An Episodic Mapping Algorithm for Mobile Robot Self-Localisation: "Meaning" through Self-Organisation[1]

Ulrich Nehmzow
Department of Computer Science
Manchester University
Manchester M13 9PL
United Kingdom
u.nehmzow@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 ambiguous and redundant sensory information is used to generate meaningful information to a mobile robot.

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").

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.

The only information available to a mobile robot to achieve this fundamental prerequisite to navigation 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 (see section 2) that allows the robot to identify its current position within the mapping of the environment, based on current and preceding perceptions of the world. Using a quality metric we compare the episodic mapping with a static mapping paradigm ([11, 13, 12], which does not exploit temporal information.

[1]International workshop on "Computation for Metaphors, Analogy and Agents", Aizu 1998.

1.2 Related Work

The mapping used here is similar in many ways to hippocampal mappings found in rats. In particular, place cells[2] 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 ([10, 11]).

Self-organising mechanisms for static sensor signal clustering have been used for robot localisation before ([11, 13, 8, 20]): 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 ([12]). 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.

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[3], 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 [19]). 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 perceptual signature of 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 ([5]).

2.2 Why topological mappings?

There are two main shortcomings to 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" for a much longer time than in a navigation system based on the current perception only, because the erroneous (freak) perception is retained for n timesteps, where

[2]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.

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

n is the number of past perceptions used for localisation. Such freak perceptions do not normally occur in computer simulations, but they can occur frequently when a real robot interacts with the real world, because of idiosyncratic sensor properties (e.g. 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 signals will indeed result in distorted perceptions, but even those distorted perceptions are normally still similar enough to a target perception to remain identifiable.

2.3 First layer: Static mapping

The mapbuilding component used in the static mapbuilding paradigm is a two dimensional self-organising feature map (SOFM, [7]) 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 to the first self-organising feature map is generated. This input vector contains the raw sensor readings from the robot's infrared sensors. 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 simulations to produce perceptual aliasing — the aim here is to find ways of achieving localisation even under very difficult circumstances.

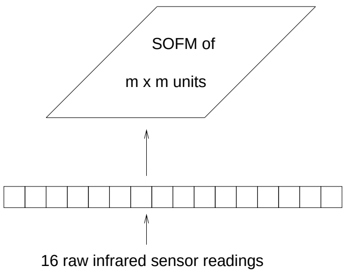

Figure 1: THE STATIC MAPPING MECHANISM: THE SOFM CLUSTERS THE CURRENT SENSORY PERCEPTION AND THUS GENERATES THE STATIC MAPPING.

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.

2.4 Second layer: Episodic mapping

The episodic mapping paradigm uses two layers of self organising feature maps (see figure 2). Layer one is the layer described in subsection 2.3.

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

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.

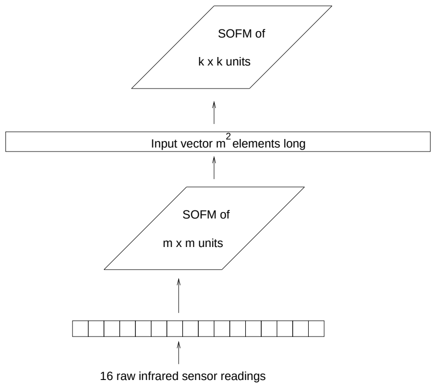

Figure 2: THE EPISODIC MAPBUILDING MECHANISM: THE FIRST LAYER SOFM CLUSTERS THE CURRENT SENSORY PERCEPTION, THE SECOND LAYER SOFM CLUSTERS THE LAST τ PERCEPTIONS AND THUS GENERATES THE EPISODIC MAPPING.

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 [12]). This is a desirable property, because freak sensor signals occur regularly in real robotic systems, and mapping system dependent on constant and reliable sensory perception would therefore be very brittle.

3 Experiments

3.1 Experimental Procedure

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 through two different environments. The different mapping schemes were then applied to the same data. This ensures that the input data to each mapping scheme is *identical* throughout all experiments.

3.2 Experiments in environment 1

In the first experiment 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 3).

Of the 366 data points, 120 were used for the initial training of the networks[4], i.e. the mapbuilding phase, and the remaining 246 data points were used for the evaluation of the localisation ability.

Localisation performance was assessed using the entropy H(L|R), which determines the correlation between locus and map response (the lower H(L|R), the stronger that correlation, i.e. the better the localisation performance). A detailed discussion of this measure can be found in [5].

The results of figure 4 show that for history lengths between 2 and 7 ($2 \leq \tau \leq 7$) the episodic mapping mechanism performs better than the static one[5]. (The case of $\tau = 1$ (i.e. no episodic information used) is included in the experiments for comparison — it ought to produce similar results to the static mapping paradigm).

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

[5]To interpret figure 4 and 6 note that localisation performance *increases* with *decreasing* H(L|R).

