RatSLAM: A Hippocampal Model for Simultaneous Localization and Mapping

M. J. Milford, G. F. Wyeth, D. Prasser
School of Information Technology & Electrical Engineering
University of Queensland
Brisbane, Australia
{milford,wyeth, prasserd} @itee.uq.edu.au

***Abstract*— The paper presents a new approach to the problem of Simultaneous Localization and Mapping – SLAM – inspired by computational models of the hippocampus of rodents. The rodent hippocampus has been extensively studied with respect to navigation tasks, and displays many of the properties of a desirable SLAM solution. RatSLAM is an implementation of a hippocampal model that can perform SLAM in real time on a real robot. It uses a competitive attractor network to integrate odometric information with landmark sensing to form a consistent representation of the environment. Experimental results show that RatSLAM can operate with ambiguous landmark information and recover from both minor and major path integration errors.**

Keywords-SLAM; hippocampus; mobile robot;

I. INTRODUCTION

In order for a robot to navigate intelligently within a large scale environment, it must possess a means of storing information about past experience, and have the ability to use that information to make decisions about suitable behavior. Over the past decade, there has been considerable interest in solving this problem by building an internal map of the environment, and navigating based on estimates of localization taken from that map. This methodology has come to be known as Simultaneous Localization and Mapping – SLAM. Typical approaches involve the use of grid representations [1], landmark representations [2] or topological representations [3]. Each of these approaches typically have complimentary strengths and weaknesses, and the search for a good solution is the subject of much current research.

This work presents an entirely new SLAM system, RatSLAM, that has been derived from models of the hippocampal complex in rodents. While the detailed models of the function and dynamics of the rodent hippocampal complex remain the subject of debate, there are areas of key agreement. Most notably, it is generally agreed that rodents have *place fields*, patterns of neural activity that correspond to locations in space. The place fields are modulated by the activity of the rodent as it moves about, and also by visual stimulus. This corresponds with the problem of correlating odometric and range or vision sensors in a mobile robot; the problem at the heart of SLAM. Place fields are not grids: they do not form Cartesian representations of the environment. Nor are place fields strictly topological: rodents are able to interpolate between locations to find shorter paths, indicating that the place fields have some properties of space. Place fields are not related solely to visual landmarks either: rodents can still navigate effectively in the dark [4]. The rodent hippocampal complex appears to use the properties of grid based, topological and landmark representations to its advantage.

The RatSLAM system uses an approximate computational model of the hippocampal complex based on competitive attractor networks. This follows other recent computational models of the rodent hippocampus which use competitive attractor networks as the basis for the representation [5-7]. The packet (or packets) of activity in the competitive attractor network represent(s) the belief(s) of the robot with regard to its own pose. Movement of the robot modulates the dynamics of the network, causing the activity packet to change and hence update the pose estimate. Sensory cues become associated with activity packets. Once the associations between sensory cues and pose estimates are learnt, the sensory cues will influence the position of the activity packet to update the pose estimate of the robot.

By using a competitive attractor network structure, RatSLAM builds a representation that is part grid and part topological. Elements that are close in the network are likely to be close in space, but the actual connectivity and sense of the network is defined by the behavior of the robot between elements. Furthermore, the system has one of the main strengths of landmark based systems. RatSLAM can take ambiguous visual input and maintain and propagate multiple pose hypotheses simultaneously. Network dynamics allow these hypotheses to compete with each other until visual input during competition can strengthen the belief in one or more of the possible pose hypotheses.

This paper proceeds with the following structure. The next section (II) describes details of the RatSLAM architecture. Section III gives details of the algorithms and equations that define the activity of the competitive attractor network. Section IV describes the experimental methodology used to investigate the properties of RatSLAM when employed on a Pioneer 2-DXE. Section V describes the results of the experiments, with brief conclusions given in Section VI.

II. THE RATSLAM ARCHITECTURE

A. Overall System

Fig. 1 shows the basic model. The robot's pose is represented by the activity in a competitive attractor network called the *pose cells*. Wheel encoder information is used to

perform path integration by injecting activity into the pose cells thereby shifting the current activity packets. Vision information is converted into a local view representation which if familiar, injects activity into the particular pose cells that are associated with that specific local view.

