

Spatial Cognition for Robots

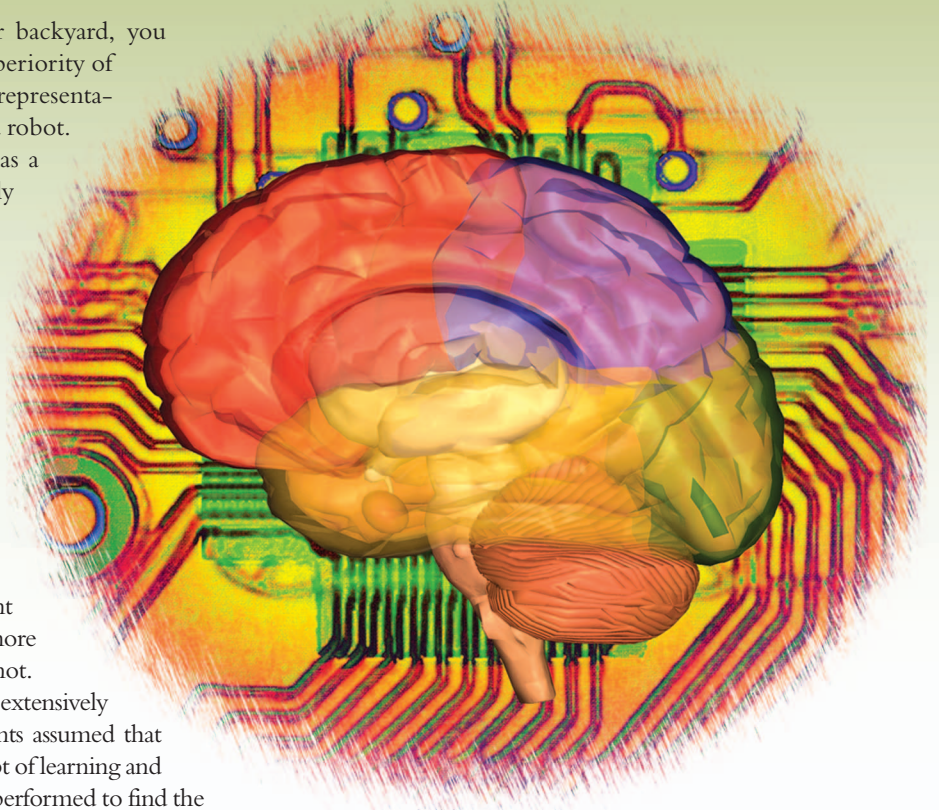
Robot Navigation from Biological Inspiration

BY GORDON WYETH
AND MICHAEL MILFORD

If you see a rat scuttling through your backyard, you might want to stop and consider the superiority of the rat at creating and exploiting spatial representations compared with the most advanced robot.

Chances are that the rodent you see has a nest that is many hundreds of meters, possibly kilometers, from your backyard, and yet the rodent can unerringly return to its home. If your yard has some ripe seeds or fruits (and no predatory pets), the rat may return some days later, further demonstrating its ability to store and recall the spatial layout of its range. The rat runs under leaves and through drains, with few clear landmarks in a world that is under constant perceptual change in terms of appearance, texture, and odor. Clearly, the rat can build a map over large ranges in a real-world environment under constant change and use and maintain that map more than its 2–3 year lifetime. As yet, a robot cannot.

The spatial abilities of rats have been extensively studied for several decades. Early experiments assumed that the rat's ability was based on a generic concept of learning and memory, and many behavioral studies were performed to find the limits of the rodent's abilities [15], [16], [23]. The breakthrough came in 1971 with the discovery of the rat's cognitive map in the cells of the hippocampus, a region of the midbrain common to all mammals, including rodents and humans [16]. Certain cells in the hippocampus were shown to be active only when the rodent was at a quite specific place in the world and were consequently dubbed *place cells* [15]. Subsequent experiments with place cells have demonstrated that rats do indeed localize and that rodents combine odometric information (from vestibular sensors and motor commands) with perceptual cues (from vision, whisking,



© ARTVILLE, LLC.

and olfaction) to simultaneously build a map and localize from the map. More recently, further cells tied to spatial ability have been discovered, such as the head-direction cells [18] and grid cells [7], which are slowly forming a more complete picture of the mechanisms used for spatial cognition in the rodent brain.

As roboticists, the notion of a cognitive mechanism that binds odometric and perceptual information into maps suitable for navigation is certainly appealing. There is a large and ever-developing body of research addressing the issue of simultaneous localization and mapping (SLAM), with some impressive results [5], [14],

[22]. However, the rat's spatial cognitive mechanism has some advantages over current SLAM solutions. First, the rat seamlessly integrates mapping, localization, and task execution throughout its lifetime by constantly building, maintaining, and using its spatial representations. SLAM solutions, in contrast, typically have discrete map building and usage phases and cannot easily accommodate changes in the environment. Second, the rat's perceptual cues have, at best, weak geometric properties, and odometry interpreted from foot motion and vestibular information would appear to be highly prone to noise. Despite the lack of accurate geometric information, the rat can maintain a spatial representation suitable for path planning over extensive ranges [4].

In our work, we have sought to build a model of the rodent brain that is suitable for practical robot navigation. The core model, dubbed *RatSLAM*, has been demonstrated to have exactly the same advantages described earlier: it can build, maintain, and use maps simultaneously over extended periods of time and can construct maps of large and complex areas from very weak geometric information [11], [13]. Our work contrasts with other efforts to embody models of rat brains in robots. While others have strived for biological fidelity [1], [9], we have been interested in robotic practicality. Other research that has been focused on practical outcomes [2], [6] has not been conducted at the same temporal and spatial scales as our studies with RatSLAM.

In this article, we describe the key elements of the known biology of the rat brain in relation to navigation and how the RatSLAM model captures the ideas from biology in a fashion suitable for implementation on a robotic platform. We then outline RatSLAM's performance in two difficult robot navigation challenges, demonstrating how a cognitive robotics approach to navigation can produce results that rival other state-of-the-art approaches in robotics.

How a Rat Brain Works

Evidence for the neural basis of spatial encoding has principally been gleaned from recordings of the activity in a single cell (or small group of cells) when the rodent moves through space. A complete pattern of brain activity cannot be compiled in this way, but the correlation between rodent pose and single-cell firing over time provides strong insight into the nature of the biological system. A review of the current understanding of the neural basis of spatial encoding can be found in [10].

