Spatial Cognition and Neuro-Mimetic Navigation: A Model of Hippocampal Place Cell Activity

Angelo Arleo and Wulfram Gerstner

Centre for Neuro-Mimetic Systems, MANTRA,
Swiss Federal Institute of Technology Lausanne,
CH-1015, Lausanne, EPFL, Switzerland.

Received: 02 July 1999 / Accepted in revised form: 20 March 2000

Abstract. A computational model of hippocampal activity during spatial cognition and navigation tasks is presented. The spatial representation in our model of the rat hippocampus is built on-line during exploration via two processing streams. An allothetic vision-based representation is built by unsupervised Hebbian learning extracting spatio-temporal properties of the environment from visual input. An idiothetic representation is learned based on internal movement-related information provided by path integration. On the level of the hippocampus, allothetic and idiothetic representations are integrated to yield a stable representation of the environment by a population of localized overlapping CA3-CA1 place fields. The hippocampal spatial representation is used as a basis for goal-oriented spatial behavior. We focus on the neural pathway connecting the hippocampus to the nucleus accumbens. Place cells drive a population of locomotor action neurons in the nucleus accumbens. Reward-based learning is applied to map place cell activity into action cell activity. The ensemble action cell activity provides navigational maps to support spatial behavior. We present experimental results obtained with a mobile Khepera robot.

1 Introduction

As the complexity of the tasks and the perceptual capabilities of biological organisms increase, an explicit spatial representation of the environment appears to be employed as a cognitive basis to support navigation [25]. In rodents, hippocampal *place cells* exhibit such a spatial representation property. Recordings from single place cells in the rat hippocampus [24, 25] show that these neurons fire as a function of the rat's spatial location. A place cell shows action potentials only when the animal is in a specific region of the environment, which defines the *place field* of the cell. Place cells have been observed in the *hippocampus proper* (CA3 and CA1 pyramidal cells) [24, 41], and in other extra-hippocampal areas such as the *dentate gyrus* [16], the *entorhinal cortex* [29], the *subiculum* [36], and the *parasubiculum* [38].

Correspondence to: Angelo Arleo, angelo.arleo@epfl.ch

In addition, recent experimental findings show the existence of *head-direction cells*, neurons whose activity is tuned to the orientation of the rat's head in the azimuthal plane. Each head-direction cell fires maximally when the rat's head is oriented in a specific direction, regardless of the orientation of the head with respect to the body, and of the rat's spatial location. Thus, the ensemble activity of head-direction cells provides a neural allocentric compass. Head-direction cells have been observed in the hippocampal formation and in particular in the *postsubiculum* [39], in the *anterior thalamic nuclei* [1, 17], and in the *lateral mammillary nuclei* [18].

Place coding and directional sense are crucial for solving spatial learning tasks. Hippocampal lesions seriously impair the rat's performance in spatial tasks (see [31] for an experimental review). This supports the hypothesis that the hippocampus plays a functional role in rodent navigation, and that it provides a neural basis for spatial cognition and spatial behavior [24, 25, 41, 20].

Hippocampal place fields are determined by a combination of environmental cues whose mutual relationships code for the current animal location [25]. Experiments on rats show that visual cues are of eminent importance for the formation of place fields [17]. Nevertheless, rats also rely on other allothetic non-visual stimuli, such as auditory, olfactory, and somatosensory cues [15]. Moreover, place cells can maintain stable receptive fields even in absence of reliable allothetic cues (e.g., in the dark) [28]. This suggests a complex architecture where multimodal sensory information is used for learning and maintaining hippocampal place fields. In the dark, for instance, idiothetic information (e.g., proprioceptive and vestibular stimuli) might partially replace external cues [10].

We present a computational model of the hippocampus which relies on the idea of *sensor-fusion* to drive place cell activity. External cues and internal self-generated information are integrated for establishing and maintaining hippocampal place fields. Receptive fields are learned by extracting spatio-temporal properties of the environment. Incoming visual stimuli are interpreted by means of neurons that only respond to combinations of specific visual patterns. The activity of these neu-

rons implicitly represents properties like agent-landmark distance and egocentric orientation to visual cues. In a further step, the activity of several of these neurons is combined to yield place cell activity. *Unsupervised Hebbian learning* is used to build the hippocampal neural structure incrementally. In addition to visual input we also consider idiothetic information. An extra-hippocampal path integrator drives Gaussian-tuned neurons modeling internal movement-related stimuli. During the agent-environment interaction, synapses between visually driven cells and path-integration neurons are established by means of Hebbian learning. This allows us to correlate allothetic and idiothetic cues to drive place cell activity. The proposed model results in a neural spatial representation consisting of a population of localized overlapping place fields (modeling the activity of CA1 and CA3 pyramidal cells). To interpret the ensemble place cell activity as spatial location we apply a *population vector coding* scheme [14, 41].

