

A COMPUTATIONAL NEUROSCIENCE PERSPECTIVE OF THE HIPPOCAMPAL PLACE CELL PHENOMENON

P.G. Madhavan, Ph.D.

Assistant Professor

Electrical Engineering Department

Purdue University at Indianapolis

Neurology Department

Physiology & Biophysics Department

Indiana University School of Medicine

ABSTRACT

Computational neuroscience is a newly emerging field formed by the intersection of electrical engineering, biomedical engineering, physiology, biophysics, computer science, mathematics and anatomy. The brain is considered to be a computing device and is studied at three interacting levels - (i) computational theory, (ii) algorithms and (iii) implementation. After a brief introduction to the methodology of computational neuroscience, we demonstrate the methods by modeling the hippocampal place cell phenomenon. The hippocampus is an area of the brain that has been demonstrated to be involved in spatial cognitive processing. Place cells in the hippocampus represent spatial locations based on landmarks in their environment. After a detailed review of the hippocampal place cell phenomenon, we present an artificial neural network model as well as a computational model of this phenomenon. We also point out their use in developing new and robust methods of map building and navigation for autonomous mobile vehicles when position information is not externally supplied.

1. INTRODUCTION

As we approach the twenty-first century, there is real hope that a fairly clear understanding of how the human brain works can be achieved in the coming decades. To make this discovery a reality, there is a developing consensus that multi-disciplinary skills have to be brought to bear on the problem. Among the multiple disciplines involved, electrical engineering, biomedical engineering, physiology, biophysics, computer science, mathematics and anatomy have come together to create the emerging discipline of **Computational Neuroscience**.

To quote from the article in *Science* by Sejnowski, Koch and Churchland [45], "One of the major research objectives of computational neuroscience is to discover the algorithms used in the brain". This is typically achieved by "studying simplified models of the brain which can provide a conceptual framework for isolating the basic computational problems and understanding the computational constraints that govern the design of the nervous system". Physiology and anatomy provide the experimental data; biophysics and mathematics provide the analytical tools; and electrical engineering, biomedical engineering and computer science provide the link between theory and practice. Successful computational models are firmly based on available data and they evolve as new experimental data become available.

In this article, we approach the study of the "place cell" phenomenon demonstrated by the hippocampal region of the animal brain from the perspective of computational neuroscience. This will serve as a case study where the principles of computational neuroscience such as model development, interaction between anatomical data and the quantitative model and the use of artificial neural networks will be demonstrated.

1.1 Computational Neuroscience - A Brief Introduction

The scientific activity currently undertaken within the realm of computational neuroscience can be traced back at least to the emergence of "Cybernetics" in 1948 [51]. The attempts to explicitly incorporate quantitative methods in the study of the brain have accelerated from that time onwards. Despite the fact that the primary advances in the study of the brain in the intervening years have been in the areas of electrophysiology, neurochemistry and molecular biology, quantitative modeling has seen spurts of growth. Some landmark events during this time are the development of the formal neuron [20], rise of perceptrons and neural networks [41], fall of perceptrons [25] and re-emergence of neural networks in the late 80's. In fact, the enthusiasm about the field and the anticipation of impending breakthroughs in our understanding of the brain have prompted the declaration of the 1990's as the "Decade of the Brain" in the United States.

1.2 What is Computational Neuroscience?

Simply stated, it is the study of the brain as a computing device. One can consider the whole brain or its subsystems at the neuronal or molecular levels and ask the following questions - (i) what does this subsystem compute?, (ii) what are the steps in this computation? and (iii) how is the computation implemented? Answers to these questions constitute the subject matter of the computational neuroscience field. The three questions we have posed correspond to the levels of analysis proposed by Marr [19]. He considers an information processing device at the computational theory, algorithm and hardware implementation levels.

Address for Correspondence : Room SL 2156 H, 723 West Michigan Street, Indianapolis, Indiana, U.S.A. 46202

The computational theory level involves identifying the computational purpose of the subsystem and the transformations applied on the incoming data to obtain the final representation within the subsystem. At the algorithm level, the computational steps required to perform the transformations are identified. The implementation level provides the specific details of how the algorithms are executed within the brain subsystem. Such a hierarchy has close correspondence to the organization of topics in digital signal processing. Consider the following hypothetical example. Assume that there is a subsystem of the brain that computes power spectral density (PSD). According to our hierarchy above, the approach will be to identify the theoretical method of PSD used and how the PSD is represented, the algorithm used for its calculation and the hardware implementation. More specifically, given that the input whose PSD to be calculated is ergodic, we have the possibility of defining the PSD as the Fourier Transform of the autocorrelation function (by the Wiener-Khinchine theorem) or as the appropriate limit of the expectation of the periodogram. Based on which of the two definitions of the PSD is used, the algorithm level may involve the Fast Fourier Transform, autocorrelation estimation methods and windowing and segmenting of data in the periodogram calculations. At the implementation level, similar to the hardware implementation issues we may have to consider issues such as the ICs and power supplies available, circuit lay out and word length effects, we will have to identify the molecular and synaptic mechanisms and neuronal and network interconnections that execute the various steps in the algorithms identified. A little reflection will show that the three levels of study are not independent. In the hypothetical example above, the actual implementation constraints alone may help us choose between the autocorrelation method and the periodogram method as the one actually used by the brain subsystem. This is also true in identifying the representation scheme by which the PSD values are "displayed" within our hypothetical brain subsystem. This makes clear the difficulty of a pure top-down approach, where the computational theory may be entirely valid but has no merit as a model of brain function.

The possibility that in an attempt to model some brain subsystems, we may uncover new and useful computational theories (that may or may not be valid models of brain function) should be of great interest to engineers seeking optimal solutions to practical problems. Thus, while the top-down approach may be sterile from a computational neuroscience point of view, it can be of high utility in one or more of the multiple disciplines that converge to form computational neuroscience. After all, even if airplanes do not flap their wings like birds (and hence they are unrealistic neurobiological models of bird flight), they provide a perfectly adequate means of flying most of the time!

The brief outline of the issues in computational neuroscience given above should not be taken to imply that there is complete convergence of opinions regarding

terminology, philosophical basis, methodology and other issues. In fact, the point of view that the computational metaphor of brain function is similar to the ancient clockwork metaphor of brain function updated to be current, has been expressed^[5]. More view points and discussion are available in a volume entitled, "Computational Neuroscience", edited by Schwartz^[44]. In summary, it is undeniable that there is much enthusiasm about the possibilities of computational neuroscience and there is genuine hope that the current multidisciplinary assault on one of the last scientific frontiers, the working of the brain, will reveal its secrets.

1.3 Hippocampal Place Cell (HPC) Phenomenon

It has been demonstrated that certain hippocampal neurons called "place cells" of a freely moving rat respond preferentially to a location in its already familiar environment^[2]. These cells act as markers of the rat's location in its environment. This phenomenon is similar to obtaining position fixes in navigation problems. To the uninitiated, the HPC phenomenon can be best explained by an analogy from automotive engineering. Suppose your automobile is equipped with a navigational aid which contains a small dash-mounted video screen displaying a city map and your automobile position is indicated by a flashing spot. The automobile is analogous to the rat, the video screen to the hippocampus and the flashing pixel to the place cell (with the flashing corresponding to the cell "firing"). As we will see below, this analogy does not carry through for representational and other details but is quite suggestive of the phenomenon.