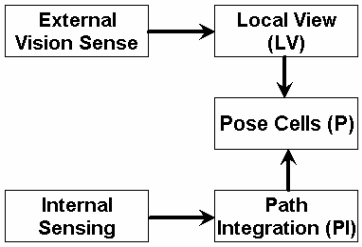

Figure 1. Pose is represented by activity in the pose cells. This pose is updated continually by path integration and local view activity input.

B. Pose Cells

The pose cells are implemented as a competitive attractor network, a type of neural network that is designed to converge to a stable pattern of activation across its units. The network units can be arranged in many configurations, but generally each unit will excite units close to itself and inhibit those further away, which leads to a clump of activity known as an *activity packet* eventually dominating. Activity injected into the network near this winning packet will tend to move that packet towards it. Activity injected far away from it will create another packet that competes with the original. If enough activity is injected the new packet can 'win' and the old packet disappear.

The RatSLAM system uses pose cells to concurrently represent the belief(s) about the location and orientation of the robot. The integration of location and orientation in a single network differs significantly from other models of the rodent hippocampus [8]. Experiments with real rodents have shown that certain cells respond maximally when a rat is at a certain location (place code cells) and that others respond when it is orientated in a certain direction (head direction cells). These results have prompted the use of separate competitive attractor networks for place code (x,y) and head direction (θ). These systems have a fundamental limitation – they cannot represent and maintain multiple beliefs in pose for any period of time. We have previously illustrated this phenomenon in [9].

By representing (x,y,θ) in the same competitive attractor network, the system can concurrently manage several pose beliefs over time. We arrange the pose cells in an (x,y,θ) arrangement for ease of visualization although there is no biological justification for this sort of ordered arrangement. This arrangement also simplifies weight assignment for path integration.

C. Path Integration

Path integration is not meant to be strictly Cartesian; the distance and bearing relationships between units has only been partially tuned to assist in visualisation and weight assignment rather than to assist the function of the network. Each cell occupies approximately 0.25m × 0.25m in area and approximately 9˚ in bearing. The weighting of connections between units within the pose cell network is described in Section III. The coarse nature of the pose cell representation means that path integration based on the pose cell network alone is far inferior to that achieved with simple odometric integration. It is the topological properties, and the relationship of pose to landmarks that maintains consistency and stability in the pose representation.

D. Local View

The robot's camera and vision processing module can see coloured cylinders and report the distance and relative bearing to the cylinder, and associated uncertainties [10]. A three-dimensional matrix of local view cells encodes the cylinder colour (type), distance and bearing. Activated local view cells are constantly being associated with the pose cells that are highly activated at that time through strengthening of weighted connections between them.

Although one of the visual parameters is distance to a cylinder, there is no geometric interpretation of distance to a landmark in our system. Rather in this scheme of artificial landmarks distance to a landmark is used to distinguish seeing a cylinder one metre away as constituting a different scene to seeing that same cylinder three metres away.

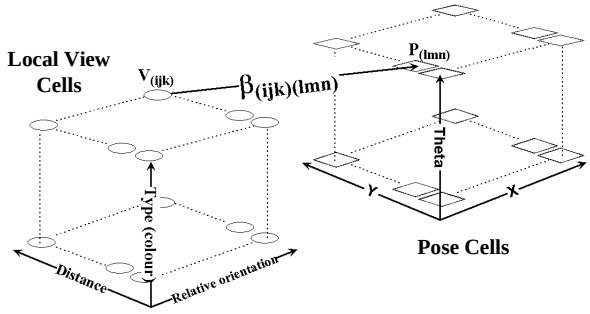

Figure 2. Illustration of the local view network and pose cell network. Units in the local view become associated with units in the pose cells through learnt weighted connections beween the two networks.

III. RatSLAM Dynamics

This section describes the RatSLAM system in operation, detailing the visual association, path integration and competitive attractor processes. The section progresses in the order in which computation is performed.

