbolic representations have been created, and the robot is mapping its environment "as it sees it".

In this way, regions of the network can be seen as representing 'perceptual landmarks' within the robot's environment, and map response can then be used for localisation.

2.4 First layer: Static mapping

The mapbuilding component used in the static mapbuilding paradigm is a two dimensional SOFM of m x m units (m=9 or m=12 in our experiments). The input to the SOFM consists of the robot's 16 infrared sensor readings. Whenever the robot has moved more than 25 cm, a 16-element input vector, containing the raw sensor readings from the robot's infrared sensors, is generated. The robot's turret maintains a constant orientation throughout the experiment to eliminate any influence of the robot's current orientation at a particular location, resulting in a unique sensory perception at each location, irrespective of the angle at which the robot approached that location. Note that the 16-element input vector does not actually convey much information about the current location of the robot. A coarse input vector such as this was deliberately chosen for these investigations to produce perceptual aliasing — the aim here is to find ways of achieving localisation even under very difficult circumstances.

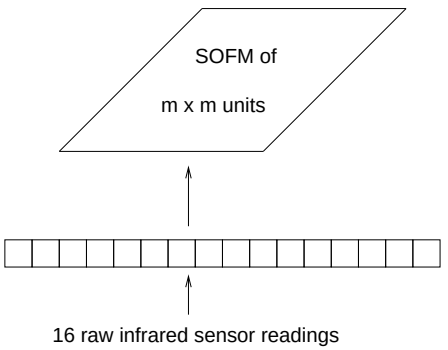

Figure 2: THE STATIC MAPPING MECHANISM: THE SOFM CLUSTERS THE CURRENT SENSORY PERCEPTION AND THUS GENERATES THE STATIC MAPPING.

As the robot moves through its environment, controlled by the operator, sensor readings are obtained and input vectors fed into the SOFM. The network clusters these perceptions according to their similarity and frequency of occurrence.

2.5 Second layer: Episodic mapping

The episodic mapping paradigm uses two layers of self organising feature maps (see figure 3). Layer one is the layer described in subsection 2.4.

Layer two is also a two-dimensional SOFM of k x k units (k=9 or k=12 in our experiments), it is trained using an input vector of m^2-element length. All elements of this vector

[4]Provided these distinct physical locations also have distinct perceptual signatures, i.e. no perceptual aliasing occurs.

are set to zero, apart from the last τ centres of excitation of layer one, which are set to "1". The value of the ("history") parameter τ was varied in our experiments.

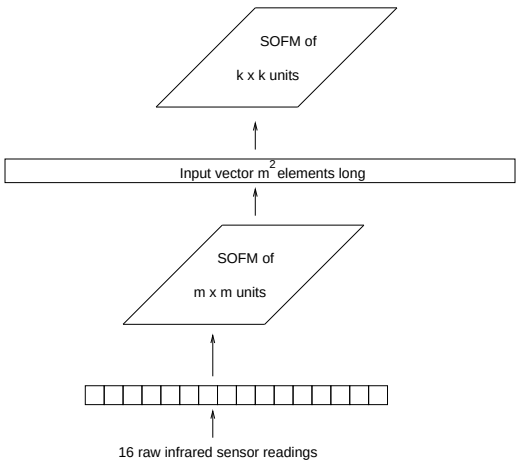

Figure 3: THE EPISODIC MAPBUILDING MECHANISM: THE FIRST LAYER SOFM CLUSTERS THE CURRENT SENSORY PERCEPTION, THE SECOND LAYER SOFM CLUSTERS THE LAST τ PERCEPTIONS AND THUS GENERATES THE EPISODIC MAPPING.

Precedence relationships between these excitation centres are not encoded. This means that the second layer of the dynamic mapbuilder — the output layer — uses information about the perceptual signature of the current location as well as temporal cues (in this case τ perceptions of the robot before it arrived at the current location), but no precedence relationships between perceptions.

The second layer SOFM performs a clustering of the last τ excitation centres observed in layer 1 in a topology-preserving manner. As the output of layer 1 is a topological mapping of all sixteen infrared sensor signals and the output of layer 2 is again a topological map, the response of the episodic mapping system is far less sensitive to freak sensory perceptions than a mapping system that uses episodes of raw sensory data as input (such as the system presented in [10]). This is a desirable property, because freak sensor signals occur regularly in real robotic systems, and mapping systems dependent on constant and reliable sensory perception therefore tend to be brittle.

3 Measuring localisation performance

3.1 An entropy-based measure of mapbuilding and localisation performance

To assess the quality of mappings achieved by the episodic mapping mechanism, and to quantify the influence of individual parameters, we used the following entropy-based measure of performance.

First, a contingency table such as the one shown in table 1 (appendix A) was obtained by logging map response[5] versus robot location[6]. To obtain the contingency table, physical space and map space were partitioned into regions ("bins"), and whenever the robot was anywhere within a region of space, obtaining a map response from a particular map region, the corresponding field of the contingency table was incremented. For example, table 1 shows that the first category of map response was perceived 24 times in total, 18 times at location A, and 6 times at location C.