In order to accomplish its functional role in spatial behavior, the proposed hippocampal model must incorporate the knowledge about relationships between the environment, its obstacles and specific target locations. As in Brown and Sharp [2], and in Burgess *et al.* [4], we apply *reinforcement learning* [37] to enable target-oriented navigation based on hippocampal place cell activity. We focus on a specific neural pathway, namely the *fornix* projection, connecting the hippocampus (in particular the CA1 region) to the *nucleus accumbens*. The latter is an extra-hippocampal structure that is probably involved in reward-based goal memory and in locomotor behavior [2, 31]. Place cell activity drives a population of *locomotor action neurons* in the nucleus accumbens [2]. Synaptic efficacy between CA1 cells and action cells is changed as a function of target-related reward signals. This results in an ensemble activity of the action neurons that provides a navigational map to support spatial behavior.

In order to evaluate our hippocampal model in a real context, we have implemented it on a Khepera miniature mobile robot (Fig. 6(b)). Allothetic information is provided by a linear vision system, consisting of 64 photo-receptors covering 36 degrees of azimuthal range. Eight infrared sensors provide obstacle detection capability (similar to whiskers). Internal movement-related information is provided by dead-reckoning (odometry). Robotics offers a useful tool to validate models of functionalities in neuro-physiological processes [27]. Artificial agents are simpler and more experimentally transparent than biological systems, which makes them appealing for understanding the nature of the underlying mechanisms of animal behavior.

Our approach is similar in spirit to earlier studies [4, 32, 31, 13, 40, 19, 2]. In contrast to Burgess, Recce and O'Keefe [4], we do not directly use metric information (i.e., distance to visual cues) as input for the model. Rather, we interpret visual properties by learning a population of neurons sensitive to specific visual stimulation. Moreover, there is no path integration in the model of Burgess *et al.*. In contrast with their model, we consider, along with vision, the path integrator as an important constituent of our hippocampal model. This allows us to account for the existence of place fields in the absence of visual cues (e.g., in complete darkness) [28]. Redish and Touretzky [32, 31] have put forward a comprehensive theory of the hippocampal functionality where place fields are important ingredients. Our approach puts the focus on how place fields in the CA3-CA1 areas might be built from multimodal sensory inputs (i.e., vision and path integration). Gaussier et *al.* [13] propose a model of the hippocampal functionality in long-term consolidation and temporal sequence processing. Trullier and Meyer [40] build a topological representation of the environment from sequences of local views. In contrast to those two models, temporal aspects are, in our approach, mainly *implicit* in the path integration. In contrast to Mallot *et al.* [19], who construct a *sparse* topological representation, our representation is rather redundant and uses a large number of place cells. Similarly to Brown and Sharp [2], we consider the cell activity in the nucleus accumbens to guide navigation. However, we do not propose an explicit model for the nucleus accumbens. Finally, similarly to Schultz, Dayan and Montague [35, 6] we consider the role of dopaminergic neurons in reward-based learning. However, we study hippocampal goal-oriented navigation in a real agent-environment context.

2 Spatial Representation in the Hippocampus

2.1 Biological Background

Fig. 1 shows the functional rationale behind the model: (*i*) External stimuli (i.e., visual data) are interpreted to characterize distinct regions of the environment by distinct sensory configurations. This results in an *allothetic (vision-based) spatial representation* consistent with the local view hypothesis suggested by McNaughton in 1989 [20]. (*ii*) Internal movement-related stimuli (i.e., proprioceptive and vestibular) are integrated over time to provide an *idiothetic (path integration-based) representation*. (*iii*) Allothetic and idiothetic representations are combined to form a stable spatial representation in the hippocampus (CA3-CA1 place fields). (*iv*) Spatial navigation is achieved based on place cell activity, desired targets, and rewarding stimulation.

Fig. 2 shows the anatomical framework underlying our computational model. The *hippocampus proper* (C-shaped structure in Fig. 2) consists of the CA3-CA1 areas. The *hippocampal formation* consists of the hippocampus proper, the dentate gyrus (DG), the entorhinal cortex (in particular, we consider superficial (sEC) and medial (mEC) entorhinal regions), and the subiculum (SC).