A. Visual Association Process

The visual association process is the key to maintaining consistent representations of pose in the face of the inconsistent representations that will arise from the coarse path integration process. The connection strengths between the local view cells and the pose cells are strengthened using Hebbian learning, by (1) using the notation from Figure 2.

$$\beta^{t+1}_{(ijk)(lmn)} = \beta^{t}_{(ijk)(lmn)} + \eta P_{lmn} V_{ijk} \quad (1)$$

The learning rate, η, is not critical in operation and was arbitrarily set to 0.05. Since only a small percentage of all the connections will have non-zero weights we encode these

This research was made possible in part by an Australian Research Council (ARC) grant.

weights in a sparse fashion, growing weights dynamically as they are needed. Full connectivity between the current implementation of about 700 local view cells and 180000 pose cells would require 1.3 x 10^8 connections. However, during a one hour real world experiment only about 800000 connections had non-zero weights resulting in a significant saving in computation.

When a familiar scene is encountered the activated local view cells project energy along these weighted connections into the pose cells (2). The amount of energy projected from the local view cells is limited by providing a hard limit on the change in each pose cell unit, σ.

$$\Delta P_{ijk} = \min\left(\sum_{l=0}^{X}\sum_{m=0}^{Y}\sum_{n=0}^{Z}\beta_{(ijk)(lmn)}V_{lmn}, \sigma\right) \quad (2)$$

The value of σ was tuned to provide a balance between maintaining the current pose estimate(s) and re-calibrating from the vision system. Assign too little importance to visual calibration and the robot is not able to recover from kidnapping or even maintain localization. Assign too much importance and the system cannot deal with the possibility of ambiguous visual input. All experiments were conducted with σ = 0.0003.

B. Path Integration

The path integration process projects the pose cell activity into cells slightly offset from the currently activated ones. If the robot is translating the activity is shifted in the *x,y* plane; if the robot is rotating activity is shifted in the θ direction. Under translation the direction of movement of activity is dependent upon the position of the cell in the θ direction. The magnitude of the movement in the *x,y* plane is dependent on the translational velocity, *v*. The movement along the θ axis is dependent on the rotational velocity, ω. Equation (3) shows the energy injected into each pose cell comes from a group of pose cells offset by the integer amounts δx_0, δy_0 and $\delta\theta_0$. The amount of activity injected is based on the product of the activity of the sending unit, *P*, and a residue component, α. The residue component is spread over a 2×2×2 cube to account the quantization effects of the grid representation. The residue is based on the fractional components of the offsets, δx_f, δy_f and $\delta\theta_f$.

$$\begin{aligned}
&\Delta P_{lmn} = \sum_{x=\delta x_o}^{\delta x_o+1}\sum_{y=\delta y_o}^{\delta y_o+1}\sum_{z=\delta\theta_o}^{\delta\theta_o+1}\alpha_{xyz}P_{(l+x)(m+y)(n+z)} \\
&\delta x_o = \lfloor k_x v\cos\theta \rfloor, \delta y_o = \lfloor k_y v\sin\theta \rfloor, \delta\theta_o = \lfloor k_\theta \omega \rfloor \\
&\delta x_f = k_x\cos\theta - \delta x_o, \delta y_f = k_y\sin\theta - \delta y_o, \delta\theta_f = k_\theta\omega - \delta\theta_o \quad (3) \\
&\alpha_{ijk} = g(\delta x_f, i-\delta x_o)g(\delta y_f, j-\delta y_o)g(\delta\theta_f, k-\delta\theta_o) \\
&g(a,b) = \begin{cases} 1-a, & b=0 \\ a, & b=1 \end{cases}
\end{aligned}$$

C. Competitive Attractor Dynamics

After the visual and path integration processes, the pose cells undergo the internal competitive attractor dynamic process. The competitive attractor dynamics ensure that the total activity in the pose cells remains constant. This is consistent with the interpretation of the pose cells as a probability distribution of pose. The activity packets located near each other move towards each other, bringing together similar pose representations. Separated activity packets representing multiple hypotheses of pose compete with each other. Global inhibition means that without visual or path integration input the activity will eventually stabilize to one stable packet.

There are four stages to the internal dynamics:

1. Excitatory update within each *x,y* layer,
2. Excitatory update between *x,y* layers,
3. Global inhibition of all cells, and
4. Normalization of pose cell activity.