In the experiments reported here, binning was performed by dividing the physical space into units of equal size. This method results in a uniform spatial resolution of the localisation

[5] "Map response" denotes the most active unit of the second layer SOFM.

[6] The robot's physical location was obtained through dead reckoning. Over relatively short distances, and logging only regions of space, rather than precise location, this is accurate enough for the experiments described here.

algorithm — a desirable property — but has the disadvantage of resulting in bin entries of different magnitude, because the robot occupies some regions of space more often than others. This unequal distribution can lead to computational problems in the contingency table analysis discussed below.

The alternative would be to construct bin sizes according to duration of travel in each bin. This would result in a more even distribution of data points per bin, but has the disadvantage that spatial resolution is no longer uniform over physical space. For this reason we adopted a spatially uniform binning method.

Entropy-based measures were then used to determine the strength of association between map response and robot location. The entropy (or average information) provides a measure of the probability that a particular signal is received, given particular contextual information. For example, if the system's response to a particular stimulus is known with absolute certainty (probability 1), then the entropy (i.e. average information) of having perceived that particular stimulus is obviously 0.

For localisation, entropy can serve as a quality metric in the following way: if any response R of the localisation system corresponds with exactly one location L in the physical world, then the entropy $H(L|R)$ is zero for that case (the "perfect" map). The larger $H(L|R)$, the larger the uncertainty that the robot is at a particular location, given some system response R.

$H(L|R)$ is defined as follows ([13]):

$$H(L|R) = -\sum_{l,r} p_{l,r} ln \frac{p_{l,r}}{p_{l.}} \tag{3}$$

with

$$p_{l,r} = \frac{N_{l,r}}{N} \tag{4}$$

and

$$p_{l.} = \frac{N_{l.}}{N}. \tag{5}$$

N is the total number of events recorded in the contingency table, $N_{l,r}$ is the number of occurrences of response r at location l, and $N_{l.}$ is the number of occurrences of any response at location l.

$H(L|R)$ can therefore be used as a metric to determine the suitability of the obtained mapping for localisation. If $H(L|R)$ is zero, perfect localisation can be achieved, i.e. a particular system response R will indicate with absolute certainly where the robot is in the world. If $H(L|R)$ is non-zero, some ambiguity regarding the robot's current location exists, the larger $H(L|R)$, the larger the ambiguity.

This measure allows quantitative comparison of two or more mapping paradigms being compared under *identical* experimental circumstances. In particular, bin sizes must be identical for all experiments. In other words, the metric allows the comparison between mapping systems, but does not provide an absolute standard which is experiment-independent. This quality metric is a useful measure for the experiments presented in this paper, because the fundamental question asked is which of two mapping paradigms performs better, under identical experimental conditions.

3.2 Evaluation of results

In all experiments reported here, we used the quality metric defined in subsection 3.1 to determine the quality of a mapping: the lower the entropy $H(L \mid R)$, the higher the map quality. A 'perfect' map has an entropy $H(L \mid R)$ of zero.

In all our experiments, the "static mapping", using a single layer self-organising feature map (see subsection 2.4), and the "episodic mapping", using a twin layer self-organising feature map (see subsection 2.5) are directly compared. Of the former we know that it does provide a feasible method for mobile robot localisation ([9, 11, 10]). The question is: does

the episodic mapping paradigm produce better maps, with respect to the criterion discussed in section 3.1?

4 Experiments

4.1 Experimental method

In order to be able to modify individual parameters and to examine the results under *identical* conditions (without resorting to computer simulations), we used *recorded* sensor data obtained by manually driving a Nomad 200 mobile robot (see figure 4) through two different environments. Both the static and the episodic mapping schemes were then applied to the same data. This ensures that the input data to each mapping scheme is *identical* throughout all experiments.

Figure 4: The Manchester Nomad 200 mobile robot, "FortyTwo". The robot has sixteen sonar sensors, sixteen infrared sensors, camera, compass, tactile sensors and onboard odometry sensors.

Furthermore, experimental parameters such as network size and bin size were varied, to determine their influence upon localisation performance.

4.2 Experiments in environment 1

4.2.1 Experimental procedure

Here, the robot was manually driven along a (more or less) fixed path in an environment containing brick walls, cloth screens and cardboard boxes. The whole route was traversed six times, and 366 data points in total were obtained, containing the robot's 16 infrared sensors and location in (x,y) coordinates (see figure 5).

Of the 366 data points, 120 were used for the initial training of the networks[7], i.e. the mapbuilding phase, and the remaining 246 data points were used for the evaluation of the localisation ability.

Localisation performance was then assessed using the performance metric described in subsection 3.1.