The hippocampus receives multimodal highly processed sensory information mainly from neocortical areas, and from subcortical areas (e.g., inputs from the medial septum via the fornix fiber bundle) [3]. We focus on neocortical inputs and in particular on the information coming from the posterior parietal cortex. Lesion data on humans and monkeys, suggest that parietal areas are involved in spatial cognition and spatial behavior

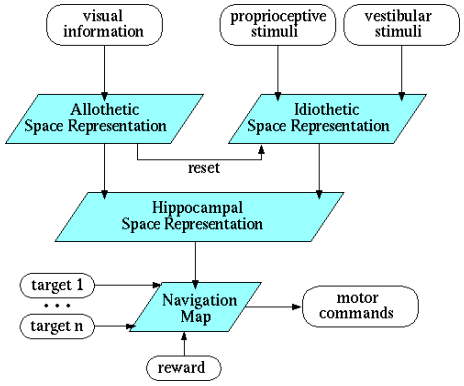


Fig. 1. Functional overview of the model. Allothetic and idiothetic stimuli are combined to yield the hippocampal space representation. Navigation is based on place cell activity, desired targets, and rewards.

[3]. The posterior parietal cortex receives inputs from visual, sensory-motor, and somatosensory cortices. This information reaches the entorhinal regions, within the hippocampal formation, via the parahippocampal (PaHi) and the perirhinal (PeRh) cortices. Finally, the entorhinal cortex projects to the hippocampus proper via the perforant path (PP) [3].

As previously mentioned, we consider the spatial representation in the CA3-CA1 areas as the result of integrating idiothetic and allothetic representations (Fig. 1). The idiothetic representation is assumed to be environment-independent. Recordings from cells in the medial entorhinal cortex (mEC) show place fields with a topology-preserving property across different environments [29, 31]. Thus, we suppose that the idiothetic representation takes place in the medial entorhinal cortex. A fundamental contribution to build the idiothetic space representation in mEC comes from the head-direction system (Fig. 2). The latter is formed by the neural circuit including the lateral mammillary nuclei (LMN), the anterodorsal nucleus of anterior thalamus (ADN), and the postsubiculum (poS) [1, 31]. Head-direction information is projected to the medial entorhinal cortex (mEC) from the postsubiculum (poS).
On the other hand, we suppose that the allothetic representation is formed in the superficial entorhinal cortex (sEC) [31]. Superficial layers of the entorhinal cortex receive spatial information about allothetic landmarks (local view) from the posterior parietal cortex, and project massively to the CA3 region via the perforant path [31].

The hippocampus proper projects its output *(i)* to the subiculum and the deep layers of the entorhinal cortex via the angular bundle, *(ii)* to several subcortical areas (e.g., the nucleus accumbens (NA)) via the fornix (FX). In particular, we consider the output of CA1 cells that reaches the nucleus accumbens via the fornix[1]. We identify the NA as the area where navigation control is achieved by means of reward-based learning [2, 32]. We consider the dopaminergic input that NA receives from the ventral tegmental area (VTA). Indeed, dopamine neuron activity codes for external rewarding stimulation [35].

[1] Actually, the fornix receives most of its inputs from the subiculum. However, experiments show that CA1 cells also project into it [31].

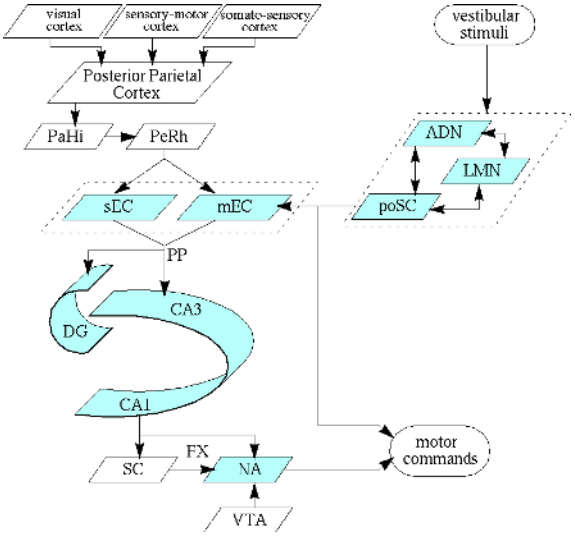


Fig. 2. A simplified overview of the anatomical counterparts of the constituents of our model. Glossary: PaHi: parahippocampal cortex, PeRh: perirhinal cortex, poSC: postsubiculum, LMN: lateral mammillary nuclei, ADN: anterodorsal nucleus of anterior thalamus (ATN), mEC: medial entorhinal cortex, sEC: superficial entorhinal cortex, DG: dentate gyrus, SC: subiculum, NA: nucleus accumbens, VTA: ventral tegmental area, PP: perforant path, FX: fornix. The hippocampus proper consists of the CA3-CA1 areas. The hippocampal formation consists of the hippocampus proper, the dentate gyrus, the entorhinal cortex, and the subicular complex [31, 3]. Adapted from Redish and Touretzky [32], and from Burgess *et al.* [3].