Cell Types

The place cells are predominantly found in the CA1–CA3 regions of the hippocampus [3]. Single-cell recordings show that the firing rate of a single place cell is strongly correlated with the rodent's absolute place at the time of firing. Place cells tend to be direction invariant: it does not matter which way the rodent is facing but only where it is located. Figure 1(a) shows the firing rate of an idealized place

cell with respect to place and absolute heading. The firing rate with respect to place (shown as intensity) forms a peak at a single location corresponding to the place within the testing environment that is associated with that place cell. The firing rate with respect to head direction (shown as the distance from the origin to the curve) is uniform for all directions, showing no association between head direction and cell firing. Recordings taken from other place cells within the same animal will be selective to other places in the testing environment.

Head-direction cells show a complementary property to place cells [18], [21]: they are place invariant and direction specific. Figure 1(b) shows an idealized head-direction cell that is equally likely to fire at any position within the testing environment but has a distinct peak in its firing activity at a particular orientation of the animal's head with respect to the global reference frame. As the name suggests, the head-direction cell's firing is dependent on the direction of the animal's head and not on the animal's body. These cells are found in various brain regions around hippocampus, particularly in postsubiculum, retrosplenial cortex, and some regions of thalamus.

Recently, a new type of spatial encoding cell called a *grid cell* has been discovered in the entorhinal cortex, an area closely related to the hippocampus proper [7]. In the shallowest layers of entorhinal cortex, grid cells show place-cell-like properties but significantly show multiple firing fields tessellated with a hexagonal pattern across the environment. The shallowest layers show a tessellation in place and no directional specificity as shown in Figure 1(c). In deeper layers, some cells have only head direction characteristics, whereas the deepest layers contain conjunctive grid cells that show the conjunction of grid and head direction characteristics as shown in Figure 1(d).

Connectivity of Regions

It is difficult to state categorically how the brain regions containing these cell types are connected and how they connect to other parts of the brain. In a recent review of spatial encoding behavior

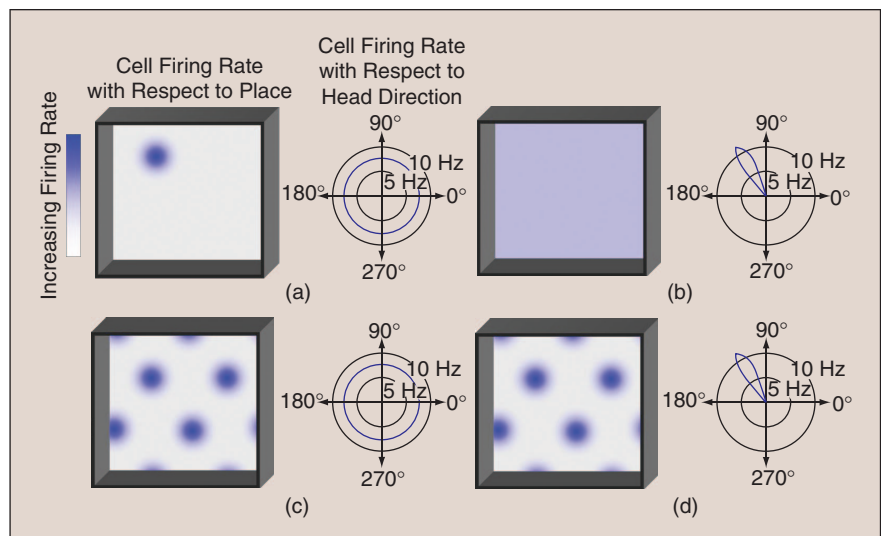


Figure 1. Idealized place and head direction specificity for various cell types involved in spatial encoding: (a) place cell, (b) head-direction cell, (c) grid cell, and (d) conjunctive grid cell.

[25], a map showing the principal connections between regions involved with the spatial encoding is shaded to show the principal cell type to be found in each region. Based on this more detailed diagram, one could broadly model the connectivity of the cell types involved in spatial encoding as shown in Figure 2.

Spatial Encoding Behavior

Recordings from spatially selective cells have shown that a rodent will observe the layout of key features in the environment to reset pose estimates. Suppose cell recordings are taken from a rodent in a plain square arena that has a single black feature on its northern wall. If the rodent is removed from the arena, the black feature moved to the eastern wall, and then the rodent is returned, new cell recordings will show directional and place specificity has been rotated by 90°. The rodent brain is using external perception to correct its pose estimation, performing a similar function to the update process in robot SLAM.

However, spatially selective cells do not rely on observations of landmarks to fire. Recordings taken in complete darkness (and with auditory and odor cues also removed) show that the spatially selective cells continue to fire even though there is no sensory stimulus. Furthermore, rodents update the pose estimate represented by the firing of spatially selective cells based on estimates of self-motion obtained from copies of motor commands and vestibular information. As with robots, pose estimates degrade with time in the absence of external cues [8], suggesting similar computation to the prediction process in robot SLAM. There exists no data on rodent performance, with respect to loop closure and larger environments, as it is difficult to make these types of recordings because of constraints in terms of recording technique and the behavioral range of the caged laboratory rats.

Unlike the robot, the rat does not appear to build any geometrical representation of its surrounds; there is no map per se. Instead, the rat relies on learnt associations between external perception and the pose belief created from the integration of self-motion cues. Nor does the rodent appear to have any indication of a probability distribution in the activity in the cells. A roboticist versed in probabilistic SLAM might expect the activity in the head-direction cells to represent a broader range of absolute heading in the absence of perceptual cues to correct bearing drift. Neural recordings, on the other hand, show a consistent range of active head-direction cells over all conditions. It is not clear whether the cell firings represent multiple or single

estimates of pose—it is impossible to tell without recordings across all cells simultaneously—so it is hard to tell whether a rat tracks multiple hypotheses in the same way that a multihypothesis Kalman filter or a particle filter might.

How RatSLAM Works

RatSLAM emulates the rat's spatial encoding behavior using three key components: the pose cells that are analogous to the rodent's conjunctive grid cells, the local view cells that provide the interface to the robot's sensors in place of the rodent's perceptual system, and the experience map that functionally replaces the place cells found in CA1–CA3. The components of RatSLAM and the interactions of the components are illustrated in Figure 3 and described briefly in the following sections. Further details of the operation of RatSLAM can be found in [11].

