Neural Model of a Grid-Based Map for Robot Sonar

Kenneth D. Harris* and Michael Recce †
Department of Anatomy and Developmental Biology
University College London
Gower Street, London WC1E 6BT, UK
Email: kdh@anat.ucl.ac.uk, m.recce@ucl.ac.uk

Abstract

A functional similarity is described between cells of an occupancy grid for robot sonar, and integrate-and-fire neurons of an artificial neural net. Using this analogy, a new grid-based mapping system for robot sonar is described, which makes use of the neural concepts of receptive fields and recurrent connections. The performance of the new network is compared to that of a previous Bayesian grid-based mapping method, and a previous feature-based mapping method.

1 Introduction

There are certain situations for which it is useful for a mobile robot to have a map of its environment. Many techniques have been suggested for the construction, representation and use of such maps.

In the simplest cases, a map is constructed by a human, digitised, and programmed into the robot. In this case, the map is known to be accurate, and only the robot's position relative to the map needs to be obtained by the robots sensors. However, this method cannot be used in circumstances where no human-constructed map was initially available, or in which the environment changes with time. In such cases the robot must construct the map itself, using its own sensors.

For a robot with a sonar sensor, two methods of map representation have been in common use up to now. In *grid-based representation* [1], the two dimensional space of the robot's environment is overlaid with a square grid. For each grid segment, a set of numbers is stored representing sensory information about the contents of that grid segment in physical space. In the simplest cases, known as *certainty grid* methods [2] [3], a single number is stored representing the probability that the segment is occupied. In *feature-based representation* [4] the map consists of a list of physical features, such as walls and pillars, and the positions of these features are represented by Cartesian coordinates.

Grid based representation has the advantage that a grid map of free space is useful for path planning. Furthermore, the probabilistic nature of grid maps allows them to represent uncertain information. By contrast in a feature-based map, features are only added to the map when their existence is known for certain. This allows grid-based maps to be built up more quickly, as it takes many sonar readings to confirm the existence of a feature, but every sonar reading can be used to update the probabilities stored in a grid-based map.

Cells of an certainty grid function similarly to neurons in an artificial neural network. In the certainty grid method, a single number is stored for each grid segment, corresponding to the probability that the segment is occupied. When a sonar reading suggests the segment is occupied, this number is increased, and when a sonar reading suggests the segment is free, this number is decreased. If the number exceeds a threshold, the segment is estimated to be occupied, otherwise is it estimated to be free. A grid segment may thus be described as a formal integrate-and-fire neuron, with membrane potential corresponding to the probability of occupancy, and firing occurring when the segment is estimated to be occupied. Making the analogy to neural networks allows us to use concepts such as receptive fields and lateral connections, that are familiar from the neural network and neuroscience literature.

The derivation of the rules for updating the probabilities stored in an certainty grid makes the assumption that the occupancy probabilities of all cells are independent. This assumption is, in general, not justified. Walls and empty spaces in the environment are usually continuous, and, in man-made environments, walls are usually straight. A simple grid-based approach does not make use of this prior information about the environment. In this paper we present an extension of grid based mapping in which we do not assume the occupancy of grid segments is independent. Lateral connections between the neurons for neighbouring segments reinforce patterns likely to be found in the environment, such as straight walls, and inhibit patterns that are unlikely to be found, such as random scatter.

2 The network

A robot can represent a map of space in either egocentric coordinates, meaning that the positions of features are expressed relative to the robot, or in allocentric coordinates, meaning that the positions of features

*Supported by a National Science Foundation Graduate Fellowship.