2.2 Learning Place Fields

The model system consists of a multi-layer neural architecture that models high-dimensional continuous sensory input by means of overlapping place fields. Starting with no prior knowledge, the system grows incrementally and on-line as the agent interacts with the environment. Unsupervised Hebbian learning is used to detect the low-dimensional view manifold representing the visual input space. However, since distinct spatial locations might provide identical visual stimuli, such a view manifold might be singular [19]. Hebbian learning is applied to correlate visual cues and path integration in order to remove such singularities. The combination of internal and external stimuli yields a stable state space representation. On the one hand, unreliable visual data can be compensated for by means of path integration. On the other hand, reliable visual information can be used to reset the path integrator system.

2.2.1 Representation of Visual Input

We apply a simple computational strategy to emulate the feature-extraction mechanism observed in the visual cortex. Moving up the visual pathway, visual neurons become responsive to stimuli of increasing complexity, from orientation sensitive cells, to neurons sensitive to more complex patterns, such as faces [33].

We model spatio-temporal relationships between visual cues by means of neural activity. Incoming visual stimuli are interpreted by mapping images into a *filter-activity space* (Fig. 3). We define several classes of *Walsh*-like filters[2]. Each class corresponds to a specific visual

[2] Walsh filters are simple and permit effective and low-cost feature-detection in one-dimensional visual spaces. We are cur-

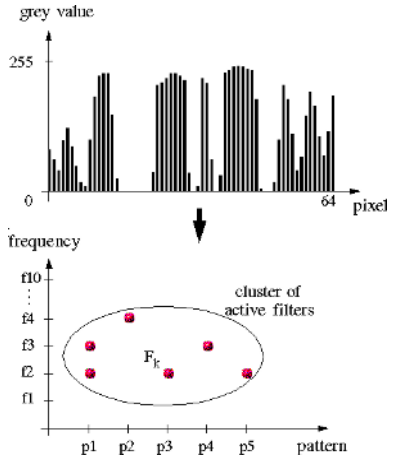


Fig. 3. Linear images (top) are mapped into a filter activity space (bottom). Along the x-axis we have different Walsh-like filters, $p_1, \ldots, p_n$, each of which responds to a specific pattern. Along the y-axis the spatial frequency of each pattern p_i is varied to represent the same pattern seen from different distances. Each image is encoded by the cluster of filters which maximally respond to that image.

pattern. The set of filters in that class corresponds to different spatial frequencies for that pattern (which endows the system with a distance discrimination property). In total we define 5 different classes of filters each containing filters at 10 different frequencies. Let F_k be one our *Walsh* filters, where $1 \leq k \leq 50$ is the index of the filter, and let l_k be its length (i.e., number of pixels covered by the filter). Given the input image $\mathbf{x} = (x_1, \ldots, x_{64})$, the response a_k of filter F_k is computed by convolution

$$a_k = \max_n \left\{ \sum_{i=0}^{l_k - 1} F_k(i)\, x_{n+i} \right\} \tag{1}$$

where $0 \leq n \leq 64 - l_k$. Since $-1 \leq x_j \leq 1$ and $F_k(i) = \pm 1$ for all i, k, the relationship $|a_k| \leq l_k$ holds.

Each neural filter F_k responds to a particular pattern. In order to detect more complex features, we consider a layer of visual cells one synapse downstream the neural filter layer. We call these neurons *snapshot cells*. The idea is to represent each image by the cluster of filters with the highest activation value, defined by Eq. 1. Let $C = 0.7 \cdot l_k$ be the threshold above which a filter F_k is considered as active. Given an image $\mathbf{x}$, the set of active filters projects one layer forward to form a snapshot cell

$$sc = \{F_k \mid a_k \geq C\} \tag{2}$$

The firing activity r_j of a snapshot cell sc_j is given by

$$r_j = \frac{\sum_{k \in sc_j} \mathcal{H}(a_k - C)}{N_j} \tag{3}$$

where $\sum_{k \in sc_j}$ sums over all the N_j filters projecting to the cell sc_j, and $\mathcal{H}$ is the Heaviside function. The normalization has been chosen so that $0 \leq r_j \leq 1$.

2.2.2 *Allothetic Representation: Place Fields in the Superficial Entorhinal Cortex*

The activity of snapshot cells depends on the current gaze direction, and does not truly code for a spatial location. In order to achieve spatial sensitivity, we apply unsupervised Hebbian learning to create a population of place cells one synapse downstream of the snapshot cell layer. We suppose that the anatomical counterpart for this neural layer is the superficial entorhinal cortex (Fig. 2). We call these neurons *sEC cells*.