Pose Cells

The pose cells are a three-dimensional (3-D) continuous attractor network (CAN) [19], [20]. The CAN, often used to model spatially selective cell networks, is a neural network that consists of an array of units with fixed weighted connections. The CAN predominantly operates by varying the activity of the neural units between zero and one rather than by changing the value of the weighted connections. In rodents, spatially responsive cells, such as place cells, fire fastest when the rat is at a certain location. In RatSLAM, the activation value of a neural unit increases when the robot approaches a location associated with that neural unit.

During operation, the pose-cell network will generally have a single cluster of highly active units: the activity packet. The center of the activity packet provides an estimate of the robot's pose that is consistent with the pose-cell network's rectangular prism structure as shown in Figure 3. Each of the three dimensions of the prism corresponds to one of the three spatial dimensions x' , y' and θ' . Primed coordinates are used to differentiate the space from that used with the experience map. The robot's pose in (x', y', θ') space maps to the active neural units in the rectangular prism structure.

The activity packet is self-maintained by local excitatory connections that increase the activity of units that are close in (x', y', θ') space to an active unit. Inhibitory connections suppress the activity of smaller clusters of activity elsewhere in the network. Connections wrap across all six faces of the pose-cell network as shown by the longer red arrows in Figure 3. The change in the cells' activity level ΔP is given by

$$\Delta P = P * \varepsilon - \varphi, \quad (1)$$

where P is the activity matrix of the network, ε is the connection matrix, and $*$ is the convolution operator. The constant φ creates further global inhibition in addition to the inhibition inherent in the connection matrix. At each time step, activation levels in P are restricted to nonnegative values, and the total activation is normalized to one.

Path Integration

Path integration involves shifting the activity packet in the pose-cell network based on odometry information. At each time step, RatSLAM interprets the odometry information to displace a

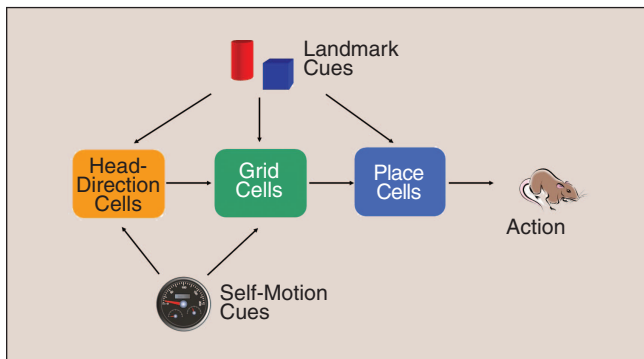


Figure 2. Connectivity of the brain regions containing various spatial encoding cell types.

copy of the current activity state in the pose-cell network. Like the excitatory and inhibitory weight matrices, the path integration process can cause a cluster of activity in the pose cells to shift off one face of the pose-cell structure and wrap around to the other as is shown in both packets in Figure 3, one of which is wrapping across the θ' boundary, the other across the y' boundary. Recording from a single cell under path integration will create firing fields with rectangular tessellations, similar to the triangular tessellations seen in grid cells (see Figure 1).

Local View Cells

The local view cells represent a sparse classification vector based on what the robot is seeing. Each local view cell becomes active when the robot sees a distinct visual scene. Multiple local view cells can be simultaneously active to varying degrees in regions with perceptual aliasing, and there is no competition imposed between the cells.

RatSLAM increases the strength of connections between local view cells and pose cells that are active simultaneously. In other words, RatSLAM learns an association between a visual scene and the robot pose. During a loop-closure event, the familiar visual scene activates local view cells with learnt connections to the pose cells representing the pose where the visual scene was first encountered. Because of the attractor dynamics of the pose cells, a single visual scene is not enough to force an immediate change of pose; several consecutive and consistent views are required to update the pose. The attractor dynamics temporally and spatially filter the information from the local view cells, providing rejection of spurious loop-closure events.

The connections between the local view cells and pose cells are stored in a connection matrix β , where the connection between the local view cell V_i and the pose cell $P_{x',y',\theta'}$ is given by

$$\beta_{i,x',y',\theta'}^{t+1} = \max(\beta_{i,x',y',\theta'}^t, \lambda V_i P_{x',y',\theta'}^t), \quad (2)$$

where λ is the learning rate. When a familiar visual scene activates a local view cell, the change in pose cell activity, ΔP , is given by

$$\Delta P_{x',y',\theta'} = \frac{\delta}{n_{act}} \sum_i \beta_{i,x',y',\theta'} V_i, \quad (3)$$

where the δ constant determines the influence of visual cues on the robot's pose estimate, normalized by the number of active local view cells n_{act} . Figure 3 represents the moment in time when a strongly active local view cell has injected sufficient activity into the pose cells to cause a shift in the location of the dominant activity packet. The previously dominant activity packet can also be seen, which is less strongly supported by a moderately activated local view cell.

Experience Mapping for Path Planning

Over time, the (x', y', θ') arrangement of the pose cells corresponds decreasingly to the spatial arrangement of the physical environment. Often, when a loop-closure event occurs, the odometric error introduces a discontinuity into the pose cell's representation of space, creating two sets of cells that might represent the same area in space. Similarly, the wrapping connectivity leads to pose ambiguity, where a pose cell encodes multiple