O'Keefe and his collaborators^[29, 39, 30, 31] proposed the theory of "hippocampal place cells". O'Keefe defined the place cell as a cell whose firing rate or pattern of activity varies as a function of the animal's location in an environment by connecting together several multisensory inputs, each of which can be perceived when the animal is in a particular location^[33]. Experiments such as the T-maze with distal visual cues in which a rat was allowed to roam^[30] revealed that place cells in the CA1 area fired with increasing frequency when the rat was at specific locations^[39]. Variations of the shape of the maze, rewards and visual cues also produced similar increases in the firing frequency of complex spike cells^[32]. Further examples of such studies, without being exhaustive, are the radial maze experiments of McNaughton et al.^[21] in which they found additional correlations to direction and those of Breese et al.^[4] where they found that the significance of a location has an important bearing on place cell firing along with the work of Muller and Kubie^[26] where they speculated about the existence of a relationship to the kinematic aspects of animal motion. Best^[2,3] summarizes previous work to conclude that spatial cognitive processing occurs in the hippocampus and that the hippocampal place cells provide a sufficient substrate to serve as a spatial cognitive map.

To arrive at the essence of the place cell phenomenon, experiments where information from

experimental enclosure, proximal cues, rewards and such were designed to be negligible were performed by Muller et al.^[27]. Using careful experimental and data analysis methods, they showed that place cell firing is location-specific rather than behavior-specific^[26]. As suggested by Nadel^[28], a hierarchy of more and more complex information can be built up onto location-specific information so as to include the effects of rewards at certain locations, motion between locations^[26,42], temporal discontinuity between locations^[8] and other such correlates. It is to be noted that there are competing theories of hippocampal place cell function such as working memory^[37], temporary memory^[40], memory indexing^[50] and declarative memory^[48], among others. Critically analyzing the range of articles commented on by Nadel^[28], it appears that **the weight of evidence is on the side of the theory which considers the complex spike cells in the hippocampus acting as an intermediate-term storage for the representation of location-related information.**

Various studies have shown that the location-related information that is represented in the hippocampus is derived from distal cues in the environment. Considering the dependence on cues for information, Muller and Kubie^[26] have shown that in most cases, (a) rotation of cues causes an equal rotation of the place field, (b) cue size changes leaves the place field relatively unaffected, (c) increase of enclosure size increases the size of the place field and the appearance of a new place field and (d) change of enclosure shape causes an unpredictable shift of the place field, indicating that the information stored in the place cell is derived from distal cues.

In considering the various representational structures that are possible, the following experimental observations are significant. It has been reported that the notion of neighbourhood is not observed in the place cell field and the environment, i.e., two nearby place cells do not have nearby place fields in the experimental space^[34] or that the representation is "non-topographic". This is hard to explain if a tacit assumption is made that the flat experimental surface will be mapped onto a flat sheet of complex spike cells that is curved into the shape of a "C" in the hippocampus. The geometric property^[38] that is not considered in this type of analysis is that when the flat experimental surface is represented by feature-based cues, its representation in the hippocampus is no longer flat and, depending on the non-linearity of the mapping, the neighborhoods are no longer conserved. Another observation that has a significant bearing on the structure of the representation in the hippocampus is that the same place cell often represents more than one place field^[12]. The proposed representation must have the capability to account for such multiplicity without compromising accuracy of representation. In all of place cell studies, the number of primary features or functions derived from them is small, less than 3 or 4. The difficulty with representing any more features in a 3-dimensional physical brain (or 4-dimensional, considering time as the 4th dimension^[49]) is obvious. Nevertheless, this is an artificial constraint and many

more than 3 or 4 feature functions can be relevant in a redundant and hence fault-tolerant representation of the environment. The "neighborhood" problem, the "multiplicity" problem and the desirability of representing more than three features prompt our development of a non-orthogonal representation system. If the "Cartesian" system where axes are orthogonal is discarded, we can consider systems where the axes are concentric circles or radial lines or parallel lines, to name a few possibilities. The lamellar architecture of the hippocampus suggests a parallel co-ordinate system of representation.

The next important step in modeling the place cell phenomenon is the identification of cue-derived features. From the place cell experimental data available so far, the determination of the precise information that is derived from the cues has been difficult. It has been shown that one of the simplest features, the radial distance to each cue, is probably not the cue-derived feature that is stored^[27]. Speculations about features range from primary features such as angles between cues^[46] and the retinal area of the cue image^[53] to functions of primary features such as distance to and angle between a suitably defined allocentric co-ordinate system based on the cues^[36] and "aggregate predictions"^[42]. Turning our attention to the development of representation during the exploratory phase, there is a large body of work which implicates the mechanisms of Long Term Potentiation (LTP) or Long Term Enhancement (LTE) in the increase in firing rates of the hippocampal pyramidal cells^[1,23]. It has been suggested that the LTP or LTE mechanism will permit intermediate-term (several weeks) activation of cells that have highly convergent inputs^[22]. This is an adequate mechanism for the development of representation in the hippocampus during exploration.

From a computational/theoretical point of view, several proposals have been put forward to explain the place cell phenomenon. Zipser^[53] considered it to be a pattern recognition problem and developed a 2-stage neural network to model the system. Some of the drawbacks of this model which stem mainly from the inability to properly explain the cue removal effects are discussed by Muller and Kubie^[26]. A neural network approach^[43] has been recently published to model classical conditioning effects using Grossberg's theory^[7]. McNaughton^[24] proposed another neural network for the place cell phenomenon using a matrix-associator model. This associator codes the local view and a transition matrix corresponding to the movement to be made is computed. O'Keefe^[35,36] proposes a computational theory where certain parameters based on the cues are calculated to obtain an allocentric co-ordinate system and features based on the allocentric system are derived. Recently, a neural network simulation which reproduced the place cell activity remarkably well has been reported^[47]. The feature common to all these models is that they provide a "global" model for the hippocampal activity. As a first step, such an approach is appropriate except that specific feature transformation properties,

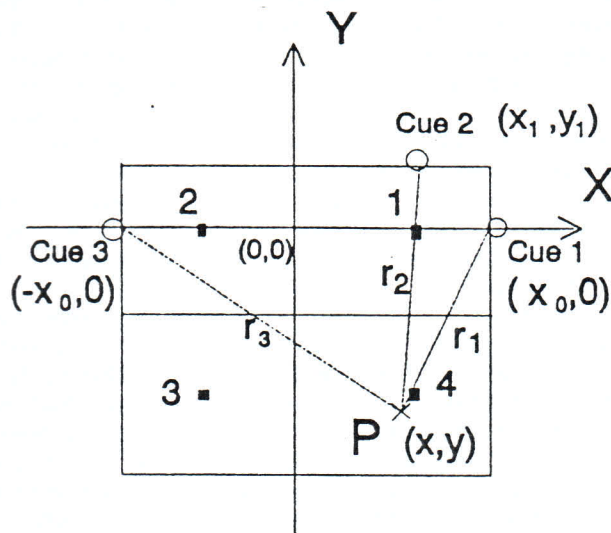


Figure 1 : Experimental surface with cue locations and grid centers marked. $x_0 = 1$; $x_1 = y_1 = 0.5$.