†New address: Department of Computer Science, New Jersey Institute of Technology, University Heights, Newark NJ 07102-1982, USA

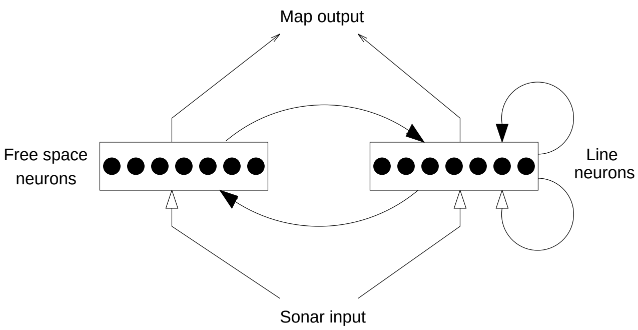

Figure 1: The network architecture. Open arrows indicate excitatory connections, filled arrows represent inhibitory connections. The network consists of two sets of topographically arranged neurons, the free space neurons and the line neurons. Both sets receive excitatory feed-forward input from sonar, and there are excitatory and inhibitory recurrent connections between the neurons.

and of the robot are expressed in a fixed, world-centred frame. Both types are in common use in robotics. In this paper we use egocentric coordinates, inspired by a previous neurobiological theory which proposes that an egocentric representation is used in the brain [5].

The space surrounding the robot is divided into a 10 centimeter square grid, and for each grid segment there are several neurons, coding for different types of features in the environment. The neurons are therefore topographically arranged, with each neuron representing the presence of a preset feature type in a preset region of egocentric space. When we refer to the "position" of a neuron, we mean the position of its underlying grid segment in egocentric space. The neurons have feed-forward inputs derived from the sonar sensor, and recurrent inputs coming from other neurons. All neurons are of integrate and fire type, with a membrane potential x_i, binary output y_i, and feed-forward input z_i obeying the following equation:

$$x_i(t+1) = x_i(t) + z_i + \sum_j w_{ij} y_j(t)$$

$$y_i(t) = \Theta(x_i - 1.0)$$

The weights w_{ij} of the recurrent connections are fixed, and do not learn. Map information is stored in the activation pattern of the neurons, rather than in the synaptic weights. The integrate and fire type neurons allow sonar input to be accumulated over several time steps, and therefore from several different robot positions. Note that in this model the integrated membrane potential does not passively decay. All neurons have a fixed threshold of 1.0 and no bias. This threshold is sufficiently high so that on average many inputs are required to make the neuron active.

Figure 1 shows a schematic diagram of the network described here. The network consists of two types of neurons. For each egocentric grid segment there is a free space neuron, and a set of line neurons. The firing of a free space neuron signals that the corresponding region of egocentric space is unoccupied. For each grid segment there are 16 line neurons, parametrised by a preferred orientation which ranges from 0^o to 180^o. The firing of a line neuron signals that a straight wall of the correct orientation passes through the grid segment. The encoding of the orientation of features in the environment is consistent with the many forms of edge coding cells that have been found in the visual system of animals, except that in this model they are encoding the egocentric location of the walls in the environment, instead of the retinotopic location of edges in a visual scene.

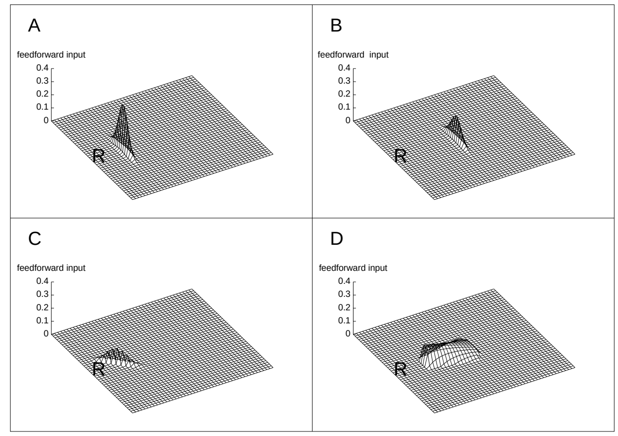

Figure 2: Receptive fields of various neurons. The figures show the excitatory input to the neuron due to a single sonar return, as a function of sonar beam reflection position. The letter "R" denotes the position of the robot. The square mesh has a $10cm$ spacing. A) Receptive field of the line neuron for the grid segment $1m$ east of the robot with north-south orientation. B) Receptive field of the line neuron for the grid segment $1.5m$ east of the robot with north-south orientation. C) Receptive field of the line neuron for the grid segment $1m$ east of the robot with northwest-southeast orientation. D) Receptive field of the free space neuron for the grid segment $1m$ east of the robot.