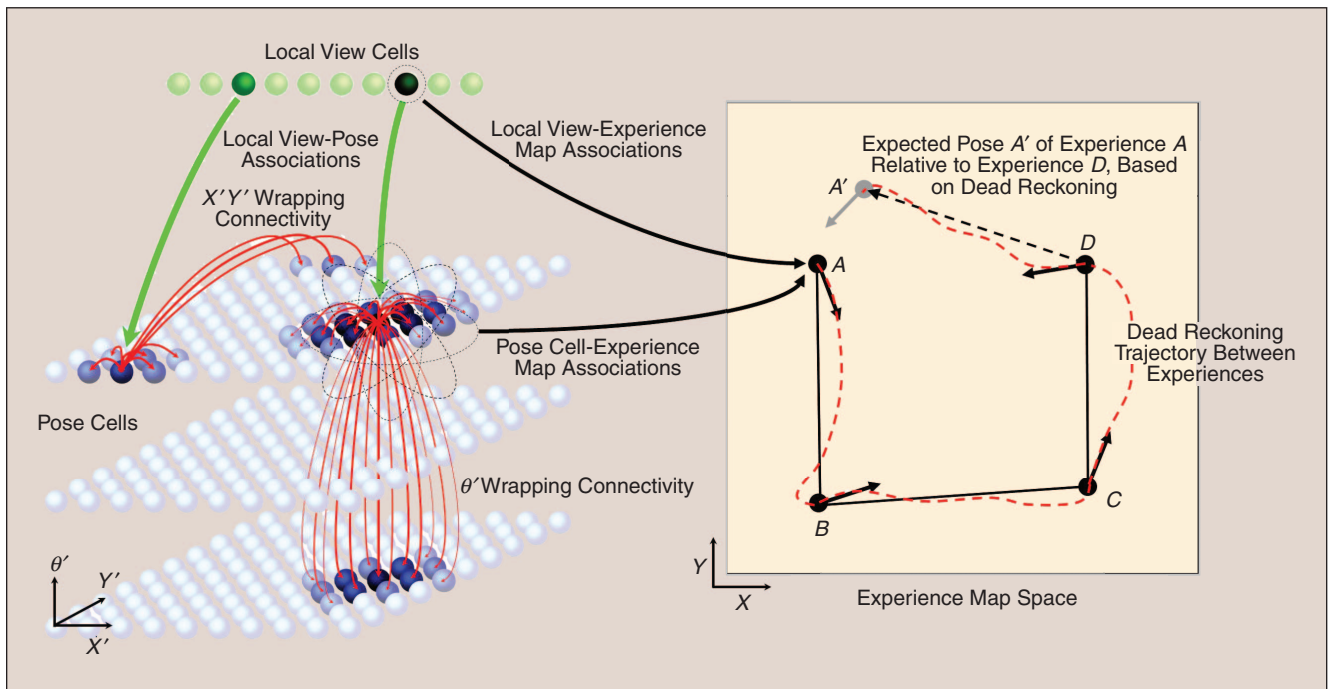


Figure 3. The RatSLAM system. Each local view cell is associated with a distinct visual scene in the environment and becomes active when the robot sees that scene. A 3-D CAN forms the pose cells, where active pose cells encode the estimate of the robot's pose. Each pose cell is connected to proximal cells by excitatory (red arrows) and inhibitory connections with wrapping across all six faces of network. Intermediate layers in the (x', y') plane are not shown. The network connectivity leads to clusters of active cells known as activity packets. Active local view and pose cells drive the creation of experience nodes in the experience map, a semimetric graphical representation of places in the environment and their interconnectivity.

locations in the environment, forming the tessellated firing fields seen in grid cells in the rat. Planning with this representation is clearly not possible. Rather than try to correct these discontinuities and ambiguities within the pose cells, we combine the activity pattern of the pose cells with the activity of the local view cells to create a topologically consistent semimetric map in a separate coordinate space called the *experience map*.

The experience map contains representations of places combined with views, called *experiences* (e^i), based on the conjunction of a certain activity state, P^i , in the pose cells and an active local view cell, V^i . Links between experiences, l^{ij} , describe the spatiotemporal relationships between places. Each experience is positioned at a location \mathbf{p}^i , a position that is constantly updated based on the spatial connectivity constraints imposed by the links. Consequently, the complete state of an experience can be defined as the three-tuple:

$$e_i = \{P^i, V^i, \mathbf{p}^i\}. \quad (4)$$

Figure 3 shows the region of pose cells and the single local view cell associated with the currently active experience A .

Transition links, l^{ij} , encode the change in position, $\Delta\mathbf{p}^{ij}$, computed directly from odometry and the elapsed time, Δt^{ij} , since the last experience was active:

$$l^{ij} = \{\Delta\mathbf{p}^{ij}, \Delta t^{ij}\}, \quad (5)$$

where l^{ij} is the link from the previously active experience e_i to the new experience e_j . The temporal information stored in the link provides the travel time between places in the environment, which is used for path planning. Path planning is achieved by integrating the time values in the transition links starting at the robot's current location to form a temporal map. The fastest path to a goal experience can be computed by performing steepest gradient ascent from the goal experience to the current location.

Experience Map Maintenance

A score metric S is used to compare how closely the current pose and local view states match those associated with each experience:

$$S^i = \mu_p |P^i - P| + \mu_v |V^i - V|, \quad (6)$$

where μ_p and μ_v weight the respective contributions of pose and local view codes to the matching score. If any experience matching scores are below the threshold, the lowest scoring is chosen as the active experience and represents the best estimate of the robot's location within the experience map.

If the activity state in the pose cells or local view cells is not sufficiently described by any of the existing experiences ($\min(\mathbf{S}) \geq S_{\max}$), a new experience is created using the current pose and local view cell activity states. The odometry information defines the initial location in experience space of a newly created experience:

$$e_j = \{P^j, V^j, \mathbf{p}^i + \Delta\mathbf{p}^{ij}\}. \quad (7)$$

When loop closure occurs, the relative position of the two linked experiences in the map will typically not match the odometric transition information between the two as shown

in the discrepancy between experience A and A' in Figure 3. The experience map relaxation method seeks to minimize the discrepancy between odometric transition information and absolute location in experience space by applying a change in experience location $\Delta\mathbf{p}^i$:

$$\Delta\mathbf{p}^i = \alpha \left[\sum_{j=1}^{N_f} (\mathbf{p}^j - \mathbf{p}^i - \Delta\mathbf{p}^{ij}) + \sum_{k=1}^{N_t} (\mathbf{p}^k - \mathbf{p}^i - \Delta\mathbf{p}^{ki}) \right], \quad (8)$$

where α is a correction rate constant, N_f is the number of links from experience e_i to other experiences, and N_t is the number of links from other experiences to experience e_i . Equation (8) is applied iteratively at all times during robot operation; there is no explicit loop-closure detection that triggers map correction. The effect of the repeated application of (8) is to move the arrangement of experiences in experience map space incrementally closer to an arrangement that averages out the odometric measurement error around the network of loops.

Indoor Results

We used RatSLAM as the basis of an autonomous mobile robot system, which could perform mock delivery tasks over a two-week period in a real world, nonstatic environment with a minimum of human intervention [24]. The environment was a busy, functioning floor of an office building with cubicles, laboratories, corridors, offices, and a kitchen. The environment was by no means static, with moving people, changing door states, rearrangement of furniture and equipment, and a range of trolleys that intermittently appeared and disappeared during deliveries, maintenance, and cleaning operations. Perceptually, the environments also changed, with the most significant example being the day–night time cycles, which had an especially significant impact in areas with many windows. There were no modifications to the environment, such as placing of navigation beacons, and no instructions given to the occupants of the building.