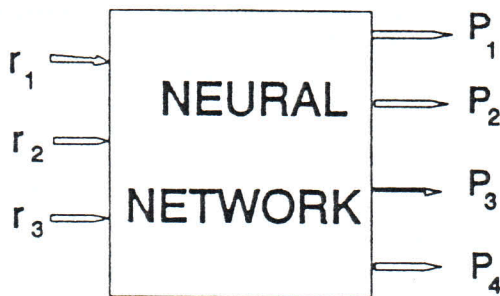


Figure 2 : Neural network with cue distance as inputs and place cells as outputs.

representational details and directly testable hypotheses are not considered at this global modeling level.

To summarize, we view the complex spike cells in the hippocampus as an intermediate-term storage for the representation of location-related information. In the sequelae, we present both a global neural network model and a neurobiologically-plausible computational model.

1.4 HPC Phenomenon - Artificial Neural Network Model

Artificial neural network (ANN) models are important tools in the development of computational models. Simulation of input-output relationships of large groups of neurons using ANNs aid us in understanding the global functions and in some cases, display correspondences to the underlying neural structures^[54]. In this spirit, we attempt to simulate the HPC phenomenon using ANNs.

In our simulation of the place cell phenomenon, we consider the mapping of locations in the experimental environment (where the rat is free to roam) into the firing of individual neurons. Our output feature is not firing frequency but simply the firing state (or ON/OFF state) of the neuron. Nevertheless, this establishes the connection between a physical location and a neuron, which is the hallmark of the place cell phenomenon. Such a mapping can be accomplished with a multilayer perceptron artificial neural network.

The multilayer perceptron using the backpropagation learning rule is so widely known that no details of the method need be given^[15]. Our perceptrons have sigmoidal non-linearities and the thresholds were kept fixed. In addition to the input and output layers of processing elements, our network has one hidden layer of processing elements. In our simulation, the rat's location is represented by any point on the 2-dimensional space within the boundary (called "experimental surface") shown in figure 1. The x-y co-ordinate system to identify any point is shown as well as the cues on the periphery and their x-y co-ordinates as Cue i (x_i, y_i). Since there are infinite number of points (locations) within the experimental surface and only finite number of place cells, we discretize the surface and make the number of locations that the rat can visit finite. To keep the neural network learning process tractable, we take the rather extreme case where there are only four locations that the rat can occupy on the surface (shown as 1, 2, 3 and 4 in figure 1). These are the centers of the four grids that the experimental surface was discretized into. In our simulations, there will be only four place cells which corresponds to grid centers, 1 to 4.

As we have seen earlier, there is a significant amount of evidence that the animal uses information derived from the cues to locate itself in space. We have decided, for this preliminary study, to use the most obvious feature based on cues, the distances of the rat to the 3 cues. Therefore, any point, P can be defined as a vector with 3 elements in the cue-co-ordinate system, i.e., $P = (r_1, r_2, r_3)$, where r_i , the distance to Cue i , is a real, non-negative number. As we have mentioned earlier, there are four place cells corresponding to the four grid centers. Ideally, these neurons fire only when the rat is at one of the four grid centers. The firing is indicated by the neuron outputs changing states from a "0" to a "1" and therefore, they are binary quantities. The place cell neurons, P_i 's, can be considered as "labeled lines" indicating grid occupancy. In other words, when the rat is at the center of grid 3, only the place cell 3 will be in the "1" state while the others will be in the "0" state (i.e., $P_3 = 1$; $P_1 = P_2 = P_4 = 0$).

The multilayer perceptron we used has 3 inputs corresponding to r_1 , r_2 and r_3 and four outputs corresponding to place cells, P_1 , P_2 , P_3 and P_4 . This is shown in figure 2. Our neural network has one hidden layer with seven nodes. A certain amount of experimentation was necessary to find these parameters. During training, the permissible locations are only the grid centers and therefore the training set contained four

DISTANCE MSE DISTRIBUTION

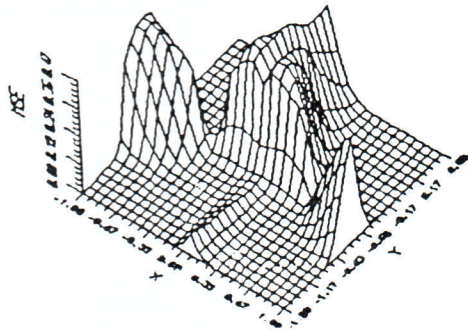


Figure 3 : Error surface during testing. Cues 1 to 3 are at (1.0), (0.5, 0.5) and (-1.0) respectively.

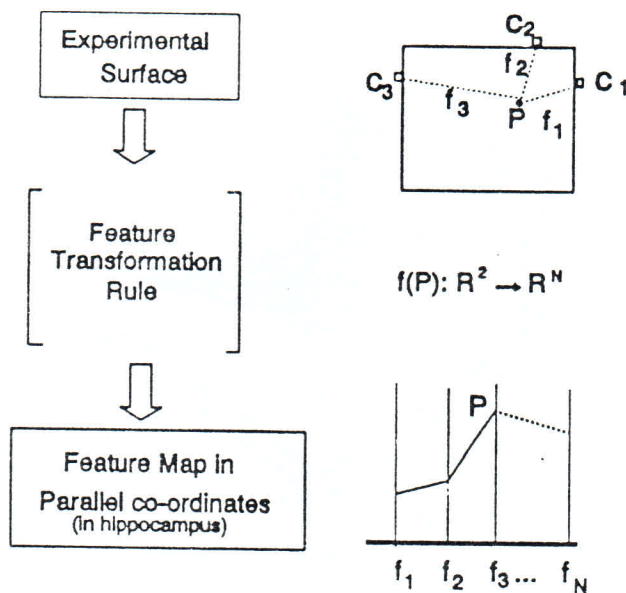


Figure 4 : Transformations in the place cell system

input-output pairs. This training set was repeatedly presented to the neural network and the mean absolute error between the observed and desired P_i 's monitored. The training was stopped when this value was less than 10^{-2} . This happened, on average, after approximately 24,000 iterations of the training set.