Every time the robot is at a new location, all simultaneously active snapshot cells are connected to a newly created sEC cell. Each new synapse is given a random weight in $(0, 1)$. Let i and j be indices for sEC cells and snapshot cells, respectively. If r_j is the firing activity of a snapshot cell j, then

$$w_{ij}^{new} = \mathcal{H}(r_j - \epsilon)\, \mathrm{rnd}_{0,1} \tag{4}$$

where $\epsilon = 0.75$ is the activity threshold above which a snapshot cell is considered to be active. The firing rate r_i of a sEC cell i is given by the average activity r_j of its presynaptic neurons j

$$r_i = \frac{\sum_j w_{ij} r_j}{\sum_j w_{ij}} \tag{5}$$

Once synapses are established, their efficacy is changed according to a Hebbian learning rule

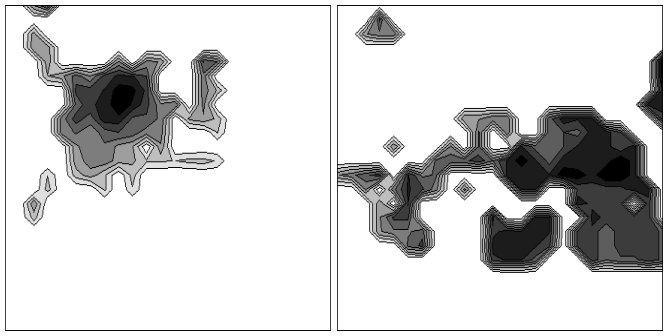
$$\Delta w_{ij} = r_j\, (r_i - w_{ij}) \tag{6}$$

where j is the index of the presynaptic neuron. If the robot is visiting a spatial location, it first checks whether there are already sEC cells coding for this location. New connections from snapshot cells to new sEC cells are created only if

$$\sum_i \mathcal{H}(r_i - \epsilon) < A \tag{7}$$

that is, only if the number of sEC cells activated at that location does not exceed a threshold A. Eq. 7 is a mere algorithmic implementation. We believe, however, that in some way rodents must have a possibility to detect novelty. Eq. 7 allows the system to control the redundancy level in the resulting spatial representation. We call the learning scheme defined by Eqs. 4, 6, and 7, an *unsupervised growing network* (see, e.g., [11]).

By definition, each sEC cell is driven by a set of snapshot cells whose activities code for visual features



(a) (b)

Fig. 4. Two examples of receptive fields of cells in our superficial entorhinal layer. The darker a region, the higher the firing rate of the cell when the robot is in that region of the environment. (a) The visual input is reliable, so that the maximal activity is confined to a localized spot in the environment. (b) The receptive field has multiple peaks indicating that similar visual stimuli occur at different locations.

rently implementing our model on a two-dimensional vision system by using biologically inspired Gabor filters [12].

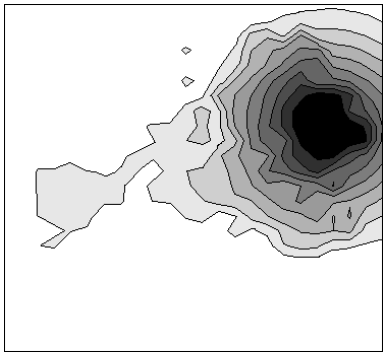


Fig. 5. A sample of place field of a place cell in our CA3-CA1 hippocampal layer. When the robot is in the region of the black spot the firing rate of the cell is maximal. Notice the gaussian-like tuning curve, which is compatible with single cell recordings from real place cells.

of the environment. As a consequence, the activity of a sEC cell depends on the combination of multiple visual cues. This results in an ensemble sEC cell activity coding for spatial locations. Fig. 4 shows two examples of place fields in the superficial entorhinal layer of the model. The darker a region, the higher the firing rate of the cell. Fig. 4(a) shows that the cell is activated only if the robot is in a localized region of the environment. Thus, the robot may use the center of the field (the darkest area) for the self-localization task. On the other hand, Fig. 4(b) shows a cell with multiple subfields. The activity of this sEC cell encodes an ambiguous visual input: the multi-peak receptive field identifies different spatial locations which yield similar visual stimuli. About 70% of the cells in our superficial entorhinal layer are of type (a), and about 30% of type (b). As previously mentioned, a way to solve the ambiguities of cell-type (b) is to consider along with the visual input the internal movement-related information provided by the path integrator (i.e., dead-reckoning), which is the topic of Sec. 2.2.3.