The external input z_i of a neuron is calculated as follows. The neuron receives an excitatory input for every sonar return, and the inputs from all the sonar returns are added to produce the total input for the neuron. For a given sonar return, the strength of input the neuron receives depends on the position of the sonar reflection point in robot-centred space, the type of neuron, and its position in the grid (and orientation for line neurons). Example graphs of input strength versus reflection point position are shown in figure 2 for various neurons. A given neuron's input is close to zero except in a certain region of space. Following the neuroscience literature, we will call this region the *receptive field* of the neuron. For a line neuron, the receptive field is elongated, directed along the neurons preferred orientation. The exact formula for the input is a product:

$$z_{line} = C_{line} \times F_{pos} \times F_{ang} \times F_{RCF}$$

C_{line} is a constant, equal to 0.5 for the simulations described below. F_{pos} is a positional factor which gives the receptive fields their oblong shape. It is given by a double Gaussian:

$$F_{pos} = e^{-\left(d_{par}^2/2\sigma_{par}^2 + d_{perp}^2/2\sigma_{perp}^2\right)}$$

where d_{par} is the distance of the sonar reflection point from the line neuron position in the direction parallel with the neuron's orientation, d_{perp} is the distance of the sonar reflection point from the line neuron position in the direction perpendicular to the neuron's orientation, and σ_{par} and σ_{perp} are constant distances equal to $300mm$ and $40mm$ respectively.

F_{ang} is an angular factor,to take into account the fact that sonar is preferentially reflected perpendicularly from specular walls by increasing the input of those line segments whose orientation is perpendicular to the reflected sonar beam. It is given by a Gaussian:

$$F_{ang} = e^{-\Delta\theta^2/2\omega^2}$$

where $\Delta\theta$ is the angular difference between the sonar beam angle and a perpendicular to the neurons orientation, and ω is the sonar beam half-width, 15^o.

F_{RCF} is a *range confidence factor* [6], whose purpose is to reduce the contribution of long range readings, which are more likely to be specular. It has the form

$$F_{RCF} = \left(1 - \frac{returning\_range}{max\_detect\_range \times range\_weight}\right)^k$$

where $returning\_range$ is the measured length of the sonar beam, max_detect_range is the maximum measurable length ($2.53m$ for our sensor), and $range\_weight$ and k are constants equal to 1.1 and 0.8 respectively (values given in [6].

Figure 2a shows the input strength as a function of sonar reflection position, to a neuron from a grid segment $1m$ to the east of the robot with orientation in the north-south direction. Figure 2b shows the input to a neuron $1.5m$ to the east of the robot, with the same orientation. The overall strength of input to this neuron is lower, due to the range confidence factor. Figure 2c shows the input strength to a neuron from the same grid segment as that in figure 2a, but oriented at a 45^o angle. This neuron also has a lower overall input strength than the neuron of figure 2a, due to the angular factor.

The length of a line neuron receptive field is much greater than the grid size. The line neurons therefore perform line detection in a similar manner to a Hough transform. For a given wall in the environment, sonar beams will be reflected from many points along the wall, all lying in a straight line, and will therefore give a large input to a line neuron whose orientation is collinear with the wall.