The system was deployed on a Pioneer 3 DX robot (shown in Figure 4) equipped with a panoramic imaging system, a ring of forward facing sensors, a Hokuyo laser range finder, and encoders on the wheels. All computation and logging was run onboard on a 2-GHz single-core computer running Windows XP. The robot's typical operation speed was 0.3–0.5 ms⁻¹. The robot operated continuously for 2–3 h between recharging cycles. Because of the need for supervision of the robot during operation (to ensure the integrity of the experiment), the total active time in a typical day was limited to one to four recharge cycles. To capture the effect of around-the-clock operation, experiments were conducted across all hours of the day.

Interface to RatSLAM

RatSLAM requires an odometry measure to perform path integration in the pose cells and an external perception system to selectively activate the local view cells. Odometry was calculated from the encoders on the robot's wheels. External perception for the local view system was driven entirely from the vision system with the robot's laser range finder and sonar systems being used only to support obstacle avoidance and corridor centering behaviors for autonomous operation.

The vision system used panoramic images obtained from an IEEE-1394 camera, mounted at the central rotation axis of a Pioneer 3DX robot, and facing vertically upward at a parabolic mirror. Some typical images are shown in Figure 5. Automatic adjustment for global illumination changes was achieved in hardware through gain and exposure control, and patch normalization was used to reduce local variation. Because this is an indoor robot application, we assume that the ground surface is locally flat and that the robot is constrained to the ground plane. The recognition process starts with the current unwrapped, patch normalized $w \times h$ pixels panoramic image with the w dimension aligned with the ground plane in the real world, and the h dimension aligned with the vertical plane. Image similarities between the current image and template images are calculated using the cross correlation of corresponding image rows. For each row, this correlation can be performed efficiently in the Fourier domain, in which multiplication is equivalent to convolution in image space:

$$C = F^{-1} \left[\sum_{y=1}^h F(\tilde{i}_y) \cdot F(r_y) \right], \quad (9)$$

where $F()$ is the Fourier transform operator and i_y and r_y are the pixels rows at y in the current and template images, respectively. The value of the maximum real correlation coefficient gives the quality of the match m :

$$m = \max(\text{Re}(C)). \quad (10)$$

A new image template and local view cell is created if the best match for all current image-template image pairings is below a threshold m_{\min} . Fourier transform coefficients are calculated only once for an image and stored by the vision system. The match quality scores for each image pair are used to set the activation levels for the corresponding local view cells:

$$V_i = \max(m_i, 0) \quad \forall i. \quad (11)$$

Experience Map Performance

The experience map is the core global representation that forms the basis for navigation performance. In Figure 6(a), the experiences are plotted as a green circle at the (x, y) coordinates in the stored position \mathbf{p} . Linked experiences are joined with a blue line. The spatial coherence of the maps is evident from comparison with the floor plans shown in Figure 6(b). The maps are not globally accurate in a Cartesian sense as is evident from the change in scale, but global accuracy is not necessary for effective navigation. The local spatial integrity of the map is important for selection of appropriate navigation behaviors, whereas global topological integrity is important for path planning.

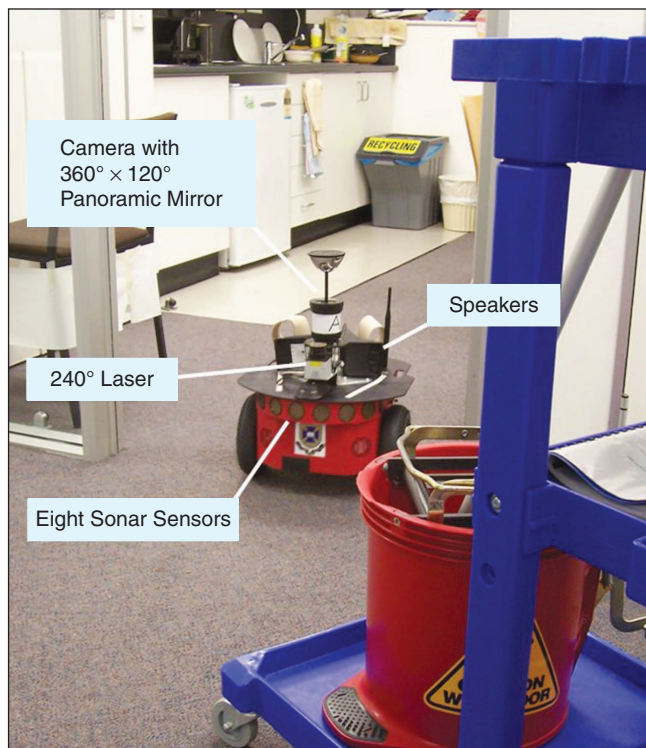


Figure 4. Robot in the indoor test environment.

Path-Planning Performance

For the application of RatSLAM as an autonomous delivery robot, the most important indicator of performance is the success rate for planning and executing paths to the delivery locations. The robot roamed the office building for eight days, with 32 h of active operation in both day and night cycles. During that time, the robot performed 1,089 mock deliveries to six locations and autonomously navigated to and docked with its charger 21 times. The time to complete a delivery

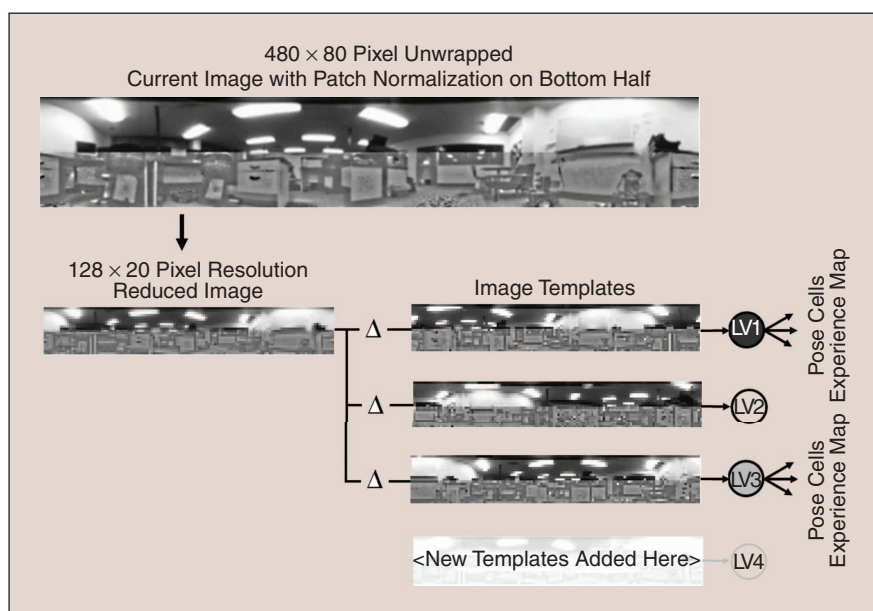


Figure 5. Visual interface to the local view cells.