4.2.2 Experimental results in environment 1

Using the data obtained in environment 1, we compared three different implementations of the episodic mapping mechanism with the static mapping mechanism. In all cases, the episodic mapping mechanism outperformed the static mapping mechanism.

[7] The first layer network only was trained with the first 20 data points, the remaining 100 data points were used to train both nets.

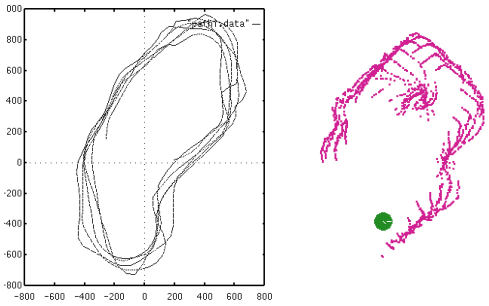

Figure 5: Actual trajectory taken by robot in environment 1 (left), and accumulated infrared sensor readings (right, environment 1 "as the robot sees it"). Dimensions in the diagram are in units of 2.5mm.

In the first experiment, we used self organising feature maps of 12x12 units. To construct the contingency table for this experiment, the network output space was partitioned into 16 bins, and the physical space in which the robot operated was partitioned into 15 bins (see figure 6).

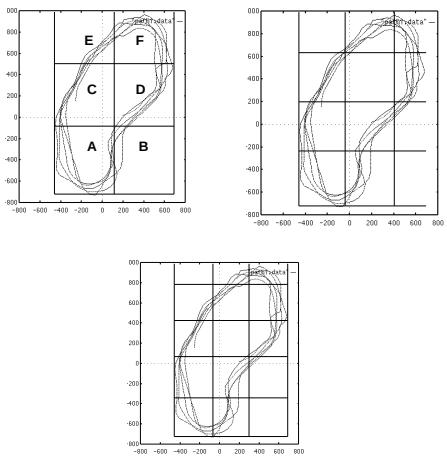

Figure 6: Partitioning of environment 1 into 6, 12 and 15 location bins respectively. Dimensions are given in units of 2.5mm.

The results obtained are shown in figure 7.

It shows that for all values of τ up to $\tau = 7$ episodic mapping produces a lower $H(L \mid R)$ than static mapping, that is better localisation performance (the case $\tau = 1$ has been included for verification of the experimental results: a history length of 1 means that only the current perception is taken into account. This is essentially identical to the static case, with the difference that in the static case the sixteen sonar readings provide more information than the one excitation centre of layer 1 that is used as input to layer 2 in episodic mapping. The expected result, therefore, is that episodic mapping with $\tau = 1$ always produces slightly worse results than static mapping — as is indeed the outcome in all experiments bar experiment 4, where both methods produce very similar results).

The conclusion to draw from this experiment, then, is that episodic mapping produces better localisation performance than static mapping, up to a certain maximum value of τ, indicating that too much episodic information is confusing, rather than helpful.

In the second experiment, the spatial resolution (i.e. the localisation precision) was reduced to 12 bins. As would be expected, localisation performance improved (because there is less opportunity for error). The difference to experiment 1, however, was small (figure 8), and essentially, the findings of experiment 2 confirm those of experiment 1.

In the final experiment conducted in environment 1, we used smaller networks, and reduced the bins size further. The results of this experiment are shown in figure 9, a

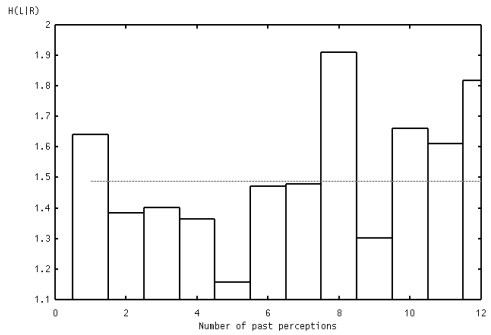

Figure 7: Experiment 1. Results obtained in environment 1, using 12x12 networks, partitioned into 16 bins. The physical space of 2.87m x 4.30m was divided into 15 bins (see figure 6). The single layer network achieves H(L|R)=1.49 in this experiment (indicated by horizontal line).

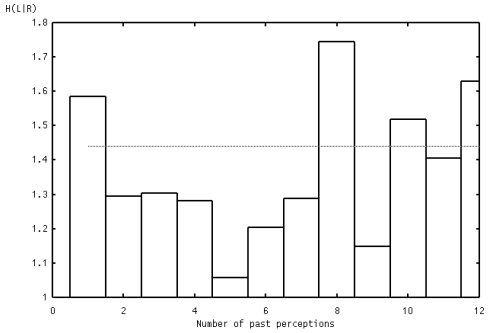

Figure 8: Experiment 2. Results obtained in environment 1, using 12x12 networks, partitioned into 16 bins. The physical space of 2.87m x 4.30m was divided into 12 bins (see figure 6). The single layer network achieves H(L|R)=1.44 in this experiment (indicated by horizontal line).