The trained neural network was tested by placing the test rat at grid centers initially. As can be expected from the low mean absolute error of 10^{-2} , the agreement was quite good. For example, if the test rat is at the grid 3 center, the output vector, (P_1, P_2, P_3, P_4) should be equal to (0, 0, 1, 0). We obtained (0.0014, 0.01, 0.99, 0.0099), which shows good agreement with the desired vector. To study the ability of the network to generalize from the specific examples used for training, we tested the grid identification ability of our trained network with unseen positions. For this test, the rat was allowed to

occupy any of 21×21 equally spaced-points on the same x-y plane as shown in figure 1. Among these 441 points, only the 4 grid centers had been seen before. The property of the simulation that is being tested would be the ability to sample the space and then generalize that information to all points in space. For example, if the test rat is at location (-0.8, -0.9) in x-y co-ordinates, i.e., near the boundary of grid 2 and 3 in grid 3, we expect the output vector to be (0, 0, 1, 0). When tested, we obtained (0.001, 0.33, 0.69, 0.0086). The error can be computed as the mean absolute error between the desired and actual output vectors. This large error reflects the uncertainty or difficulty in deciding to which grid the point (-0.8, -0.9) belongs. The mean absolute error for all 21×21 points on the plane is plotted in figure 3. Notice that around grid centers, the errors are very low, whereas near grid boundaries, the errors are quite high, as was seen in the numerical example above. In figure 3, the grid closest to cues C1 and C2 is grid 1 (on the far right side). Moving counter-clockwise, we have grids 2, 3 and 4. The following observations can be made from this figure: (i) the error is very low at not only the grid centers, but also for fairly large regions around them, thus showing the ability to generalize; (ii) the error is large at the center of the experimental surface at x-y co-ordinate position (0, -0.5). The ambiguity as to which of the four grids the point belongs to is a maximum here; (iii) the error near internal boundaries between any two grids is quite large. The boundary that runs parallel to the x-axis through the center of the experimental surface has an undulating profile, the reasons for which are not quite clear at this point; and (iv) the error profile along the y-axis has a distinct pattern, viz., the error is high between grids 1 and 2 whereas it is much lower between grids 3 and 4. Consider the boundary between grids 1 and 2. Relative to cues C1 and C3 (disregarding C2 for the moment), the points at the boundary between grids 1 and 2 are symmetric along this line and hence ambiguity exists between points in the upper and lower halves. Therefore, the error profile is higher. Near the boundary between grids 3 and 4, this ambiguity does not exist because symmetric points are outside the experimental surface.

The feasibility of simulating the place cell phenomenon has been demonstrated by this study. Even with a rather arbitrary choice of the cue feature (distance to the cue), the multilayer perceptron network is able to abstract the essential features of the phenomenon and show interesting abilities to generalize. This model allows us to duplicate various physiological experiments such as cue removal and study its effect on the error surface. We have performed simulations [11, 16] where cue features such as angles between the cues and features based on an allocentric co-ordinate system [36] are employed.

2. HPC PHENOMENON - COMPUTATIONAL MODEL

From a computational neuroscience point of view, the previous model lacks the wealth of internal details

necessary to form testable hypotheses that lead to physiological experiments. In this section, we develop a model consistent with known hippocampal anatomy and physiology based on which certain experiments to test our model are proposed. From the place cell experimental data available so far, the determination of the precise information that is derived from the cues has been difficult. However, the determination of features derived from the cues and their representation in the hippocampus are intricately interconnected. We advance the hypothesis that location information is represented in the hippocampus based on a parallel co-ordinate system.

We considered the place cell representation problem to be basically **geometric** in nature, where points on the plane on which the animal roams around are mapped into a sheet of cells in the hippocampus, considering the 2-dimensional case for the present [17]. The overall transformation in the animal brain has **two steps**. The first step involves obtaining a cue-based representation, called "**feature space**", of the experimental surface. The second step is the mapping of the "feature space" into the 2-dimensional place cell sheet using a parallel co-ordinate representation system.

The theory of place cell computational system proposed is represented in the flow chart in figure 4. The left hand column represents the theoretical steps and examples corresponding to each step are shown on the right. The overall problem is cast as a geometric problem of mapping from the 2-dimensional space to the N-dimensional space, i.e., R^2 to R^N . Appropriate features from cues are extracted and used in representing the surface. The choice of features defines the transformation rule used. The cornerstone of the place cell system is the representation scheme. We propose the parallel co-ordinate system. Its ability to explain observed physiological data and many desirable properties such as avoidance of dimensionality problems and graceful degradation of stored information will be discussed below. The right hand column of the flowchart shows examples of the blocks in the left hand column. The animal is at point P. The features, f_i are any appropriate features derived from the cues (distances to the cues, for example). Mathematically, $f(P): R^2 \rightarrow R^N$ corresponds to the feature transformation rule. In the parallel co-ordinate system, the line segments represent the point, P.

3. FEATURE TRANSFORMATION

In our analysis, the animal can be at any point on the 2-dimensional space within the boundary (called "experimental surface") shown in figure 5. The x-y co-ordinate system to identify any point is shown as well as the cues on the periphery and their x-y co-ordinates as Cue 1: (x_1, y_1) . There is significant amount of evidence that the animal uses information derived from the cues to locate itself in space. What type of information is derived is a matter of debate and there are various speculations in the literature [53]. We have decided, for this preliminary study, to use 2 feature sets; distance and angle features.

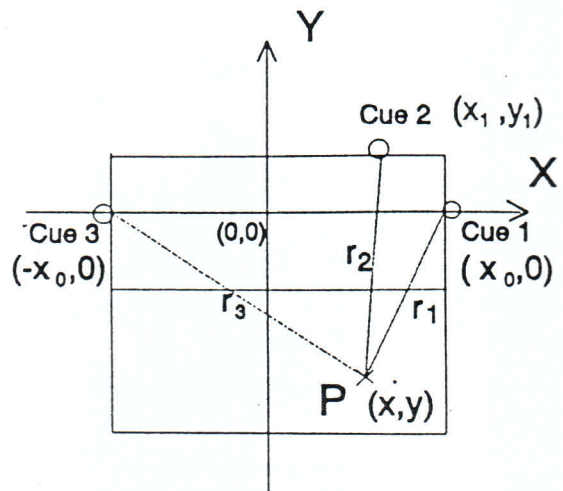


Figure 5 : Experimental surface; distance features.
 $x_0 = 1; x_1 = y_1 = 0.5$

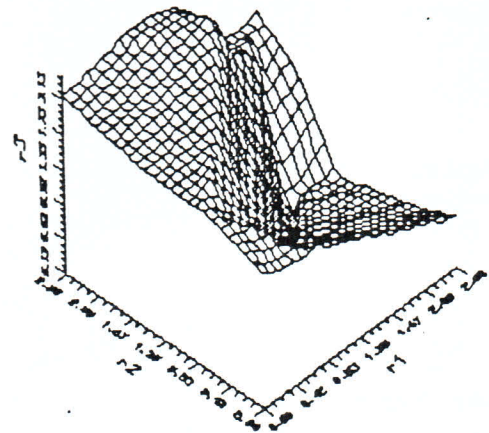


Figure 6 : Features space for distance features.

(i) *Distance Features*: The distances-to-the-cues features are one of the most direct and simple representations that can be studied. Even though distances may not be the features actually used by the animal, analysis using distances is very instructive because of its simplicity. Any point, P can be defined as a vector with three elements in the cue-co-ordinate system, i.e., $P = (r_1, r_2, r_3)$, where r_i , the distance to Cue i , is a real, non-negative number.