1) Internal X-Y Layer Update

A two dimensional discrete Gaussian distribution was used to create the excitatory weights, ε. The weighted connections project the activity from each cell *P* to all other cells in the N_x by N_y layer.

$$\Delta P_{jk} = \sum_{a=0}^{N_x}\sum_{b=0}^{N_y}\varepsilon_{(a-j)(b-k)}P_{ab} \quad (4)$$

2) Inter-Layer Update

A one dimensional Gaussian distribution is used to form the weights, δ, which cause excitation between layers. The field of influence of a layer is about 45° (or two layers each side) – set by γ.

$$\Delta P_{ijk} = \sum_{c=k-\gamma}^{k+\gamma}\delta_{(c-k)}P_{ijc} \quad (5)$$

Connections between layers represent links between cells with similar angular orientations. As such there is wraparound of connections in the theta direction – the 'top' layer in Fig. 2. excites both the layers directly below it and the layers at the 'bottom' of the diagram.

3) Global Inhibition

Because multiple pose hypotheses (represented by multiple activity packets) require time to compete and be reinforced by further visual input, inhibition is relatively gentle and rival packets can co-exist for significant periods of time. The level of inhibition decreases as cell activation increases. The inhibition constant φ controls the level of global inhibition and is set to 0.004. Activation levels are limited to non-negative values.

$$P_{ijk}^{t+1} = \max\left[P_{ijk}^{t} + \varphi\left(P_{ijk}^{t} - \max(\mathrm{P})\right), 0\right] \quad (6)$$

4) Normalization

The last step is normalization which maintains the total activation level after visual and path integration input at one.

$$P_{ijk}^{t+1} = \frac{P_{ijk}^{t}}{\sum_{x=0}^{N_X}\sum_{y=0}^{N_Y}\sum_{z=0}^{N_{THETA}}P_{xyz}^{t}} \quad (7)$$

An example of the competitive attractor dynamics in operation is illustrated in Fig. 3. This figure shows two rival

packets, where one packet wraps around through the top of the theta axis.

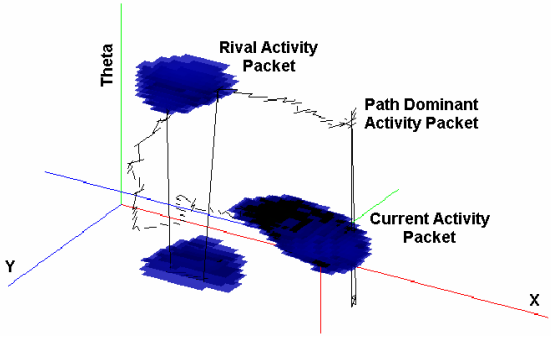

Figure 3. Snapshot of pose cell activity during an experiment. Note the current activity packet is smeared indicating that it is moving. The rival activity packet here will not win unless it receives reinforcement from further visual input.

IV. Experimental Setup

RatSLAM has been tested on a Pioneer2-DXE robot. The robot carries a 400 MHz AMD K6-2 processor that performs on board processing of the vision and interfaces with the motion control system. A 1.1 GHz Pentium III laptop runs the pose cell network, interfacing over a wireless link. Using this hardware all processes were updated every 200 ms.

The main test arena was a carpeted corridor area in a campus building, with dimensions of approximately 20 m × 10 m. The only modifications made to the environment were the addition of cardboard boxes at exits to prevent the robot leaving the test arena and the placement of coloured cylinders (Fig. 4). The coloured cylinders were used as artificial landmarks, with the vision system tuned to recognize rectangular areas of consistent colour as landmarks [10]. The cylinders were consistently visible when between one and three metres away, with a distance uncertainty of about 10%. Cylinders positioned with bright light sources behind them (such as windows) were almost invisible to the robot from some orientations but seen easily from others. Rectangular patches of sunlight on the floor were sometimes picked up as being cylinders and learnt as landmarks. The vision sensor was a forward-facing CCD camera with an effective field of view of about 40°.

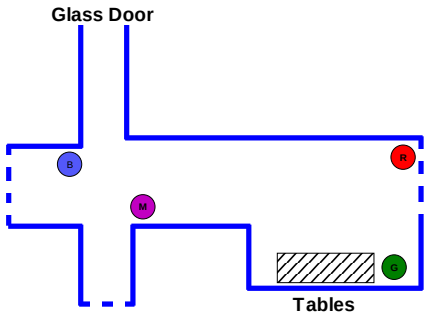

Figure 4. Plan view of testing arena. Dotted lines indicate where cardboard boxes have been used to block corridors. Cylinder locations are shown with exaggerated size. The colors of the cylinders were red (R), green (G), magenta (M) and blue (B).