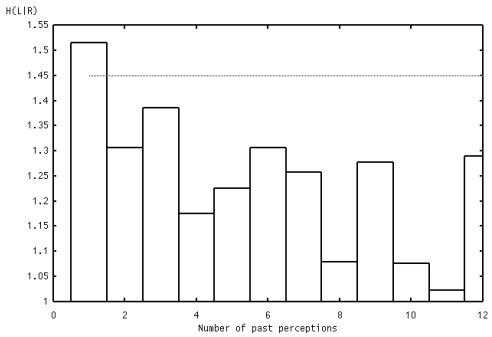

Figure 9: Experiment 3. Results obtained in environment 1, using a 9x9 network, partitioned into 9 bins. The physical space of 2.87m x 4.30m was divided into 6 bins (see figure 6). The single layer network achieves H(L|R)=1.45 in this experiment (indicated by horizontal line).

discussion of results follows in section 4.4. The contingency table for this experiment is shown in table 1 in appendix A. Again, the earlier findings are confirmed. In this case, episodic mapping produces better localisation performance than static mapping in all cases, regardless of the value of τ.

4.3 Experiments in environment 2

4.3.1 Experimental procedure

For a second set of experiments, a route in a different environment containing cluttered furniture (desks, chairs), brick walls, and open space was traversed nine times, and 456 data points in total were obtained by manually driving the robot. 160 of these data points were used for training the networks[8], the remaining 296 data points were used to evaluate localisation performance.

Environment 2 was less structured than environment 1, in that it contained a larger variety of perceptually distinct objects, and more clutter. It was also bigger, and the robot's path in it is longer than in environment 1. Figure 10 shows the robot's path through this environment, and the robot's perception of it.

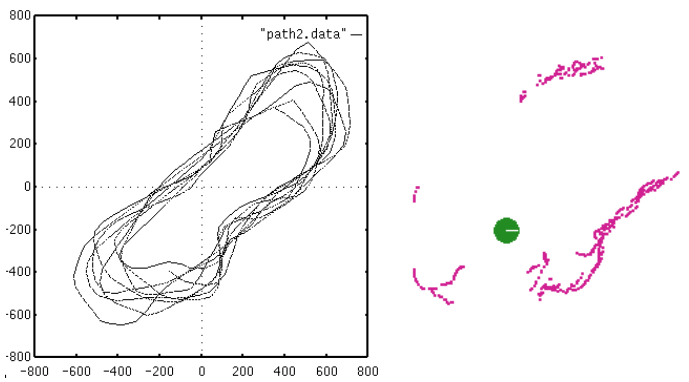

Figure 10: Robot trajectory in environment 2 (left) and accumulated infrared sensor readings obtained by the robot in environment 2 (right, environment 2 "as the robot sees it"). Dimensions are in units of 2.5mm.

[8]The first 20 data points were used for training the first layer network alone.

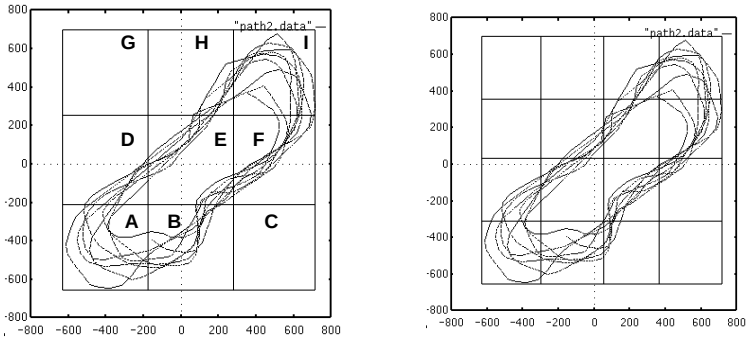

Figure 11: Partitioning of environment 2 into 9 and 16 location bins respectively. Dimensions are given in units of 2.5mm.

4.3.2 Experimental results in environment 2

In the first experiment conducted in environment 2, we used networks of 9x9 units, dividing both network space and physical space into nine regions (figure 11) to build the contingency table.

The results of this experiment are shown in figure 12. As in the experiments in environment 1, episodic mapping produces better localisation performance than static mapping, in this case for all values of τ (the subsequent experiments in environment 2 confirm, however, that there is a maximum useful length of τ, beyond which the performance of episodic mapping decreases).

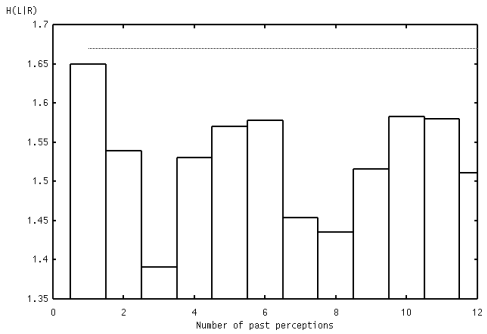

Figure 12: Experiment 4. Results obtained in environment 2, using 9x9 networks, partitioned into 9 bins. The physical space of 3.37m x 3.36m was divided into 9 bins (see figure 11). The single layer network achieves H(L|R)=1.67 in this experiment (indicated by horizontal line).