Considering figure 5, assume, without loss of generality, that the x and y axes are so located that Cue 1 is at $(x_0, 0)$ and Cue 3 is at $(-x_0, 0)$. Cue 2 is at (x_1, y_1) . Using the distances to the three cues from any point, $P(x, y)$, we can write $P(x, y) = P(r_1, r_2, r_3)$

$$\begin{aligned} \text{where } r_1 &= \pm \sqrt{x_0^2 + x^2 + y^2 - 2x_0x} \\ r_2 &= \pm \sqrt{(x - x_1)^2 + (y - y_1)^2} \\ r_3 &= \pm \sqrt{x_0^2 + x^2 + y^2 - 2x_0x} \end{aligned}$$

In figure 6, we see the experimental surface mapped into the feature space. The surface is curved. The consequence of the curvature is that the notion of "neighborhood" becomes different in the experimental surface and feature space, i.e., two points next to each other in the feature space may correspond to two points far apart on the experimental surface. In this example where the surface in the feature space is highly warped, such non-linearity is very large. Two points on either side of a deep fold in the feature space will be much farther apart on the experimental surface. Different cue configurations will produce different amounts of warping of the surface in the feature space.

(ii) Angle Features : Using other information from the cues, we can attempt to see which is most suggestive of the observed physiological non-linearities. A second attempt was made using angles between cues as the features as shown in figure 7, where

$$a_1 = \cos^{-1} \frac{r_1^2 + r_2^2 + r_{12}^2}{2r_1r_2}$$

$$a_2 = \cos^{-1} \left[\frac{r_2^2 + r_3^2 + r_{23}^2}{2r_2r_3} \right]$$

$$a_3 = 2\pi - (a_1 + a_2)$$

are the angle features. Here, r_1 , r_2 and r_3 are distances to the cues as in figure 5 and r_{12} and r_{23} are the distances between cues 1-2 and 2-3, respectively.

The feature space corresponding to the experimental surface is shown in figure 8. As can be seen, the feature space is not very warped and the correspondence of neighborhoods will not be very non-linear in this case. Comparison of feature surfaces for distances and angles in figures 6 and 8, respectively, reveals their strikingly different properties. In figure 6,

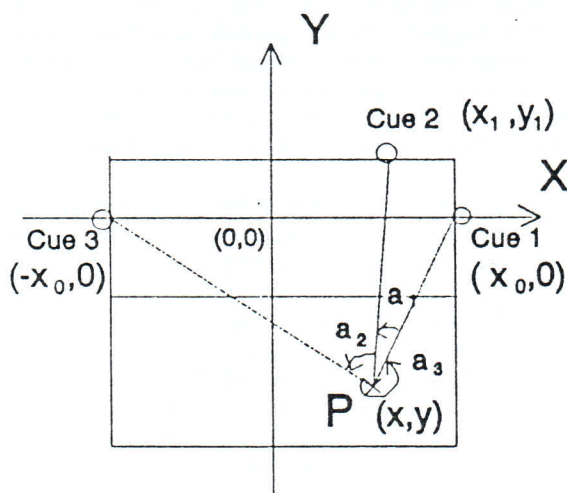


Figure 7. : Experimental surface; angle features. $x_0 = 1$;
 $x_0 = y_1 = 0.5$.

due to significant non-linearity of the surface, two points close by will be far apart on the experimental surface.

It is to be noted that **the choice of the features actually used by the animal is impossible to determine** with this analysis. What it allows us to do is to study all the properties of the corresponding transformations and predict place cell recording events which can then be tested in actual animal experiments. The transformation and the corresponding feature set that best predicted the recorded place cell firing for locations on the experimental surface should be selected as the one used by the animal.

4. PARALLEL CO-ORDINATE REPRESENTATION

The experimental surface that the animal is exposed to can contain more than 2 or 3 cues. From each of the large number of cues, more than one feature can be extracted. Representation of these large number of features (say, equal to N) becomes exceedingly difficult if we insist that each feature be represented along orthogonal axes. Not only is this N -dimensional space hard to visualize, but also the more pertinent question of how they may be represented in a 3-dimensional (or 4 using time as an axis) brain becomes impossible to answer. Indeed, it could be argued that current theories [36,53] consider only two or three features because of this "demon of dimensionality".

The use of a parallel co-ordinate system [9] is best explained by the example in figure 9. Suppose that the animal is on a rectangular experimental surface with three cues as shown in figure 9(a). The numerical values shown are examples of lengths. The point, P can be represented either in x - y co-ordinates, $P(x,y) = P(1,3)$, or by features assumed here to be distances to the cues, $P(r_1, r_2, r_3) = P(3,2,5)$. In figure 9(b) and (d), we show the orthogonal co-ordinate representation and in (c) and (e), the parallel co-ordinate representation. Another categorization is that figure 9(b) and (c) show the **cartesian space** where any point is represented by its

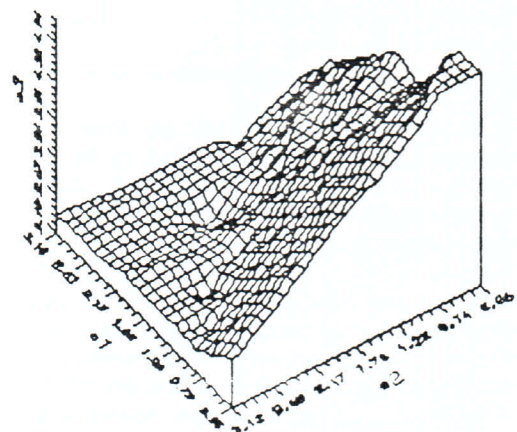


Figure 8 : Feature space for angle features

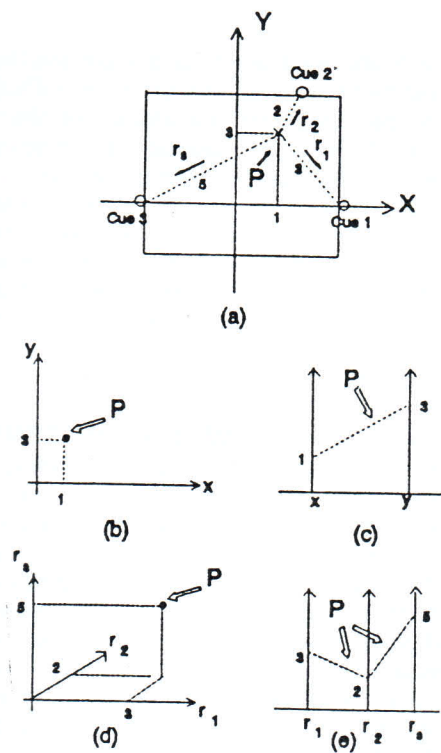


Figure 9 : Experimental surface and various representations of a point, P

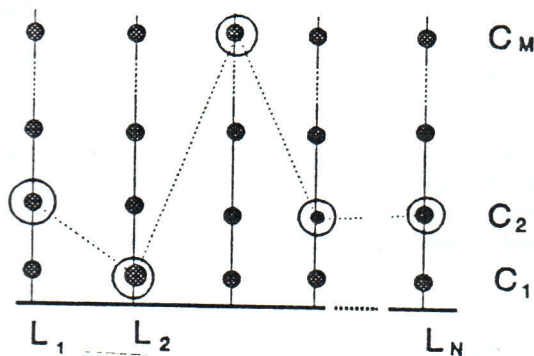


Figure 10 : Parallel co-ordinate representation of features

(x,y) co-ordinates whereas (d) and (e) show the **feature space** where any point is represented by its features, in this case distances to the three cues. **All the four representations in figures 9(b), (c), (d) and (e) are completely equivalent.**

We propose that any large number, N , of features extracted from the cues in the experimental surface can be represented in a parallel co-ordinate system. In parallel co-ordinates, we will have N vertical axes and the point, P , will be a sequence of line segments or what we call a "trace", as shown in the figures 4 and 9(e). Such a sequence of line segments is very reminiscent of an "engram". What we propose is that, **in the hippocampus, such traces code the location**

information in terms of features derived from the cues.