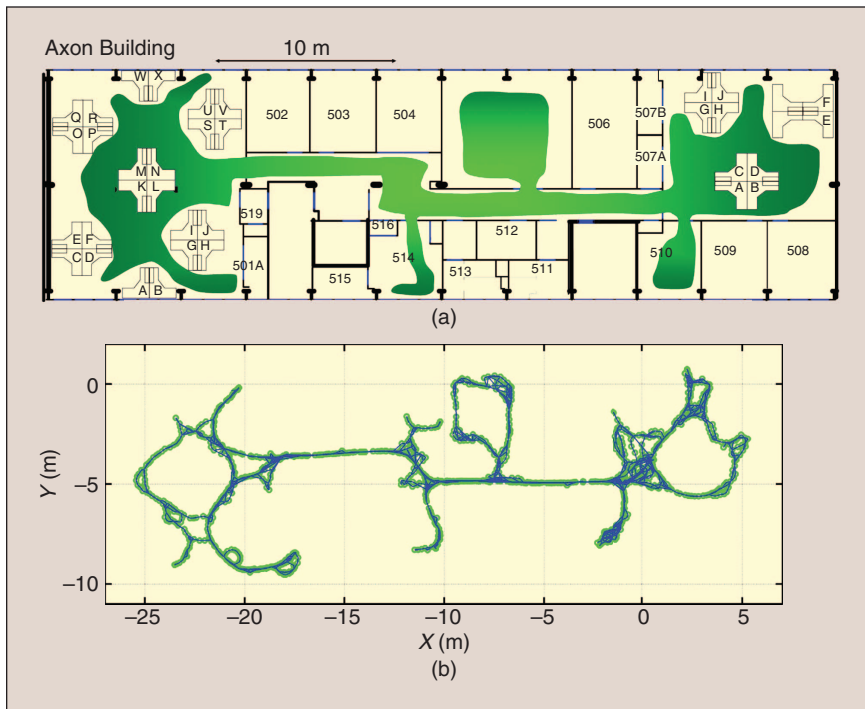


Figure 6. (a) Plan view of the testing environment. (b) The experience map state after 5 h in the testing environment (rotated for comparison).

typically varied from 1–3 min depending on the length of the path. Figure 7 shows an example of the path-planning process in action. The robot failed only to complete a delivery once because of an inconsistency in low-level navigation routines in

in our more detailed publication of this experiment [24].

Outdoor Results

Although it is clear from the tests in the office environment that RatSLAM can create usable maps that can be maintained over long periods, we were interested to test the ability of RatSLAM in mapping very large and complex environments. To test the capacity of the system, we set out to map the entire suburb of St. Lucia in Brisbane, Australia [11], [12]. St. Lucia is a challenging environment to map with many contrasting visual conditions: busy multilane roads, quiet back streets, wide-open campus boulevards, road construction work, tight leafy lanes, monotonous suburban housing, highly varied shopping districts, steep hills, and flat river roads. The road network consists of 51 inner loops of varying size and shape and includes more than 80 intersections including some large roundabouts.

The road network was filmed by mounting a notebook computer on the roof of a car [see Figure 8(a)] and driving the car at normal traffic speeds so that each street in St. Lucia was visited at least once with most streets visited multiple times. The test took place on a typical fine spring Brisbane day starting in the late morning. Images were obtained at a rate of 10 frames/s from the camera and saved to disk as a movie for offline processing.

Interface to RatSLAM

In this experiment, both the local view cells and path-integration processes in RatSLAM were driven from the vision system. The vision algorithms use scanline intensity profiles much like the profiles used in autonomous visual steering [17]. The scanline intensity profile is a one-dimensional (1-D) vector formed by summing the intensity values in each pixel column and then

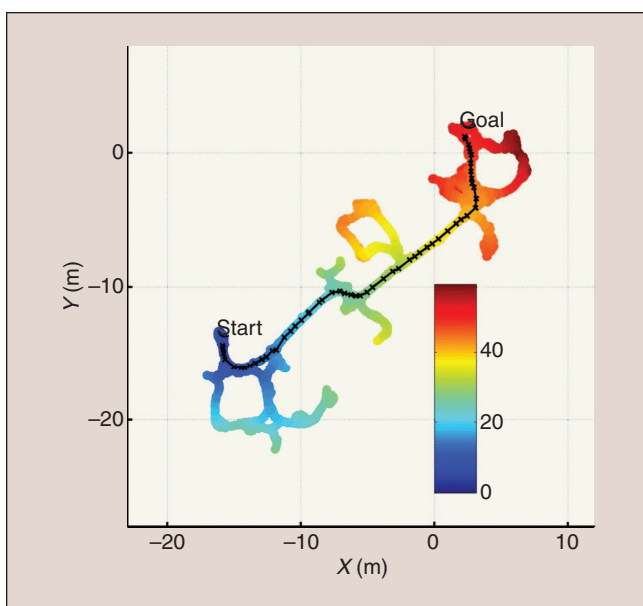


Figure 7. Path planning, showing the temporal map and planned route. The temporal map shows the predicted travel times to any other location in the environment from the robot's current (start) location, whereas the route is calculated from a gradient climbing process. The temporal map and goal route are calculated continually enabling path planning to adapt to map changes instantaneously.

an open laboratory area. The robot timed out the delivery after 5 min and reported the failure before autonomously recovering and resuming operation.

The RatSLAM algorithm does not have separate learning and recall cycles; it can update the map while being used to plan and enact navigation to goals. Because the use and update of the map operates in parallel, there is no need for user intervention to apply changes to the map or to tell the robot to learn new features. In the experiment, the robot continuously updated and augmented its representations during operation over changes caused by day–night cycles or by rearrangement of furniture and accessories in the office. The office environment was represented using approximately 1,000 visual templates and 1,500 experiences. The complete RatSLAM algorithm, including visual processing and map updating, was continuously performed at approximately 7 Hz for the duration of the experiment. Further details of the results have been presented

normalizing the vector [see Figure 8(b)]. This profile is used to estimate the rotation and forward speed between images for odometry and compare the current image with previously seen images to perform local view calibration. Rotation information is estimated by comparing successive scanline intensity profiles and finding the pixel shift that maximizes the similarity between the profiles. Forward speed is estimated from the residual change in the successive profiles once rotation has been removed. Similarity in intensity profiles is also the basis for detecting whether the current image is already associated with a local view cell or it should form the basis of a new local view.