We then evaluated localisation performance using 12x12 unit networks, at increased spatial resolution. The results are shown in figure 13. Here, the maximum useful value of τ is eight previous perceptions, with optimum performance for $\tau = 4$. And again, episodic mapping outperforms static mapping by a wide margin.

Finally, we reduced localisation precision to nine distinct regions, and obtained the results shown in figure 14. The contingency table for this experiment is shown in table 3 in appendix A. This final experiment confirms earlier findings about the performance of both mechanisms, and the existence of a maximum useful value of τ.

4.4 Discussion of results

4.4.1 Analysis of contingency tables using the entropy-based metric

In both environments, and for all experiments conducted, episodic mapping outperforms static mapping, provided not too long a history is incorporated into the input vector. Here,

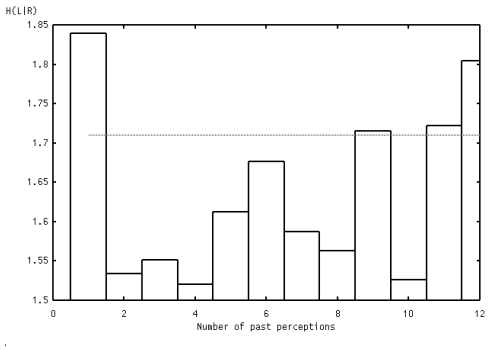

Figure 13: Experiment 5. Results obtained in environment 2, using 12x12 networks, partitioned into 16 bins. The physical space of 3.37m x 3.36m was divided into 16 bins (see figure 11). The single layer network achieves H(L|R)=1.71 in this experiment (indicated by horizontal line).

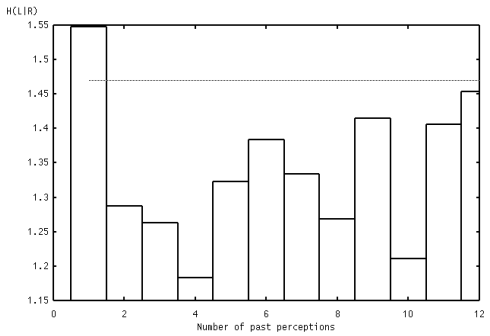

Figure 14: Experiment 6. Results obtained in environment 2, using 12x12 networks, partitioned into 16 bins. The physical space of 3.37m x 3.36m was divided into 9 bins (see figure 11). The single layer network achieves H(L|R)=1.47 in this experiment (indicated by horizontal line).

for all $2 \leq \tau \leq 7$ episodic mapping performs better than static mapping. The explanation for this observation is obvious: Exploiting temporal information does help to disambiguate sensory perceptions obtained in different locations, but including too much episodic information confuses, rather than aids, and localisation performance decreases as a consequence.

The optimum value of τ is dependent on bin sizes, but lies between 3 and 5 in most cases. Note that the optimum value can be determined in real time by the robot itself, as the contingency table and $H(L \mid R)$ is available to the robot.

The choice of network size and bin size appear to be non-critical. In all cases episodic mapping can outperform static mapping.

The performance metric introduced in section 3.1 can be applied to any mapbuilding system that generates categorical data, and therefore provides a tool of comparing different paradigms, as well as determining the influence of any process parameters, independent of the actual paradigm used.

4.4.2 Direct analysis of the contingency tables

The statistical measure defined in section 3.1 determines the strength of the correlation between map response and the robot's physical location. Besides taking this measure as an indication of localisation performance, we also looked at the contingency tables directly to determine how useful both static and episodic mapping would actually be for localisation.

The four contingency tables of experiments 3 and 6 (see appendix A) are redrawn in figure 15, here indicating entries which comprise more than 30% and more than 50% of a row's entries graphically.