For a free space neuron, the receptive field consists the region of space from which the sonar beam would not have been reflected if the the part of space the neuron corresponds to was occupied. This region consists of all points which are further away from the robot than the neuron, and whose bearing from the robot is less than one sonar beam width different from the neurons bearing. The strength of the input is given by:

$$z_{fs} = C_{fs} \times F_{pos} \times F_{RCF}$$

where C_{fs} is a constant equal to 0.2, F_{RCF} is as above, and F_{pos} is a positional factor given by

$$F_{pos} = 1 - \Delta\phi^2/\omega^2$$

where $\Delta\phi$ is the angular difference between the sonar beam and the bearing to the neuron, and ω is the beam half-width as above. An example of a free space neurons receptive field is shown in figure 2d.

The recurrent connections of the network are summarised in figure 3. A line neuron is connected to all other line neurons lying in a rectangular box the size of its receptive field (size $2\sigma_{par} \times 2\sigma_{perp}$). If the second neuron has the same orientation, the connection is excitatory, with weight 0.01. If the second neuron has a different orientation, it is inhibitory, with weight -0.02. The aim of these connections is to enforce collinearity of line segments corresponding to a single wall in the environment. A line neuron also inhibits with weight -0.3 all free space neurons lying in its receptive field, and receives an inhibitory connection from them of weight -0.2. The aim of these connections is ensure that specular reflections do not cause line neurons to erroneously fire in unoccupied regions, or free space neurons to erroneously fire in occupied regions.

At each time-step, a sonar scan is taken, the feedforward inputs are calculated, and the membrane potential of the neurons are updated. Any neurons which

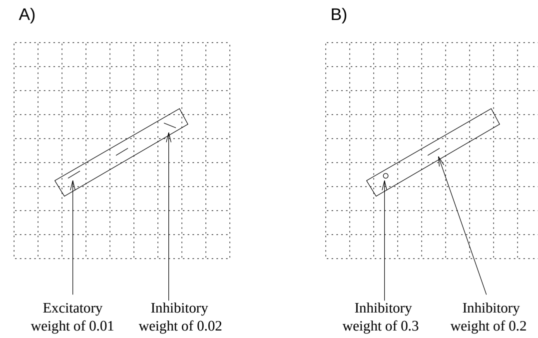

Figure 3: This figure illustrates the recurrent connections between neurons. A) A line neuron is connected to all line neurons for grid segments within its receptive field. If the two line neurons have the same orientation, the connection is excitatory, otherwise it is inhibitory. B) A line neuron inhibits all free space neurons in its receptive field, and is reciprocally inhibited by them.

exceed their threshold will fire. The recurrent connections are then activated once, and a map is output on the basis of the neural firing pattern. The robot then makes a movement. After movement, the entire neural activation pattern, including sub-threshold potentials, is shifted through the distance the robot has moved, so it remains in an egocentric coordinate system. A new time-step then begins.

When the robot is initially placed in an environment all of the neurons have membrane potential zero. The initial sonar scan provides sub-threshold inputs to several neurons in the network. When the robot moves, this potential is transferred to the neurons corresponding to the new egocentric position of the features that caused the initial sub-threshold potential, by the shifting mechanism just described. A second scan receiving reflections from the same features will activate these same neurons, allowing sonar information from several viewpoints to be integrated.

3 Experimental Verification

We tested the above map-making system with sonar data collected from a mobile robot. The robot, ARNE, which was used to gather the data, has a 25 cm diameter circular base, and two wheels which allow it to move forward or backward, and rotate on the spot. ARNE's wheels are equipped with shaft encoders that provide a measure of the distance moved and the rotation angle of a turn. ARNE has a single sonar sensor on a motorised pivot, which is used for obstacle avoidance during movement.