Experience Map

The experience map was created offline but in real time from playback of the data set as a movie. An aerial photograph of the test environment is shown in Figure 9(a) with the route driven in the experiment highlighted in green. The experience map shown in Figure 9(b) contains 12,881 individual experiences and 14,485 transitions between experiences. The map captures the overall layout of the road network as well as the finer details such as curves, corners, and intersections. Because the mapping system only had a perceptual rather than absolute measure of speed, the map is neither geometrically consistent on a global scale nor could it be expected to be.

In creating the experience map, RatSLAM closed and organized multiple interlocking loops including one loop of more than 5,000 m in length. The noisy nature of the visual odometry created an accumulated odometric error of approximately 1,200 m when this loop was closed. Loop closure in RatSLAM is achieved by accumulating evidence based on the appearance of a sequence of visual scenes rather than through geometric processing.

Discussion

In this project, we found that building a mapping system based on an understanding of the highly competent spatial cognition system evidenced by the rat's remarkable navigation abilities has produced new insights into developing long-term and large-scale navigation competence in a robot. The operation of a rat brain clearly has some similarities and many differences to the way that robots currently perform the task of mapping and navigation. RatSLAM is an attempt to build a practical robotic system that can take advantage of the points of difference highlighted by studies of the rat brain, while still being cognizant of the realities of navigation with a robot. The outcome is a system that performs well at some of the most challenging problems in robotic navigation.

The success, we believe, is due to the spatial representations at the heart of the RatSLAM system. The maps are less accurate in their geometry than the maps one might produce from a state-of-the-art SLAM system, but flexibility in geometry helps to create a system that can cope with noisy input, deal with a changing environment, and accommodate increasing complexity. The pose-cell representation, inspired by the grid cells of the rat, forms a powerful spatiotemporal filter for the noisy data association information from the local view matching process. The coherent and maintainable

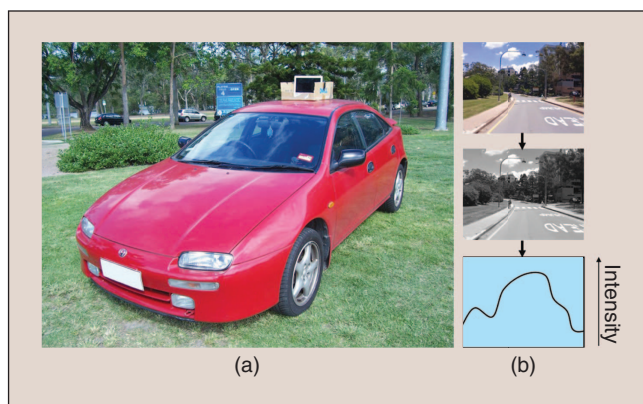


Figure 8. (a) The car with the laptop mounted on the roof and (b) the scanline intensity profile of a typical image.

topology of the experience maps, inspired by the place cells in the hippocampus, provides a resource for straightforward path planning and execution—a major application of robotically acquired maps.

We are continuing to test the limits of the RatSLAM application. How little sensory information can be used? How light can the computational load be made? How sparsely can the pose cells and experiences be packed? Our approach to answering these questions is to develop alternative solutions for components in the RatSLAM system, moving away from the system's biological origins and toward an engineering approach. We hope to create a new generation of lightweight and low-cost mapping and navigation systems that can be deployed in the new generation of domestic and service robots.

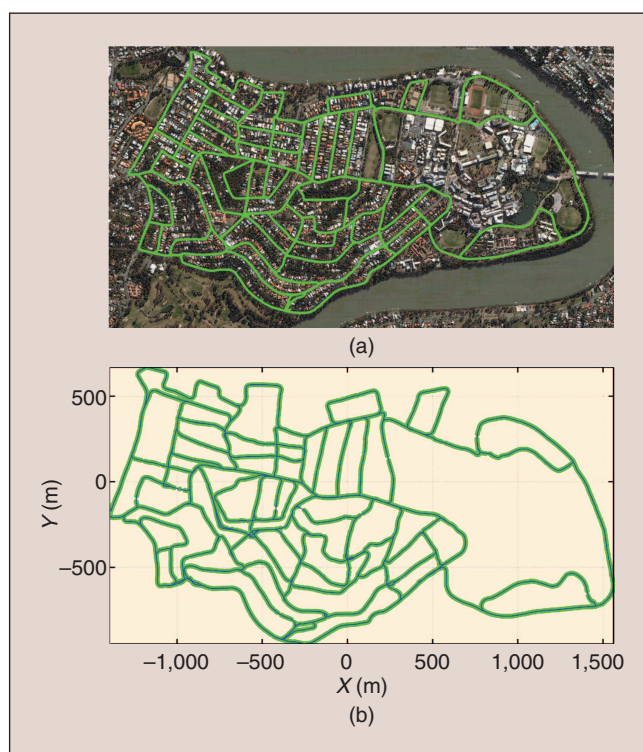


Figure 9. (a) Aerial photo of St. Lucia and (b) corresponding experience map (from [11]).

Our work also proceeds in the other direction—toward biology. We are working closely with neuroscientists to develop high-fidelity spiking models of neural systems that bridge the gap between RatSLAM's abstract model of hippocampus and the reality of neural activity in the rodent brain. A model at this level requires many more details of the biology to be implemented, such as the correct relative numbers of neurons and synapses, the timing of spiking behavior, and the overlying brain oscillations such as the theta rhythm. This work is more challenging in its complexity, as each neuron within the model is a dynamical system in its own right. By testing large-scale in silico simulations of the hippocampus in a robot platform, we hope to reveal further mysteries of spatial cognition in the mammalian brain.

Keywords

Neurorobotics, biologically inspired robots, learning and adaptive systems, SLAM.