The robot performed wall following to circumnavigate the test environment. Additional parallel behaviors included obstacle avoidance and homing on the cylinders. The tests were performed during the day and as such there was a reasonable volume of human traffic through the area. The robot sometimes had to go around people who tried to obstruct it, and coped with obscured landmarks or false landmarks generated by people in brightly coloured pants.

The robot was not given any information about the layout of the environment before the start of the test. Learning, recall and map maintenance all occurred concurrently. There was no user intervention during the tests, and no “phases” of learning in the trial. All tests were of 40 minute duration.

Figure 5. Testing arena and Pioneer robot. The Pioneer is equipped with a ccd camera, wheel encoders and eight sonar sensors spread over the front half of the robot. Coloured cylinders are used as artificial landmarks.

V. Results

A. *Performance Indicators*

RatSLAM does not produce strictly Cartesian representations of the environment; rather it creates a topologically consistent representation that has some Cartesian properties. Consequently, performance indicators that illustrate consistency rather than Cartesian accuracy have been developed to evaluate the effectiveness of RatSLAM. The first of these indicators is consistency in measured trajectory over time. Uncorrected odometry from the robot gradually drifts in an unbounded fashion over time – a successful test produces trajectories that remain bounded. The second indicator is consistency in the perceived location of the cylinders based on the robot’s localization in the pose cells. If the robot is consistently localized then the landmark position estimates should be tightly clustered.

B. *Baseline Performance of Odometry*

The initial tests recorded the baseline performance of the odometry system using the Pioneer’s internal pose estimation. Note that the internal estimate is *not* related to the RatSLAM path integration process – RatSLAM is inherently noisy in its pose estimates because of quantisation effects. The test illustrated in Fig. 6 and Fig. 7 shows that there is significant odometric drift over the 40 minutes of operation.

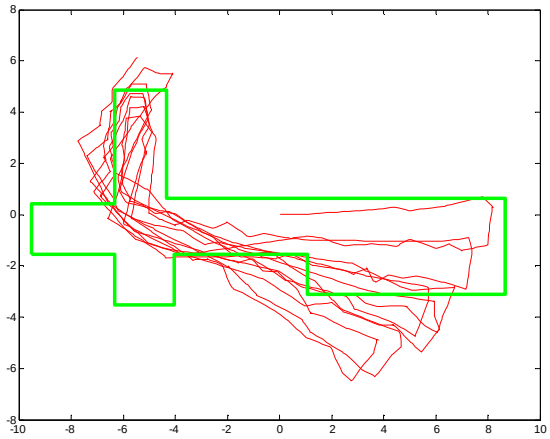

Figure 6. Uncorrected odometry trajectory.

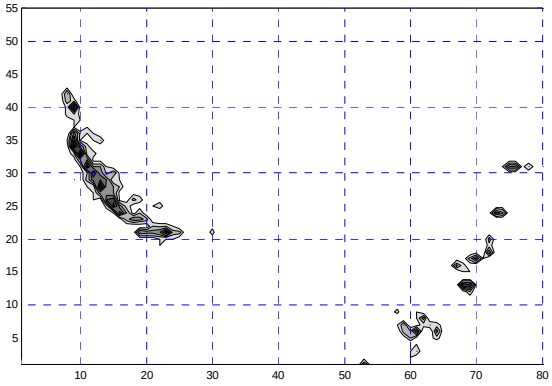

Figure 7. Mapped cylinder positions based on uncorrected odometry only.

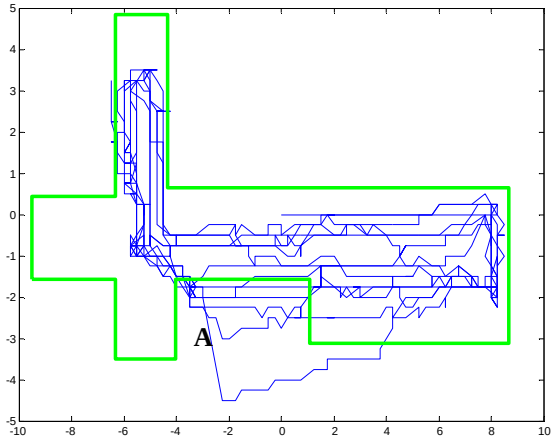

Figure 8. RatSLAM trajectory – unique landmarks.