In order to asses the performance of our model in comparison to other methods, we also implemented a Bayesian grid based mapping system. We found that the grid based methods of Elfes [1], Moravec [2], and Cho [3] performed poorly on sonar data collected from our robot due to specular reflection from smooth wood walls in the environment. We therefore implemented the method of Lim and Cho [6], which was designed to overcome the problems of specular reflection. In this method, one number is stored for every grid segment that corresponds to the probability of occupancy of the segment. Also a set of numbers are stored that correspond to the orientation probabilities of the segment, given that the segment is occupied. For the parameter C (unspecified in [6], we used the value 0.001, as we found this gave best results. In order to produce output that is visually comparable to that produced by our model, we thresholded the occupancy probabilities. If a segment has occupancy probability less than 0.3, it is judged to be empty, if it has occupancy probability greater that 0.7, it is judged occupied, otherwise it is marked unknown. Occupied cells were drawn on the output maps as line segments, with orientation given by the maximum of the orientation probability stored for the segment.

We also compared the performance of our model to that of the model of Lee and Recce [7], [8]. In this system, a non-probabilistic grid-based map is used to represent free space, but the position of walls in the environment is represented by sets of Cartesian coordinates for their endpoints.

We tested the three models on identical copies of

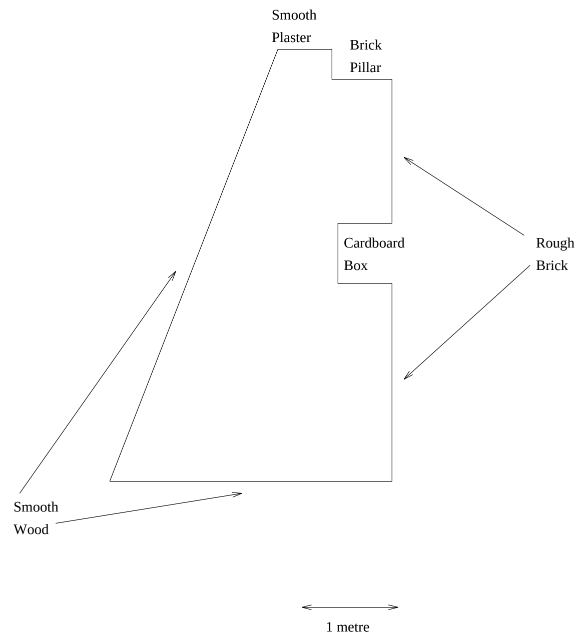

Figure 4: The environment used to collect sonar data. The environment contains a mix of specular and non-specular surfaces.

sonar data collected from several runs in several different environments. Figure 5 shows the maps produced 10 and 30 time-steps into a run in the environment shown in figure 4. During the data collection runs, the robot followed a simple wall following strategy [7], maintaining a distance of 35cm between itself and the wall on its right. During each time-step the robot attempted to move one length along the wall. It stopped short of this full length step if an obstacle was detected in its path. The location of the robot at each of the time steps is shown by an unfilled circle in figure 5. After each movement ARNE makes a sonar range scan, taking measurements every 18°. The maximum sonar range measurement is 2.5 metres, which is shorter than the length of the environment.

By 10 time-steps, all 3 systems produced a partial map of the environment, and by 30 time-steps, they produced a fairly complete map. The neural and Bayesian methods produce broadly similar maps. However, the effect of the lateral connections can be seen in the neural method. For example there is a greater coherence and collinearity of line segments that correspond to walls. Also in the Bayesian method there are incorrectly identified line segments in the middle of free space regions (see figure 5), that would cause the robot to make unneeded detours, and could make free space regions unreachable.

The feature based method is slightly slower than the other two at accumulating a map, but is the most accurate at representing the positions of the walls. This is because of its non-probabilistic nature, whereby it must be sure of the existence of a feature before adding it to the map. For example, even by