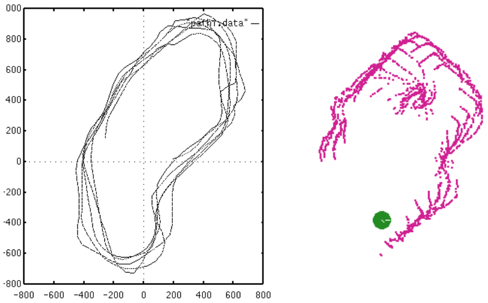

Figure 3: Actual trajectory taken by robot in environment 1 (left), and accumulated infrared sensor readings (right, environment 1 "as the robot sees it"). Dimensions in the diagram are in tenths of inches.

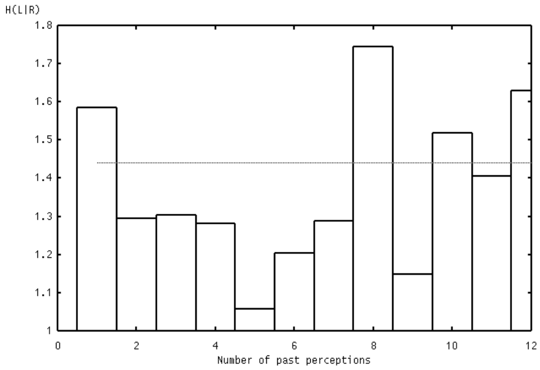

Figure 4: Results obtained in environment 1, using a 12x12 network, partitioned into 16 bins. The physical space of 2.87m x 4.30m was divided into 12 bins. The single layer network achieves H(L|R)=1.44 in this experiment (indicated by horizontal line). The lower H(L|R), the better localisation performance.

Once history length becomes too long, no benefits are obtained by considering temporal aspects, in fact including too many past perceptions reduces the quality of the mapping. Our current explanation for this observation is that beyond an optimal point the inclusion of further episodic information becomes a "confusing" influence due to noise, which means that it is not sufficient to simply increase sensor resolution or temporal resolution to solve problems of perceptual aliasing.

3.3 Results obtained in environment 2

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[6], 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 5 shows the robot's path through this environment, and the robot's perception of it.

Results are shown in figure 6. As before (section 3.2), the episodic mapping mechanism produces a better mapping than the static one, until a critical value of $\tau = 9$ is reached.

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

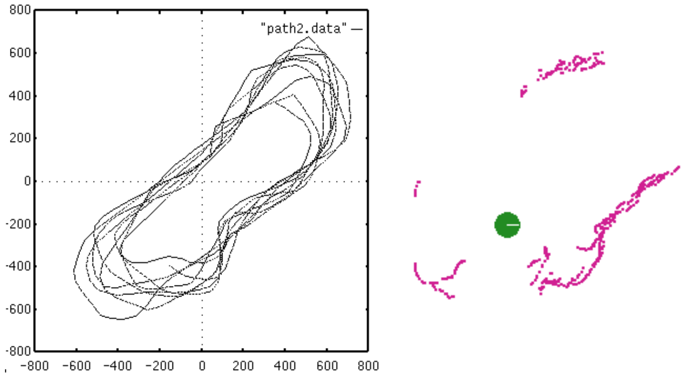

Figure 5: Robot trajectory in environment 2 (left) and accumulated infrared sensor readings obtained by the robot in environment 2 (right, environment 2 "as the robot sees it"). Dimensions are in tenths of inches.

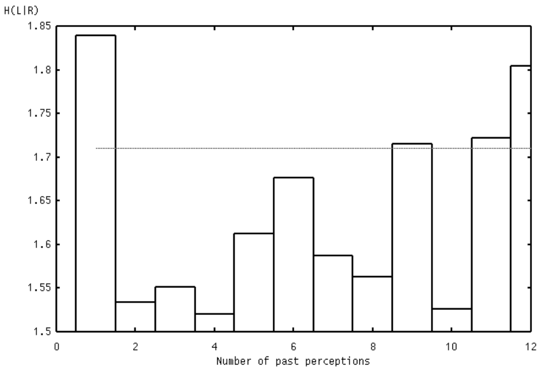

Figure 6: Results obtained in environment 2, using a 12x12 network, partitioned into 16 bins. The physical space of 3.37m x 3.36m was divided into 16 bins. The single layer network achieves H(L|R)=1.71 in this experiment (indicated by horizontal line). The lower H(L|R), the better localisation performance.

These observations are similar to those made through the experiments conducted in environment 1.

4 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.

This paper presents 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.

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. Provided not too much episodic information is included (i.e using a short term memory, rather than a medium or long term one) the episodic mapping performs better than the static one.

The method of double topological mapping discussed here has the advantage of being 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, it is therefore conceivable that the robot determines the optimal episode length automatically. It is also conceivable that the optimal episode length τ can be used to characterise robot-environment interaction. Both these conjectures are currently under investigation.

Also, although we use previous perceptions for the episodic mapping, we do not encode the temporal ordering of those perceptions. Whether using this additional information would produce even better mappings again 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.

[5] Tom Duckett and Ulrich Nehmzow, *Mobile Robot Self-Localization and Measurement of Performance in Middle-Scale Environments*, to appear in J. Robotics and Autonomous Systems, 1998.

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

[8] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

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

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