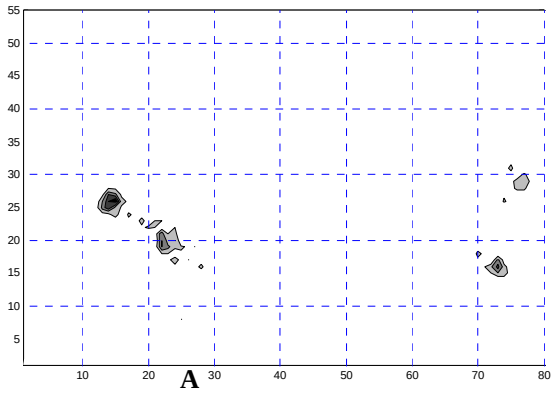

Figure 9. Mapped cylinder positions using RatSLAM – unique landmarks.

C. Unique Landmarks

The first set of tests run using the RatSLAM system used unique landmarks. The landmarks were unique in that each of the four cylinders in the test arena were of a different colour. There was still a level of ambiguity in the landmarks during this test however. The cylinders look identical from any orientation, meaning that a unique local view representation corresponds to a helix of possible pose representations. This singularity is resolved by the ability of the pose cells to account the context of the local view with respect to the robot's pose.

The results from an indicative test run are shown in Fig. 8 and Fig. 9. In contrast to the unbounded drift in localisation seen in Fig. 6, the RatSLAM system bounds the trajectories, as shown in Fig. 8. Errors in path integration that were not corrected by visual cues caused the robot to become lost (perhaps because the cylinders were temporarily obscured by passers-by). However once visual input was received the robot re-localized, as shown at point A. The mapped cylinder positions in Fig. 9 show tight clustering, which indicates that the robot maintained consistent localisation throughout the test.

D. Ambiguous Landmarks

In the second set of test runs, the landmark information was further ambiguated by the use of two identical landmarks. The magenta cylinder in Fig. 4 was replaced with a red cylinder. In this test, the pose cells were required to maintain multiple pose hypotheses over a period of time until one hypothesis could be strengthened by evidence from a unique landmark.

The results from an indicative test run are shown in Fig. 10 and Fig. 11. The trajectory of the robot remains bounded in a fashion similar to the tests with unique landmarks. The plot in Fig. 10 shows two clear cases of significant re-localization from path integration errors. The system is, of course, continuously re-localising in a less dramatic fashion. The errors at points B and C were due to significant wheel slippage caused when the robot clipped an obstacle. In both cases the robot was able to effectively re-localize to a consistent location from visual cues. Fig. 11 also shows tight clustering of the cylinder locations, indicating consistent localization. The additional cluster that appears in the top left corner came from a patch of sunlight that had similar appearance to a blue cylinder. The system was able to use this further ambiguous feature to its advantage, rather than the feature causing any degeneration in system function.

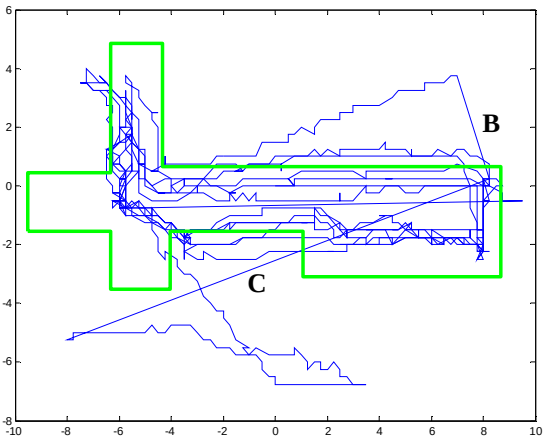

Figure 10. RatSLAM trajectory with non-unique landmarks. Note the two obvious relocalization jumps at B and C after the robot has travelled for a while without visual input and built up a significant path integration error.