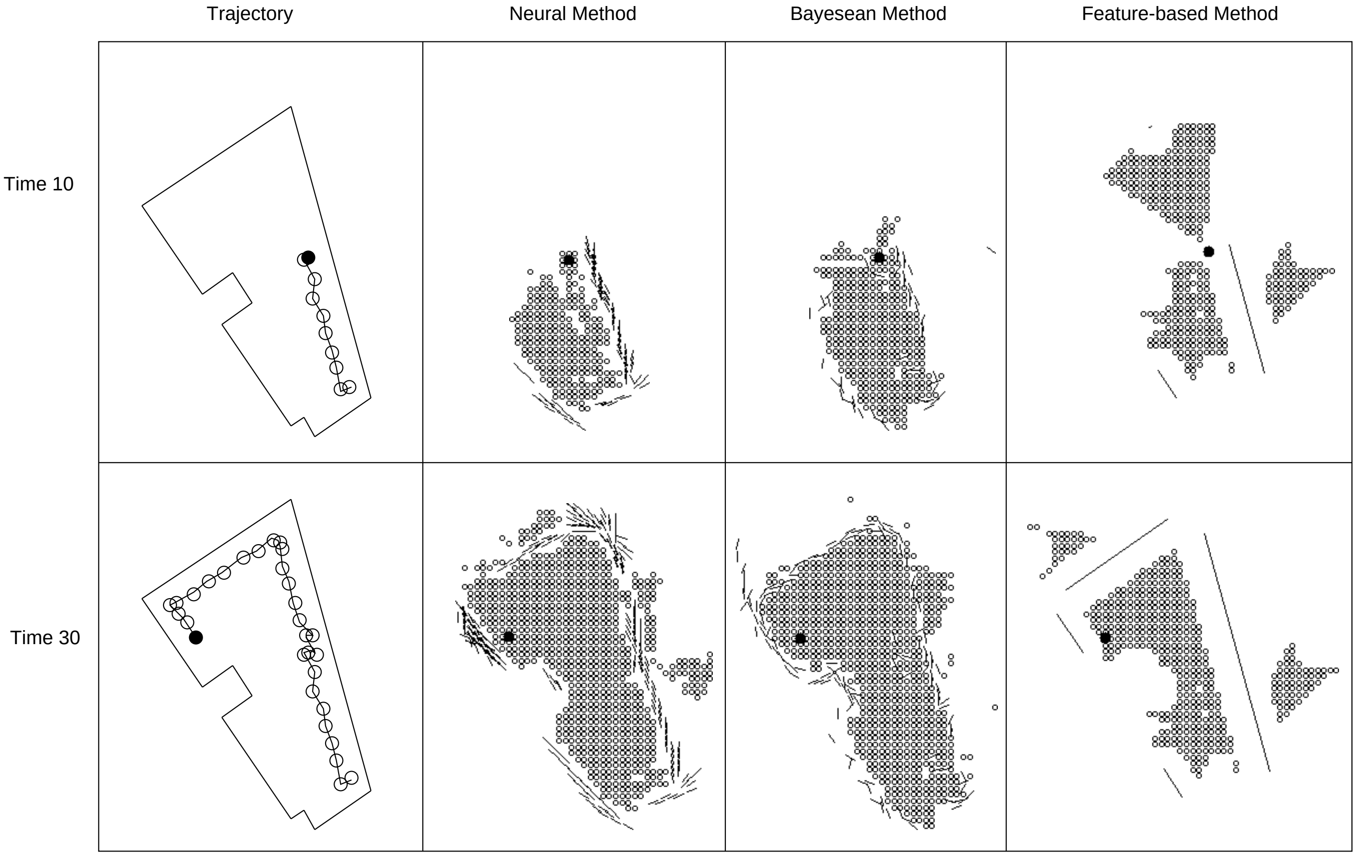

Figure 5: The maps produced by the neural method described in this paper, the Bayesian method of [6], and the feature based method of [7], from the same sonar data, after 10 and 30 time-steps. Open circles represent areas of space marked as free, and line segments or lines represent walls. The filled circle represents the robot position.

time 30, this method has not detected the cardboard box or smooth plaster wall. Note that the smaller size of the free space area with this method is due to a "safety zone" of one robot diameter placed around the walls by the mapping software, rather than a failure to detect free space.