Place fields in our model of sEC are non-directional. This is due to the fact that sEC cells bind together the several snapshot cells that correspond to the north, east, south, and west views. Experimental data show that place cells tend to have directional place fields (i.e., their firing activity depends on head direction) in very structured arenas (e.g., linear track mazes and radial narrow arm mazes [22]). On the other hand, when the rat can freely move over two-dimensional open environments (e.g., the arena of Fig. 6(a)) place fields tend to be non-directional [23]. In order to obtain directionally independent place fields in our model, the system takes four snapshots corresponding to the north, east, south, and west views at each location visited during exploration [4]. Thus, each visited location in the environment is characterized by four snapshot cells, which are bound together to form a non-directional local view. On the other hand, in a linear track maze the rat always runs in the same direction. If we would model this by taking a *single* view only, then we would get directionality.

2.2.3 *Idiothetic Representation: Place Fields in the Medial Entorhinal Cortex*

In this paper we do not present an explicit model for the path integrator system [8]. We simply define extra-hippocampal neurons, namely *path-integration cells* (*PI cells*), whose activity provides an allocentric spatial representation based on dead-reckoning [21]. Thus, as the robot moves, the activity of the PI cells changes according to proprioceptive stimuli and the robot's orientation provided by the head-direction system. The firing rate r_p of a PI cell p is taken as a Gaussian

$$r_p = \exp\left(\frac{-(\mathbf{p}_{dr} - \mathbf{p}_p)^2}{2\sigma^2}\right) \tag{8}$$

where $\mathbf{p}_{dr}$ is the position estimated by dead-reckoning, $\mathbf{p}_p$ is the center of the field of cell p, and σ is the width of the Gaussian field. In the current implementation, the value of the dead-reckoning position $\mathbf{p}_{dr}$ is evaluated by direct mathematical integration of the movement signals (wheel turns). The activity of the PI cells is environment-independent, that is place fields of PI cells do not change from environment to environment [32]. We suppose that the spatial representation provided by the PI place fields takes place in the medial entorhinal cortex [29] (Fig. 2).

Our PI cell assembly could be interpreted as one of the charts of the *multichart path integrator* proposed by McNaughton *et al.* [21]. A chart is an imaginary frame of reference appropriately mapped into the environment and where each cell is located at the center of its place field. In the model of McNaughton *et al.* several charts are stored in the same recurrent network. Additional spatial reference cues trigger which chart to pick so that different charts are mapped into different environments. Our system would correspond to one finite chart. Since in this study we concentrate on a single environment only, we have not implemented how the system would switch to a new chart if it leaves the reference frame [31].

2.2.4 *Hippocampal Representation: Place Fields in the CA3 and CA1 Regions*

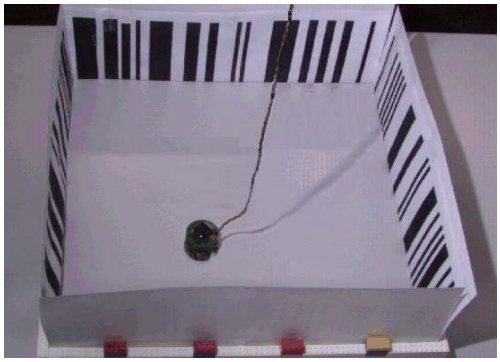
Allothetic and idiothetic representations converge onto the hippocampus proper to form a spatial representation based on CA3-CA1 place fields.

sEC cells project to CA3-CA1 neurons by means of downstream synapses that are incrementally created by applying our *unsupervised growing network* scheme (Eqs. 4, 6, and 7). Simultaneously active sEC cells are connected to create new CA3-CA1 place cells. If i and j represent CA3-CA1 place cells and sEC cells, respectively, synapses are created according to Eq. 4 and they are changed on-line by Hebbian learning (Eq. 6). The firing rate of each CA3-CA1 cell is a weighted average of the activity of its presynaptic cells (Eq. 5).

In addition, during the agent-environment interaction, Hebbian learning is used to learn synapses between PI cells and CA3-CA1 place cells. If i and p represent a place cell in the hippocampus and a PI cell, respectively, the synaptic weight w_{ip} is established according to

$$\Delta w_{ip} = r_p r_i (1 - w_{ip}) \tag{9}$$

As a consequence, the place cell activity in the CA3-CA1 layer depends on the activity of both sEC cells and PI cells. This combination of internal and external stimuli



(a)

(b)

Fig. 6. (a) The experimental setup: The 60×60 cm square arena with the Khepera robot inside. Walls are covered by a random sequence of black and white stripes of variable width, which form the visual input patterns for the system. (b) The mobile Khepera robot equipped by a linear-vision system. Eight infrared sensors provide obstacle detection capability. Two motors drive the two wheels independently. Two wheel encoders provide the dead-reckoning system. In this configuration the robot is about 7 cm tall with a diameter of about 6 cm.

yields a rather stable spatial representation. Fig. 5 shows a typical receptive field of a place cell in the CA3-CA1 layer of our model. Again, the darker a region, the higher the firing rate of the cell.