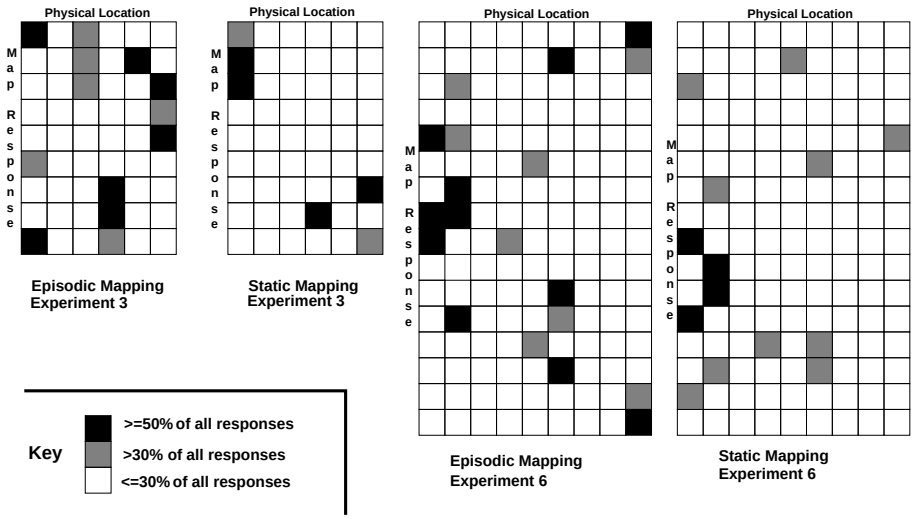

Figure 15: Assessment of the suitability of contingency tables for robot self localisation. Comparing the tables for static and episodic mapping demonstrates that episodic mapping provides a clearer correlation between physical location and map response. (The table values are shown in appendix A).

When a particular map response (rows in contingency table) is obtained, good localisation is possible if along one row one field accounts for the majority of the entries along the entire row: this field indicates the physical location (columns in contingency table) the robot is at.

As can be seen from figure 15, episodic mapping produces more "decisive" contingency tables. In the contingency tables for static mapping, for many rows no column attracts a large number of entries — in these cases no unambiguous localisation is possible — whereas episodic mapping generates exactly one strong candidate in every column (with very few exceptions).

This visual analysis of the contingency tables therefore confirms the findings of subsection 4.4.1.

5 Summary and Conclusion

The ability to navigate is of paramount importance for autonomous mobile robots for every task that requires goal-directed motion (i.e. all tasks that cannot be achieved by random movement). Localisation, i.e. establishing one's position within a frame of reference (a "map") is the most fundamental competence required for navigation. Mobile robot navigation, and in particular self localisation, is a hard problem, because meaningful information about the robot's position in space has to be obtained from noisy, redundant and ambiguous data. In addition, the problems of self-localisation are compounded by perceptual aliasing, the fact that most sensory perceptions of a mobile robot are not uniquely associated with exactly one position in the real world (there is usually no one-to-one mapping between physical location and sensory perception).

This paper addresses these problems by presenting a localisation mechanism for autonomous mobile robots that uses spatial *and* episodic information to establish the robot's position in the world. In the first (static mapping) stage of the process, raw sensory perceptions of the robot are processed using an unsupervised, self-organising clustering technique. The last τ perceptions of this first layer are then clustered again to encode episodic information. Through this unsupervised, self-organising process meaningful internal representations of a mobile robot's environment ("mappings") emerge, without any external intervention.

This process can be interpreted as the *emergence of an analogy*, as the internal representation resembles the robot's environment in some particulars (i.e. the robot's sensory

perception of the world is represented), without actually being identical to the real world ([14]).

An entropy-based quality metric was used to compare the two localisation paradigms, and to determine the influence of individual process parameters upon the final map quality. It could be shown that, provided not too much episodic information is included (i.e using a short term memory, rather than a medium or long term one), episodic mapping outperforms static mapping, irrespective of experimental parameters such as bin sizes or history length.

The main advantage of episodic *topological* mappings as presented here is that this method is less sensitive to freak sensor signals than an episodic mapping scheme that uses *raw* sensor data. This is advantageous for actual robot localisation, because freak sensor signal are a common occurrence when real robots are used.

There are a number of unanswered questions, subject to future research. We have shown that a maximum useful episode length exists, beyond which episodic mapping produces worse results than static mapping. The information of what constitutes the optimal episode length τ is actually available to the algorithm through the computation of the entropy $H(L \mid R)$, it is therefore conceivable that the robot determines the optimal episode length automatically. This approach is subject to ongoing research.

It is also conceivable that the optimal episode length τ can be used to characterise robot-environment interaction: if identical mapping mechanisms and identical robots are used in different environments, presumably the environment with a higher $H(L \mid R)$ is the more "difficult" environment, containing fewer perceptual cues regarding the robot's position. Characterising environments is desirable to allow independent replication of mobile robotics experiments ([12]).

Lastly, although we use previous perceptions for the episodic mapping, we do not encode the temporal ordering of those perceptions. It is likely that using this additional information would produce even better mappings. Again, this is subject to ongoing research.

References

[1] Rodney A. Brooks, *A Robust Layered Control System for a Mobile Robot*, MIT AI Memo No. 864, September 1985.

[2] N. Burgess and J. O'Keefe, *Neuronal computations underlying the firing of place cells and their role in navigation*, Hippocampus 7:749-762 (1996).

[3] N. Burgess, J. O'Keefe and M. Recce, *Using hippocampal 'place cells' for navigation, exploiting phase coding*, in Hanson, Giles and Cowan (eds.), *Advances in neural information processing systems 5*, Morgan Kaufmann 1993.