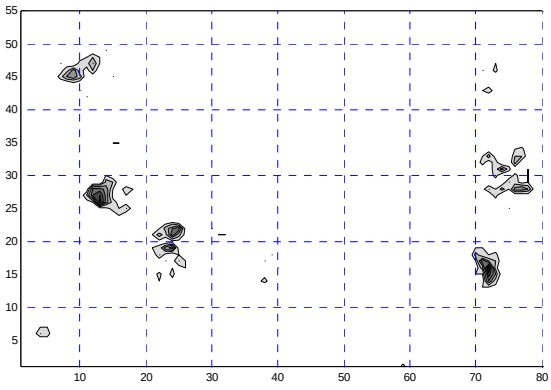

Figure 11. Mapped cylinder positions using RatSLAM – non-unique landmarks. Note the correspondence with the cylinder positions in Fig. 4 and the learnt cylinder at top left corresponding to a patch of sunlight.

E. Discussion

The experiments have clearly shown that RatSLAM is able to build consistent representations of its environment. Any consistent and useful representation of the environment can be called a "map", so the system is clearly mapping even though the map does not follow a strict Cartesian coherence. Furthermore, the robot is able to localize itself with respect to the map in an incremental on-line fashion. Consequently, it is clear that RatSLAM is performing Simultaneous Localization and Mapping.

While this work shows that the RatSLAM hippocampal model can perform SLAM, it opens questions in terms of operation of the system in wider settings. There are issues in computability of the large competitive attractor networks needed to operate in larger environments. "Closing the loop" on large circular arenas will be a good test of the topological properties of the pose cell representation. Methods of using the map for goal oriented behavior require investigation. Methods currently under investigation follow the ideas of a "cognitive map" [4]. In other work, we propose to replace the cylinder landmark local view system with a system inspired by hierarchical models of the visual cortex in mammals. It is intended then to perform SLAM in an unmodified indoor environment.

VI. CONCLUSION

RatSLAM captures the benefits of mapping and localization using a model of the rodent hippocampus. The pose cell representation at the heart of RatSLAM captures the benefits of grid-based, topological and landmark based representations, giving the robot a sense of space that is not rigidly tied to a Cartesian grid. The system has been shown, by experimentation on a real robot, to create consistent representations of its environment in an on-line incremental fashion without user intervention. Furthermore, RatSLAM can resolve ambiguous landmark data, even when subject to large odometric errors.

[1] S. Thrun, "Learning Maps for Indoor Mobile Robot Navigation," *Artificial Intelligence*, 1997.

[2] J. Nieto, J. Guivant, and E. Nebot, "Real Time Data Association for FastSLAM," presented at International Conference on Robotics & Automation, Taipei, Taiwan, 2003.

[3] Y. Endo and R. C. Arkin, "Anticipatory Robot Navigation by Simultaneously Localizing and Building a Cognitive Map," presented at Intelligent Robots and Systems, Las Vegas, 2003.

[4] D. Redish, *Beyond the Cognitive Map*. Massachusetts: Massachusetts Institute of Technology, 1999.

[5] A. Arleo, "Spatial Learning and Navigation in Neuro-Mimetic Systems: Modeling the Rat Hippocampus," in *Information Science*. Milan: University of Milan, 2000, pp. 198.

[6] D. Redish, A. Elga, and D. Touretzky, "A coupled attractor model of the rodent head direction system," *Network Computer Neural Systems*, vol. 7, pp. 671-685, 1996.

[7] S. M. Stringer, E. T. Rolls, T. P. Trappenberg, and I. E. T. de Araujo, "Self-organizing continous attractor networks and path integration: two-dimensional models of place cells," *Computational Neural Systems*, vol. 13, pp. 429-446, 2002.

[8] A. Arleo and W. Gerstner, "Modeling Rodent Head-direction Cells and Place Cells for Spatial Learning in Biomimetic Robotics," presented at Sixth International Conference on the Simulation of Adaptive Behaviour (Animals to Animats6), 2000.

[9] M. J. Milford and G. Wyeth, "Hippocampal Models for Simultaneous Localisation and Mapping on an Autonomous Robot," presented at Australian Conference on Robotics and Automation, Brisbane, Australia, 2003 [in press].

[10] D. Prasser and G. Wyeth, "Probabilistic Visual Recognition of Artificial Landmarks for Simultaneous Localization and Mapping," in *Proceedings of the 2003 IEEE International Conference on Robotics and Automation*. Taipei, 2003.