About 3% of our CA3-CA1 place cells show multiple subfields. This is consistent with experimental single-unit recordings data which show that about 5% of observed cells have multiple subfields within a single environment [30].

2.3 Population Vector Coding

The proposed model yields a spatial representation consisting of a large number of overlapping place fields. Fig. 6(a) shows the square arena used for the experiments with the mobile Khepera robot (Fig. 6(b)). Walls are covered by random sequences of black and white stripes of variable width. Combinations of these stripes form the input patterns for the linear vision system. During exploration (see Sec. 2.4) the robot tries to cover the two-dimensional space uniformly and densely by a population of CA3-CA1 place fields. Fig. 7 shows the distribution of CA3-CA1 place cells after learning. Each dot represents a place cell, and the position of the dot represents the center of the place field. In this experiment the robot, starting from an empty population, created about 800 CA3-CA1 place cells.

The ensemble place cell activity shown in Fig. 7 codes for the robot's location in Fig. 6(a). The darker a cell, the higher its firing rate. In order to interpret the information represented by the ensemble pattern of activity, we apply *population vector coding* [14]. This approach has

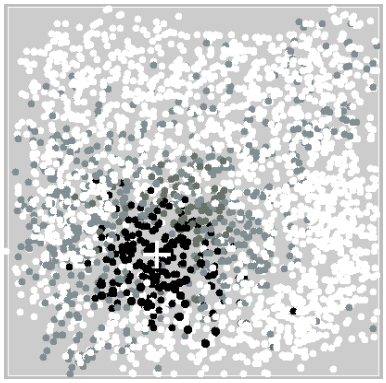


Fig. 7. The learned population of CA3-CA1 place cells. Each dot denotes the center of a place field. The darker a dot, the higher the firing rate of the corresponding place cell. The ensemble activity corresponds to the robot's location in Fig. 6(a). The white cross represents the center of mass of the population activity.

been successfully applied to interpret the neural activity in the hippocampus [41]. We average the activity of the neural population to yield the encoded spatial location. Let us suppose that the robot is at an unknown location $\mathbf{s}$. If $r_i(\mathbf{s})$ is the firing activity of a neuron i and $\mathbf{x}_i$ is the center of its place field, the population vector $\mathbf{p}$ is the center of mass of the network activity:

$$\mathbf{p}(\mathbf{s}) = \frac{\sum_i \mathbf{x}_i \, r_i(\mathbf{s})}{\sum_i r_i(\mathbf{s})} \tag{10}$$

Notice that the encoded spatial position $\mathbf{p}$ is near, but not necessarily identical to, the true location $\mathbf{s}$ of the robot. The approximation $\mathbf{p} \approx \mathbf{s}$ is good for large neural populations covering the environment densely and uniformly [34]. In Fig. 7 the center of mass (Eq. 10) coding for the robot's location is represented by the white cross.

Note that the place field center $\mathbf{x}_i$ has been made explicit for interpreting and monitoring purposes only. Associated with each place cell i is a vector $\mathbf{x}_i$ which represents the estimated location of the robot (based on dead-reckoning) when it creates the cell i. While the vector $\mathbf{x}_i$ is used in Eq. 10 for the interpretation of the population activity, knowledge of $\mathbf{x}_i$ is not necessary for navigation as discussed later in Sec. 3.

2.4 Exploration and Path Integrator Calibration

The robot moves in discrete time steps Δt which determine the frequency at which it senses the world, interprets sensory inputs, and takes an action. Experiments on rats show that, during motion, hippocampal processing is timed by a sinusoidal EEG signal of $7 - 12$ Hz, namely the *theta rhythm*. The activity of hippocampal cells is correlated to the phase of theta [26]. We assume that each time step Δt corresponds to one theta cycle of approximately 0.1 seconds, thus place cell activity is updated with a frequency of 10 Hz (the real movement of the robot is, of course, slower than this).

The robot uses a simple *active-exploration* technique which helps to cover the environment uniformly. At each time step, it chooses its new direction of motion based on the activity in the CA3-CA1 layer. If a relatively large number of neurons are currently active, it means that a well known region of the environment is being visited.