When we have N features based on the cues to represent a point on the experimental surface, we adopt the parallel co-ordinate representation as shown in figure 10, which is a generalization of figure 9(e). The magnitude of each feature is represented by M cells, C_1 to C_M , arranged along axes, L_1 to L_N in such a way that for a larger magnitude of the feature, a cell higher up will be active. The foci of activity, or the active cells, are indicated by circles around cells. These cells may or may not be interconnected as shown by the dotted lines. Based on such a representation, we identify the following desirable features: (i) "Demon of Dimensionality": we are no longer constrained to represent locations with two or three features. As is most often the case, there can be a large number of features available in the experimental region and all N of them can be represented on the N parallel axes, L_1 to L_N ; (ii) Storage Capacity:- we consider a small finite number of cells, equal to $N \cdot M$. In the parallel co-ordinate representation, each piece of information is stored as a sequence, or **trace of active cells**. The same cells could participate in many traces. Hence, the total amount of storage of information is equal to M^N . If $N = 10$ and $M = 50$, the total storage = $50^{10} \approx 10^{17}$, a very large number. If pieces of information were generated every microsecond, such a system can store all the information generated in a 100-year life span (a total of about 10^{16} pieces of information); and (iii) "Trace" Storage:- the use of many features represented on many axes provides redundancy in the representation of a location using N features. The loss of a few cells will retain that representation, giving rise to graceful degradation of specificity rather than catastrophic loss. If the active cells are interconnected, activation of a few cells will activate the complete trace.

We propose that the anatomical basis of the parallel co-ordinate representation resides in the **lamellar** structures of the hippocampus. The tightly coupled, transverse organization of interconnections in each lamella is suggestive of the parallel co-ordinate axes [49]. This organization is reminiscent of "functional columns" in the neocortex. Since feature function magnitudes are coded along the columns here, we will call them "feature columns". The position of CA3 pyramidal cells codes the magnitude of the feature function.

Our parallel co-ordinate scheme seems to satisfy most of the reported place cell data regarding representation in the hippocampus: (i) The primary experimental observation is that place cells respond to location [26] based on cue-derived information. Clearly, our hypothesis that active cells code feature functions is consistent with this observation. (ii) It is seen that same cell may have fields in more than one environment [12]. Referring to figure 10, it can be seen that an active cell may be part of more than one trace and hence may code a location in more than one environment. (iii) There is a certain discontinuity in the mapping in that nearby cells do not correspond to nearby place fields. This is consistent with our model in that two nearby

active cells may be part of two different traces. For example, in figure 10, consider axes, L_1 and L_2 . The two active cells shown correspond to a point, P_1 close to Cue2 and slightly away from Cue1, if distances to the cues are assumed to be the features. Now consider a new trace where the lowest cell on L_1 and the highest cell on L_2 are active. This point, P_2 is close to Cue1 but far from Cue2. Hence, points, P_1 and P_2 are far apart on the experimental surface, whereas active cells on L_1 are close to each other. (iv) Removal of some redundant cues do not affect the place cell activity, as shown by removal of two out of four cues [32]. Again, considering the trace shown in figure 9, if we are monitoring the active cell on L_1 , even if some other cues are removed, the cell on L_1 will be active if the corresponding cue is not removed.

To summarize, we take a unique geometric point of view in analyzing the place cell experimental observations. This leads us to a parallel co-ordinate representation system within the hippocampus. The "neighborhood" problem and representation using a large number of features which none of the current theories can account for are accommodated naturally within our parallel co-ordinate representation system. From a theoretical neuroscience point of view, the notion of a non-orthogonal representation system will permit a new approach to the study of neural representation where the architectonics will determine the type of the axes of the representation system, the parallel axes being just one example of a range of possibilities.

5. EXPERIMENTS TO TEST OUR MODEL

From a computational neuroscience point of view, for our models to be useful, we have to test whether indeed the computational steps proposed are carried out in the animal brain. Our main assumption is that the lamellae correspond to the parallel axes of the representation system. Testing this assumption is confounded by the lack of exact knowledge of the cue-derived features used in the representation. A very decisive, albeit difficult experiment can be designed to test the hypothesis that location information is represented based on a parallel co-ordinate system in the hippocampus. A second experiment is also given to identify the feature functions used. Purely for the purpose of exposition, we will assume for the time being, that primary features such as radial distances to the cues are used by the animal.

1) *Parallel co-ordinate representation:-* In the first stage of the experiment, three rectangular cues (30° sector angle) will be placed along the experimental chamber wall at 3 o'clock, 10 o'clock and 7 o'clock positions. A linear array of electrodes [14], spaced 1 mm apart will be used in these experiments. The electrode will be stereotactically implanted, longitudinally in the hippocampus. Alternatively, three separate tungsten microelectrodes can be implanted independently using stereotactic procedures. The locations are such that the three electrodes record from three adjacent lamellae in the CA3 area. During training of the rat, the electrode position can be adjusted so that at least two place cells

being recorded from will have the same place field. From figure 10, it can be seen that this corresponds to monitoring one trace.

When stable recordings are obtained, we add a 4th rectangular cue at the 1 o'clock position. This is the 2nd stage of the experiment. Based on our assumption that the radial distances to the cues are the features used by the animal, the introduction of the 4th cue should not disturb the representation of the first two cues along two adjacent lamellae in the 1st stage. Following retraining, recordings under test conditions will be performed again. If our hypothesis is tenable, the firing of the first two place cells should remain unchanged.

2) *Feature function selection:-* This is an extremely important aspect since if the feature functions used by the animal become known, the accuracy of the proposed representation can be tested conclusively.

The experimental arrangement will be similar to the earlier one using three cues. In addition, a second linear array electrode such that the two electrodes are parallel will be used (figure 11(a)). From an introductory consideration of parallel co-ordinates, it can be seen that two parallel links between two axes correspond to two points on a line of slope equal to +1 in orthogonal co-ordinates. If the two parallel electrodes are not perpendicular to the lamellar axes (horizontal in figure 11(a)), it means that the line does not pass through the origin but the line will still have a slope equal to +1.

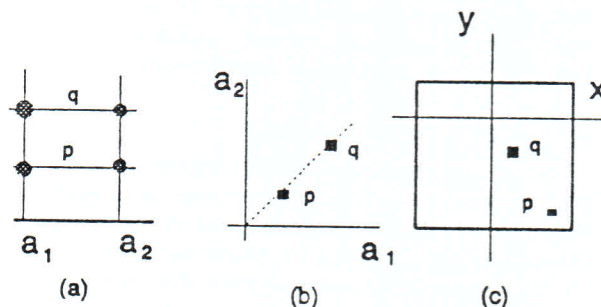


Figure 11 : Two locations, p and q, in
a) parallel co-ordinate representation, b) feature space and
c) on the experimental surface

In figure 11(a), we show the two electrodes (horizontal lines) measuring from two lamellae, marked a_1 and a_2 . If we assume that the features used are the two angles between the three cues (as in figure 7), figure 11(b) shows the feature space for the two-angle case. This is similar to figure 8 which was for three angles. Because of the warping of the feature space, the two place fields on the experimental surface corresponding to the two place cells monitored may be found centered at p and q as shown in figure 11(c). The need to analyze the feature space is now clear because in its absence we would not be able to predict the positions of the place fields on the experimental space correctly.