[4] Tom Duckett and Ulrich Nehmzow, *Mobile Robot Self-Localization and Measurement of Performance in Middle-Scale Environments*, J. Robotics and Autonomous Systems, Vol 24, Nos. 1-2, 1998.

[5] Teuvo Kohonen, *Self Organization and Associative Memory*, Springer Verlag, Berlin, Heidelberg, New York, 2nd edition, 1988.

[6] Andreas Kurz, *Constructing maps for mobile robot navigation based on ultrasonic range data*, IEEE Trans Systems, Man and Cybernetics B, Vol 26 No 2 1996.

[7] David Charles Lee, *The map-building and exploration strategies of a simple sonar-equipped mobile robot; an experimental quantitative evaluation*, PhD thesis, University College London, 1995.

[8] Maja Mataric, *Navigating with a Rat Brain: A Neurobiologically-Inspired Model for Robot Spatial Representation*, in Jean-Arcady Meyer and Stuart Wilson (eds.), *From Animals to Animats*, MIT Press 1991.

[9] Ulrich Nehmzow and Tim Smithers, *Mapbuilding using Self-Organising Networks*, in: Jean-Arcady Meyer and Stewart Wilson (eds.), *From Animals to Animats*, MIT Press, Cambridge Mass. and London, England, 1991, pp. 152-159.

[10] Ulrich Nehmzow, Tim Smithers and John Hallam, *Location Recognition in a Mobile Robot using Self-Organising Feature Maps*, in G. Schmidt (ed.), *Information Processing in Autonomous Mobile Robots*, Springer Verlag, Berlin, Heidelberg, New York, 1991, pp. 267-277.

[11] Ulrich Nehmzow and Tim Smithers, *Using Motor Actions for Location Recognition*, in F. Varela and P. Bourgine (eds.), *Toward a Practice of Autonomous Systems*, MIT Press, Cambridge Mass. and London, England, 1992, pp. 96-104.

[12] Ulrich Nehmzow and Michael Recce (eds.), Scientific Methods in Mobile Robotics, Special Issue of J. Robotics and Automation, Vol 24, Nos. 1-2, 1998.

[13] W. Press, S. Teukolsky, W. Vetterling and B. Flannery, *Numerical recipes in C*, Cambridge University Press, 1992.

[14] Merriam-Webster online dictionary, `http://www.m-w.com/`.

[15] B. Yamauchi and R. Beer, *Spatial learning for navigation in dynamic environments*, IEEE Trans Systems, Man and Cybernetics B, Vol 26 No 3, 1996.

[16] U. Zimmer, *Self-localisation in dynamic environments*, submitted to IEEE/SOFT Workshop BIES'95, Tokyo, May 95.

A Contingency tables

	Physical Location					
	A	B	C	D	E	F
M	**18 (67%)**	0 (0%)	6 (33%)	0 (0%)	0 (0%)	0 (0%)
a	0 (0%)	0 (0%)	3 (33%)	0 (0%)	**6 (67%)**	0 (0%)
p	2 (9%)	0 (0%)	8 (35%)	0 (0%)	0 (0%)	**13 (57%)**
	1 (3%)	0 (0%)	7 (24%)	0 (0%)	8 (28%)	**13 (45%)**
R	1 (2%)	1 (2%)	0 (0%)	9 (18%)	12 (24%)	**27 (54%)**
e	19 (43%)	4 (9%)	13 (30%)	3 (7%)	5 (11%)	0 (0%)
s	0 (0%)	0 (0%)	1 (25%)	**3 (75%)**	0 (0%)	0 (0%)
p	0 (0%)	0 (0%)	0 (0%)	**11 (79%)**	0 (0%)	3 (21%)
.	**16 (55%)**	1 (3%)	0 (0%)	9 (31%)	0 (0%)	3 (10%)

Table 1: Contingency table for experiment 3 (episodic mapping, $\tau = 11$). The eight situations in which first and second candidate location (given a particular map response) differ by more than 150% are shown in heavy type. Percentages of localisation obtained given response are shown in brackets.

	Physical Location					
	A	B	C	D	E	F
M	**14 (44%)**	0 (0%)	4 (13%)	6 (10%)	3 (9%)	5 (16%)
a	**11 (50%)**	2 (9%)	1 (5%)	3 (14%)	2 (9%)	3 (14%)
p	**17 (53%)**	2 (6%)	4 (13%)	5 (16%)	1 (3%)	3 (9%)
	3 (11%)	0 (0%)	8 (30%)	4 (15%)	5 (19%)	7 (26%)
R	8 (30%)	1 (4%)	3 (11%)	6 (22%)	2 (7%)	7 (26%)
e	5 (21%)	1 (4%)	4 (17%)	4 (17%)	4 (17%)	6 (25%)
s	2 (7%)	0 (0%)	6 (21%)	3 (11%)	3 (11%)	**14 (50%)**
p	1 (4%)	0 (0%)	3 (11%)	**14 (52%)**	6 (22%)	3 (11%)
.	2 (8%)	0 (0%)	6 (23%)	1 (4%)	5 (19%)	**12 (46%)**

Table 2: Contingency table for experiment 3, static mapping. The six situations in which first and second candidate location (given a particular map response) differ by more than 150% are shown in heavy type. Percentages of localisation given response are shown in brackets. See also table 1.