References

- [1] A. Arleo and W. Gerstner, "Spatial cognition and neuro-mimetic navigation: A model of hippocampal place cell activity," *Biol. Cybern.*, vol. 83, no. 3, pp. 287–299, 2000.
- [2] A. Barrera and A. Weitzenfeld, "Biologically-inspired robot spatial cognition based on rat neurophysiological studies," *Autonom. Robots*, vol. 25, no. 1–2, pp. 147–169, 2008.
- [3] P. J. Best, A. M. White, and A. Minai, "Spatial processing in the brain: The activity of hippocampal place cells," *Annu. Rev. Neurosci.*, vol. 24, pp. 459–486, 2001.
- [4] D. E. Davis, J. T. Emlen, and A. W. Stokes, "Studies on home range in the brown rat," *J. Mammal.*, vol. 29, no. 3, pp. 207–225, 1948.
- [5] G. Dissanayake, P. M. Newman, S. Clark, H. Durrant-Whyte, and M. Csorba, "A solution to the simultaneous localisation and map building (SLAM) problem," *IEEE Trans. Robot. Automat.*, vol. 17, no. 3, pp. 229–241, 2001.
- [6] C. Giovannangeli and P. Gaussier, "Autonomous vision-based navigation: Goal-orientated planning by transient states prediction, cognitive map building, and sensory-motor learning," presented at the Int. Conf. Intelligent Robots and Systems, Nice, France, 2008.
- [7] T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, and E. I. Moser, "Microstructure of a spatial map in the entorhinal cortex," *Nature*, vol. 436, pp. 801–806, 2005.
- [8] J. Knierim, H. Kudrimoti, and B. McNaughton, "Place cells, head direction cells, and the learning of landmark stability," *J. Neurosci.*, vol. 15, no. 3, pp. 1648–1659, 1995.
- [9] J. L. Krichmar, D. A. Nitz, J. A. Gally, and G. M. Edelman, "Characterizing functional hippocampal pathways in a brain-based device as it solves a spatial memory task," *Proc. Nat. Acad. Sci. USA*, vol. 102, no. 6, pp. 2111–2116, 2005.
- [10] B. L. McNaughton, F. P. Battaglia, O. Jensen, E. I. Moser, and M. B. Moser, "Path-integration and the neural basis of the cognitive map," *Nat. Rev. Neurosci.*, vol. 7, no. 8, pp. 663–678, 2006.
- [11] M. Milford and G. Wyeth, "Mapping a suburb with a single camera using a biologically inspired SLAM system," *IEEE Trans. Robot.*, vol. 24, no. 5, pp. 1038–1053, 2008.
- [12] M. Milford and G. Wyeth, "Single camera vision-only SLAM on a suburban road network," presented at the Int. Conf. Robotics and Automation, Pasadena, CA, 2008.
- [13] M. J. Milford, G. Wyeth, and D. Prasser, "RatSLAM: A hippocampal model for simultaneous localization and mapping," presented at the Int. Conf. Robotics and Automation, New Orleans, LA, 2004.
- [14] M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit, "FastSLAM 2.0: An improved particle filtering algorithm for simultaneous localization and mapping that provably converges," presented at the Int. Joint Conf. Artificial Intelligence, Acapulco, Mexico, 2003.
- [15] J. O'Keefe and D. H. Conway, "Hippocampal place units in the freely moving rat: Why they fire where they fire," *Exp. Brain Res.*, vol. 31, no. 4, pp. 573–590, 1978.
- [16] J. O'Keefe and J. Dostrovsky, "The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely moving rat," *Brain Res.*, vol. 34, no. 1, pp. 171–175, 1971.
- [17] D. Pomerleau, "Visibility estimation from a moving vehicle using the RALPH vision system," presented at the IEEE Conf. Intelligent Transport Systems, 1997.
- [18] J. J. B. Ranck, "Head direction cells in the deep cell layer of dorsal presubiculum in freely moving rats," *Abstr. Soc. Neurosci.*, vol. 10, p. 599, 1984.
- [19] S. M. Stringer, E. T. Rolls, T. P. Trappenberg, and I. E. T. de Araujo, "Self-organizing continuous attractor networks and path integration: Two-dimensional models of place cells," *Network*, vol. 13, no. 4, pp. 429–446, 2002.
- [20] S. M. Stringer, T. P. Trappenberg, E. T. Rolls, and I. E. T. de Araujo, "Self-organizing continuous attractor networks and path integration: One-dimensional models of head direction cells," *Network*, vol. 13, no. 2, pp. 217–242, 2002.
- [21] J. S. Taube, R. U. Muller, and J. J. B. Ranck, "Head direction cells recorded from the postsubiculum in freely moving rats—Part I: Description and quantitative analysis," *J. Neurosci.*, vol. 10, no. 2, pp. 420–435, 1990.
- [22] S. Thrun and M. Montemerlo, "The GraphSLAM algorithm with applications to large-scale mapping of urban structures," *Int. J. Robot. Res.*, vol. 25, no. 5–6, pp. 403–429, 2006.
- [23] E. C. Tolman, "Cognitive maps in rats and men," *Psychol. Rev.*, vol. 55, no. 4, pp. 189–209, 1948.
- [24] M. Milford and G. Wyeth, "Persistent navigation and mapping using a biologically inspired SLAM system," *Int. J. Robot. Res.*, to be published.
- [25] J. S. Taube, "The head direction signal: Origins and sensory-motor integration," *Annu. Rev. Neurosci.*, vol. 30, pp. 181–207, 2007.

Gordon Wyeth received his B.E. degree in 1989 and Ph.D. degree in 1997 in computer systems engineering from the University of Queensland, Brisbane, Australia. He is a senior lecturer in information technology and electrical engineering at the University of Queensland. He is the codirector of mechatronic engineering, with research interests in biologically inspired robot systems, developmental robotics, and robot education. He is a chief investigator on the Thinking Systems project, a joint Australian Research Council and National Health and Medical Research Council (ARC/NHMRC) funded initiative and has served as a chief investigator on several Australian Research Council and industry-funded projects. He served as president of the Australian Robotics and Automation Association (2004–2006) and has twice chaired the Australasian Conference on Robotics and Automation.

Michael Milford received his B.E. degree in mechanical and space engineering in 2002 and Ph.D. degree in electrical engineering in 2006, both from the University of Queensland. He is a research fellow at the Queensland Brain Institute and the School of Information Technology and Electrical Engineering at the University of Queensland. His research interests include biologically inspired robot mapping and navigation and the neural mechanisms underpinning navigation in animals.

Address for Correspondence: Gordon Wyeth, Information Technology and Electrical Engineering, University of Queensland, St. Lucia 4072, Australia. E-mail: wyeth@itee.uq.edu.au.