During the experiments, when stable firings from all 4 recording sites are obtained, we will test one feature space at a time from the array of solutions we have

previously developed. It is to be noted that for the same place cells monitored, place fields will be different for different feature functions. The procedure to test which features are being used by the animal is as follows:

(i) Mark the centers of place fields on the experimental surface (as in figure 11(c)) corresponding to active traces p and q.

(ii) Consider one type of feature function at a time, say angle features. Find the angles between cues at points, p and q on the experimental surface (from figure 11(c)).

(iii) On the feature space map for angles previously developed, mark the two points corresponding to the two angles obtained in step (ii) above (as in figure 11(b)).

(iv) If angle features are indeed being used by the animal, the two points on the feature space map must lie on a straight line with slope equal to +1 as shown in figure 11(b). This is so because points on a line with slope equal to +1 in orthogonal co-ordinates correspond to two parallel line segments in parallel co-ordinates (as in figure 11(a)). In general, the line of slope equal to +1 on which the two points, p and q, lie in the feature space need not pass through the origin, as already mentioned.

If the line connecting the two points in figure 11(b) does not have a slope equal to +1, the cue-derived features used by the animal are not the angles between the cues. The procedure described above can then be iterated for each feature function for which the feature space has been determined, including new feature functions or feature function combinations.

6. DISCUSSION

Our attempts to model the hippocampal place cell phenomenon provide an opportunity to investigate the various tools and procedures used in the computational neuroscience field. Artificial neural networks are perhaps the best known tool emerging from the research activities in computational neuroscience and related fields. However, its utility in brain modeling may be limited. As we saw in the case of HPC modeling, ANNs can simulate the phenomenon in its essential details but inspection of hidden layers and nodes or connection weights is not likely to lead directly to testable hypotheses. In our computational model of the HPC phenomenon on the other hand, we took an approach more consistent with the tenets of computational neuroscience. We discussed the overall transformation and representation issues at a computational theory level. Our choices at this level were constrained by the implementation (hardware) details known about the hippocampal anatomy. Our model does not address the algorithmic level in detail at this stage of its development. In any case, the success of this model as being neurobiologically useful can be declared only by conducting the proposed (or similar) animal experiments.

As we had mentioned during the brief introduction to computational neuroscience, even when the models do not hold up in physiological experiments, they may have

significant engineering value. In this case, the HPC model of spatial cognitive processing may be used as a model for machine algorithms. This may provide new and robust methods of map building and navigation for autonomous mobile vehicles when position information is not externally supplied. In the area of map building and navigation of mobile vehicles, Kuipers and Levitt^[13] address many of the issues that parallel the HPC phenomenon. The point has been made^[13] that in order to learn the structure of large-scale space, the observer has to build a cognitive map by interacting with the environment. This involves sensing cues in the environment and observing the effects of moving in this space. As is evident, this is similar to the animal developing the "place cell map" of the environment by moving and observing the environment. Once the observer has a cognitive map of the environment, navigation is possible by creating and executing a plan to travel from one place to another. The traditional methods^[6, 10] do not build maps from local observations but instead rely on non-local information such as preexisting maps, active landmark beacons or geolocating satellites. Navigation is performed by inertial or dead-reckoning schemes^[6] which can accumulate substantial position errors. In general, accurate measurements of sensor data and movement parameters are essential to the success of the traditional methods of map building and navigation. Methods based on a four-level semantic hierarchy^[13] and occupancy grid^[6] have been proposed for robust spatial information processing. In contrast to these methods, we could base a system on the hippocampal place cell phenomenon to develop a robust map building and navigation method^[18, 52] for mobile vehicles.

REFERENCES

- [1] Berger, T.W., "Long-term potentiation of hippocampal synaptic transmission affects rate of behavioral learning", *Science*, vol. 224, pp 627-630, 1984.
- [2] Best, P.J., "Hippocampal cellular activity and spatial cognitive processing", *Psychobiology*, vol. 17, pp 215-216, 1989.
- [3] Best, P.J. and Thompson, L.T., "Persistence, reticence, and opportunism of place-field activity in hippocampal neurons", *Psychobiology*, vol. 17, pp 236-246, 1989.
- [4] Breese, C.R., Hampson, R.E. and Deadwyler, S.A., "Hippocampal place cells: stereotypy and plasticity", *Journal of Neuroscience*, vol. 9, pp 1097-1111, 1989.
- [5] Daugman, J.G., "Brain Metaphor and Brain Theory", in *Computational Neuroscience*, E.L. Schwartz (ed.), MIT Press, Cambridge, 1990, pp 9-18.
- [6] Elfes, A., *Occupancy grids: a probabilistic framework for robot perception and navigation*, Ph.D. Thesis, Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, 1989.
- [7] Grossberg, S., "A neural model of attention, reinforcement, and discrimination learning", *International Review of Neurobiology*, vol. 18, pp 263-327, 1975.
- [8] Hughey, D.J., "Temporal discontinuity: alternative to, or component of, existing theories of hippocampal function?", *Behavioral and Brain Sciences*, vol. 8, pp 501-502, 1985.
- [9] Inselberg, A., "The plane with parallel coordinates", *The Visual Computer*, vol. 1, pp 69-91, 1985.
- [10] Kayton, M., "Introduction to land navigation", in *Navigation: Land, Sea, Air and Space*, M. Kayton (ed.), IEEE Press, New York, 1990, pp 65-71.
- [11] Khatri, N. and Madhavan, P.G., "ANN Simulation of the Place Cell Phenomenon Using Cue Size Ratio", *4th Conference on Neural Networks and Parallel Distributed Processing*, Fort Wayne, 1992.
- [12] Kubie, J.L. and Ranck, J.B., "Sensory-behavioral correlates in individual hippocampal neurons in three situations: space and context",