4 Discussion

In this paper, we described functional similarities between grid-based maps for robot sonar and artificial neural networks. Making use of this analogy, we proposed a new grid based mapping system making use of the neural concepts of receptive fields and lateral connections. Neurons with elongated receptive fields detect sonar beam returns corresponding to straight walls in a similar manner to a Hough transform. The lateral connections in the network are designed to reinforce maps that represent configurations likely to occur in the world, like straight walls, and inhibit configurations unlikely to occur, such as random scatter.

The line orientation receptive fields used in the model are consistent with the experimentally observed properties of neurons in the visual pathway. Furthermore cells have been found in the temporal cortex which maintain the memory of an object through sustained activity [9], similarly to the way cells in our network continue to code for features even when they are further from the robot than the maximum sonar range. Previous models [10];[11] have described neural mechanisms by which an entire neural activation pattern can be shifted to remain in egocentric coordinates after movement. This shift of activity pattern is the equivalent of dead-reckoning in egocentric coordinates. In the current implementation we simply shift the array of membrane potentials with a loop. Implementing a neural mechanism for this shift would make the network more biologically realistic but it would not contribute to its performance as an engineering method for guiding the navigation of a robot.

We compared the new network to an established grid-based mapping system, and to a feature-based mapping system, and found that the representation of walls by line segments is more coherent in the new neural method than in the Bayesian method, and that both the neural and Bayesian methods were faster at building a map than the feature based method. Quantitative measures across a range of environments are required to more completely differentiate the performance of these methods (Harris and Recce, in prep).

The new method is based on the use of a single layer of neurons, and if implemented on parallel neural hardware, the map would updated one computational time-step after the presentation of the new sensory input.

The current model represents only the first step in neural-based sonar mapping. The real strength of neural networks is their ability to learn using real world data. The net we described here, however, has fixed weights, and a large number of parameters whose values were not systematically optimised. Future work will aim to incorporate learning into neural sonar mapping algorithms, and thereby reduce the number of free parameters. However, even without learning, the neural method described here produces results comparable with a leading Bayesian method.

References

[1] A. Elfes. Sonar-based real-world mapping and navigation. *IEEE Journal of Robotics and Automation*, 3:249–265, 1987.

[2] H. P. Moravec. Sensor fusion in certainty grids for mobile robots. *AI Magazine*, 9:61–74, 1988.

[3] D. W. Cho. Certainty grid representation for robot navigation by a bayesian method. *Robotica*, 8:159–165, 1990.

[4] J. J. Leonard and H. F Durrant-Whyte. *Directed sonar sensing for mobile robot navigation.* Kluwer Academic Publishers, Cambrage, MA, 1992.

[5] M. Recce and K. D. Harris. Memory for places: a navigational model in support of Marr's theory of hippocampal function. *Hippocampus*, 6:735–748, 1996.

[6] J. H. Lim and D. W. Cho. Physically based sensor modeling for a sonar map in a specular environment. In *proceedings of the 1992 IEEE international conference on robotics and automation*, pages 1714–1719, 1992.

[7] D. Lee and M. Recce. Quantitative evaluation of the exploration strategies of a mobile robot. *Int. J. Robotics Research*, in press.

[8] D. C. Lee. *The Map-Building and Exploration Strategies of a Simple, Sonar-Equipped Mobile Robot; An Experimental, Quantitative Evaluation.* Distinguished Dissertations in Computer Science. Cambridge University Press, 1996.

[9] Y. Miyashita. Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature*, 335:817–820, 1988.

[10] J. Droulez and A. Berthoz. A neural network model of sensoritopic maps with predictive and short-term memory properties. *Proc. Nat. Acad. Sci.*, 88:9653–9657, 1991.

[11] K. Zhang. Representation of spatial orientation by the intrinsic dynamics of the head-directon cell ensemble: A theory. *J. Neurosci.*, 16:in press, 1996.