	Physical Location								
	A	B	C	D	E	F	G	H	I
	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	3 (25)	**9 (75)**
M	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	**9 (64)**	0 (0)	0 (0)	5 (36)
a	0 (0)	7 (39)	0 (0)	0 (0)	5 (28)	3 (17)	0 (0)	2 (11)	1 (6)
p	3 (16)	2 (11)	0 (0)	2 (11)	5 (26)	5 (26)	0 (0)	1 (5)	1 (5)
	2 (67)	1 (33)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
	0 (0)	5	0 (0)	3 (13)	**10 (44)**	3 (13)	0 (0)	2 (9)	0 (0)
R	0 (0)	**3 (100)**	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
e	5 (50)	5 (50)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
s	**23 (58)**	2 (5)	0 (0)	13 (33)	1 (3)	1 (3)	0 (0)	0 (0)	0 (0)
p	3 (14)	8 (38)	0 (0)	0 (0)	0 (0)	6 (29)	0 (0)	0 (0)	4 (19)
o	2 (13)	0 (0)	0 (0)	0 (0)	4 (27)	8 (50)	0 (0)	0 (0)	0 (0)
n	0 (0)	7 (54)	0 (0)	0 (0)	1 (8)	5 (39)	0 (0)	0 (0)	0 (0)
s	0 (0)	0 (0)	0 (0)	1 (5)	8 (40)	3 (15)	0 (0)	4 (20)	4 (20)
e	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	**5 (57)**	0 (0)	2 (22)	2 (22)
	6 (29)	0 (0)	0 (0)	4 (19)	3 (14)	1 (5)	0 (0)	0 (0)	7 (33)
	4 (11)	0 (0)	0 (0)	1 (3)	6 (17)	0 (0)	0 (0)	7 (19)	**18 (50)**

Table 3: Contingency table for experiment 6 (episodic mapping, $\tau = 4$). Map responses that indicate a localisation with more than 150% lead over the second candidate location are shown in heavy type. Percentages of localisation given response are shown in brackets. Compare with table 4, which shows the results obtained in the same experiment, but using the single layer network map.

	Physical Location								
	A	B	C	D	E	F	G	H	I
	10 (27)	3 (8)	0 (0)	4 (11)	3 (8)	2 (5)	0 (0)	4 (11)	11 (30)
M	4 (20)	1 (5)	0 (0%)	0 (0%)	7 (35)	4 (20)	0 (0%)	1 (5)	2 (10)
a	8 (35)	2 (9)	0 (0%)	0 (0%)	2 (9)	7 (30)	0 (0%)	0 (0%)	4 (17)
p	0 (0%)	4 (16)	0 (0%)	3 (12)	5 (20)	6 (24)	0 (0%)	4 (16)	3 (12)
	1 (3)	0 (0%)	0 (0%)	3 (9)	10 (30)	1 (3)	0 (0%)	3 (9)	**15 (46)**
	2 (13)	2 (13)	0 (0%)	0 (0%)	1 (7)	6 (40)	0 (0%)	0 (0%)	4 (27)
R	2 (25)	3 (38)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (25)	1 (13)
e	1 (9)	2 (18)	0 (0%)	0 (0%)	2 (18)	0 (0%)	0 (0%)	3 (27)	3 (27)
s	**10 (56)**	2 (11)	0 (0%)	2 (11)	1 (6)	0 (0%)	0 (0%)	2 (11)	1 (6)
p	1 (10)	**6 (60)**	0 (0%)	0 (0%)	0 (0%)	2 (20)	0 (0%)	0 (0%)	1 (10)
o	2 (25)	**5 (63)**	0 (0%)	0 (0%)	1 (13)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
n	2 (50)	1 (25)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (25)
s	3 (12)	0 (0%)	0 (0%)	10 (39)	4 (15)	8 (31)	0 (0%)	1 (4)	0 (0%)
e	0 (0%)	8 (42)	0 (0%)	0 (0%)	0 (0%)	8 (42)	0 (0%)	0 (0%)	3 (16)
	6 (32)	5 (26)	0 (0%)	0 (0%)	4 (21)	3 (16)	0 (0%)	0 (0%)	1 (5)
	5 (25)	4 (20)	0 (0%)	2 (10)	4 (20)	3 (15)	0 (0%)	1 (5)	1 (5)

Table 4: Contingency table for experiment 6, static mapping. Map responses that indicate localisation with more than 150 lead over the second candidate location are shown in bold. Percentages of localisation given response are shown in brackets. See also table 3.