- in *Neurobiology of the Hippocampus*, W. Seifert (ed.), Academic Press, London, 1983, pp 433-447.
- [13] Kuipers, B.J. and Levitt, T.S., "Navigation and mapping in large-scale space", in *Advances in Spatial Reasoning*, S. Chen (ed.), Ablex Publishing Corporation, New Jersey, 1990, pp 207-251.
 - [14] Kuperstein, M. and Whittington, D.A., "A practical 24 channel microelectrode for neural recording in vivo", *IEEE Transactions on Biomedical Engineering*, vol. 28, pp 288-293, 1981.
 - [15] Lippmann, R.P., "An introduction to computing with neural nets", *IEEE ASSP Magazine*, pp 4-22, 1987.
 - [16] Madhavan, P.G., Xu, B. and Low, W., "Simulation of Hippocampal Place Cell Phenomenon Using Multilayer Perceptrons", *Proceedings of the International Conference on System, Man and Cybernetics*, Charlottesville, 1991, pp 1611-1614.
 - [17] Madhavan, P.G., Xu, B., Penna, M. and Low, W., "Co-ordinate Transformations in the Hippocampal Place Cell Phenomenon", *Proceedings of the International Conference on System, Man and Cybernetics*, Charlottesville, 1991, pp 1615-1620.
 - [18] Madhavan, P.G. and Manatunga, D., "Map Building & Hippocampal Place Cell Phenomenon", *Computational Neuroscience Symposium*, Indianapolis, October 1992 (to be published).
 - [19] Marr, D., *Vision*, W.H. Freeman and Company, New York, 1982.
 - [20] McCulloch, W.S. and Pitts, W., "A logical calculus of the ideas immanent in nervous activity", *Bulletin of Mathematical Biophysics*, vol. 5, pp 115-133, 1943.
 - [21] McNaughton, B.L., Barnes, C.A. and O'Keefe, J., "The contributions of position, direction, and velocity to single unit activity in the hippocampus of freely-moving rats", *Experimental Brain Research*, vol. 52, pp 41-49, 1983.
 - [22] McNaughton, B.L. and Barnes, C.A., "The hippocampus, synaptic enhancement, and intermediate-term memory", *Behavioral and Brain Sciences*, vol. 8, pp 507-508, 1985.
 - [23] McNaughton, B.L., Barnes, C.A., Rao, G., Baldwin, J. and Rasmussen, M., "Long-term enhancement of hippocampal synaptic transmission and the acquisition of spatial information", *Journal of Neuroscience*, vol. 6, pp 563-571, 1986.
 - [24] McNaughton, B.L., "Neuronal mechanisms for spatial computation and information storage", in *Neural Connections, Mental Computation*, L. Nadel, L.A. Cooper, P. Culicover and R.M. Hamish (eds.), MIT Press, Cambridge, 1989, pp 285-350.
 - [25] Minsky, M.L. and Papert, S.A., *Perceptrons*, MIT Press, Cambridge, 1969.
 - [26] Muller, R.U. and Kubie, J.L., "The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells", *Journal of Neuroscience*, vol. 7, pp 1951-1968, 1987.
 - [27] Muller, R.U., Kubie, J.L. and Ranck, J.B., "Spatial firing patterns of hippocampal complex-spike cells in a fixed environment", *Journal of Neuroscience*, vol. 7, pp 1935-1950, 1987.
 - [28] Nadel, L., "The hippocampus and space revisited", *Hippocampus*, vol. 1, pp 221-229, 1991.
 - [29] O'Keefe, J. and Dostrovsky, J., "The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat", *Brain Research*, vol. 34, pp 171-175, 1971.
 - [30] O'Keefe, J., "Place units in the hippocampus of the freely moving rat", *Experimental Neurology*, vol. 51, pp 78-109, 1976.
 - [31] O'Keefe, J. and Nadel, L., *The Hippocampus as a Cognitive Map*, Clarendon Press, Oxford, 1978.
 - [32] O'Keefe, J. and Conway, D.H., "Hippocampal place units in the freely moving rat: why they fire where they fire", *Experimental Brain Research*, vol. 31, pp 573-590, 1978.
 - [33] O'Keefe, J., "A review of the hippocampal place cells", *Progress in Neurobiology*, vol. 13, pp 419-439, 1979.
 - [34] O'Keefe, J. and Speakman, A., "Single unit activity in the rat hippocampus during a spatial memory task", *Experimental Brain Research*, vol. 68, pp 1-27, 1987.
 - [35] O'Keefe, J., "Computations the hippocampus might perform", in *Neural Connections, Mental Computation*, L. Nadel, L.A. Cooper, P. Culicover and R.M. Hamish (eds.), MIT Press, Cambridge, 1989, pp 225-284.
 - [36] O'Keefe, J., "A computational theory of the hippocampal cognitive map", in *Progress in Brain Research, Vol. 83: Understanding the Brain Through the Hippocampus*, J. Storm-Mathisen, J. Zimmer and O.P. Ottersen (eds.), Elsevier Science, Amsterdam, 1990, pp 301-312.
 - [37] Olton, D.S., Becker, J.T. and Handelmann, G.E., "Hippocampus, space, and memory", *Behavioral and Brain Sciences*, vol. 2, pp 313-365, 1979.
 - [38] Pellionisz, A.J., "Neural geometry", *Proceedings of the International Joint Conference on Neural Networks*, Washington, 1989, pp 711-715.
 - [39] Ranck, J.B., "Studies on single neurons in dorsal hippocampal formation and septum in unrestrained rats", *Experimental Neurology*, vol. 41, pp 461-555, 1973.
 - [40] Rawlins, J.N.P., "Associations across time: the hippocampus as a temporary memory store", *Behavioral and Brain Sciences*, vol. 8, pp 479-496, 1985.
 - [41] Rosenblatt, F., *Principles of Neurodynamics*, Spartan Books, New York, 1962.
 - [42] Schmajuk, N.A., "Role of the hippocampus in temporal and spatial navigation: an adaptive neural network", *Behavioural Brain Research*, vol. 39, pp 205-229, 1990.
 - [43] Schmajuk, N.A. and DiCarlo, J.J., "A neural network approach to hippocampal function in classical conditioning", *Behavioral Neuroscience*, vol. 105, pp 82-110, 1991.
 - [44] Schwartz, E.L. (ed.), *Computational Neuroscience*, MIT Press, Cambridge, 1990.
 - [45] Sejnowski, T.J., Koch, C. and Churchland, S., "Computational neuroscience", *Science*, vol. 241, pp 1299-1306, 1988.
 - [46] Shapiro, M.L., Hetherington, P.A., Eichenbaum, H.B. and Fortin, W.J., "A simple PDP model simulates spatial correlates of hippocampal neuronal activity", *Society for Neuroscience Abstracts*, vol. 16, pp 473, 1990.
 - [47] Sharp, P.E., "Computer simulation of hippocampal place cells", *Psychobiology*, vol. 19, pp 103-115, 1991.
 - [48] Squire, L.R., *Memory and Brain*, Oxford University Press, New York, 1987.
 - [49] Teyler, T.J. and DiScenna, P., "The topological anatomy of the hippocampus: a clue to its function", *Brain Research Bulletin*, vol. 12, pp 711-719, 1984.
 - [50] Teyler, T.J. and DiScenna, P., "The hippocampal memory indexing theory", *Behavioral Neuroscience*, vol. 100, pp 147-154, 1986.
 - [51] Wiener, N., *Cybernetics: Control and Communication in the Animal and the Machine*, MIT Press, Cambridge, 1948.
 - [52] Wu, J., Penna, M.A., Madhavan, P.G. and Zheng, L., "Cognitive Map Building and Navigation", *4th Conference on Neural Networks and Parallel Distributed Processing*, Fort Wayne, 1992.
 - [53] Zipser, D., "A computational model of hippocampal place fields", *Behavioral Neuroscience*, vol. 99, pp 1006-1018, 1985.
 - [54] Zipser, D. and Anderson, R., "A backpropagation network that simulates response properties of a subset of posterior parietal neurons", *Nature*, vol. 331, pp 629-684, 1988.