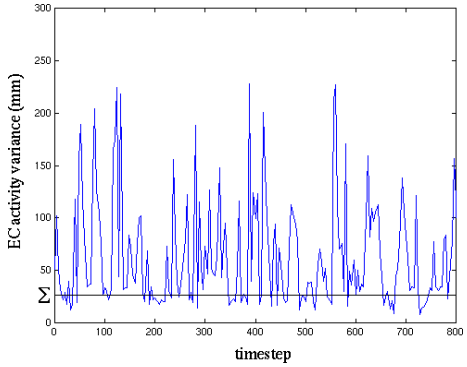


Fig. 8. The variance of the sEC cell activity around the center of mass $\mathbf{p}_{sec}$. When the variance falls below the fixed threshold Σ the spatial location $\mathbf{p}_{sec}$ is used to calibrate the robot's position.

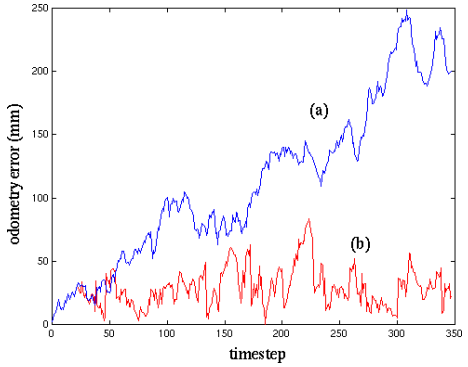


Fig. 9. Uncalibrated dead-reckoning error (curve (a)) versus calibrated robot positioning using sEC cell activity (curve (b)).

Then, a small directional change, $\Delta\phi_s$, will increase the probability of leaving that area. Conversely, a large variability of the robot's direction, $\Delta\phi_l$, is associated to low CA3-CA1 place cell activity, which results in a thorough exploration of that region. In our experiments $\Delta\phi_s$ and $\Delta\phi_l$ are randomly drawn from $[-5, +5]$ and $[-60, +60]$, respectively.

Path integration is vulnerable to cumulative errors in both biological and robotics systems [10]. As a consequence, to maintain the allothetic and idiothetic representations consistent over time, we need to bound dead-reckoning errors by occasionally resetting the path integrator. Visual information may be used to accomplish this aim [21].

The robot adopts an exploration strategy which emulates the exploratory behavior of animals [10, 5]. It starts from an initial location (e.g., the nest) and, as exploration proceeds, it creates new place cells. At the very beginning, exploration consists of short return trips (e.g., narrow loops) which are centered in the nest and directed towards the principal radial directions (e.g., north, north-east, east, and so on). This overall behavior relies on the head-direction system and allows the robot to explore the space around the nest exhaustively. Afterwards, the robot switches to a more open-field exploration strategy. It starts moving in a random direction and it uses the above active-exploration technique to update its direction at each time step. After a while, the robot "feels" the need to re-calibrate its path integrator. We do not propose a specific uncertainty model for the dead-reckoning system. We simply assume that the "need of calibration" grows monotonically as some function $n(t)$ of time t. When, after a time t_{cal}, $n(t)$ overcomes a fixed threshold n_{cal}, the robot stops creating place cells and starts following the homing vector [10, 5] to return towards the nest location. As soon as the robot finds a previously visited location (not necessarily the nest location), it tries to use the learned allothetic spatial representation to localize itself.

We take the visually driven activity of sEC cells as the signal for the calibrating process. Let $\mathbf{p}_{sec}$ be the center of mass of the sEC cell activity and let σ be the variance of the activity around it. In order to evaluate the reliability of the sEC cell activity, we consider a fixed variance threshold Σ. If σ is smaller than Σ, then the spatial location $\mathbf{p}_{sec}$ is suitable for re-calibrating the robot (Fig. 8). More precisely, we define a weight coefficient

$$\alpha = \begin{cases} 1 - \frac{\sigma}{\Sigma} & \sigma \leq \Sigma \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

and then we use it to compute the calibrated robot position $\mathbf{p}^*$

$$\mathbf{p}^* = \alpha\ \mathbf{p}_{sec} + (1 - \alpha)\ \mathbf{p}_{dr} \tag{12}$$

where $\mathbf{p}_{dr}$ is the position estimated by the dead-reckoning system. Eq. 12 is an algorithmic implementation. In the future, we would like to implement odometry calibration by applying associative learning to correlate the sEC cell activity to the PI cell activity.

Once the robot has calibrated itself, exploration is resumed and it starts creating new place cells. This technique allows the robot to explore the environment by keeping the dead-reckoning error within a bounded range. Fig. 9 shows calibrated versus uncalibrated path-integrator error during an exploration session of about 350 time steps.

Even though this case has never occurred in our experiments, during the homing behavior the robot might reach the nest without having re-calibrated its path integration (i.e., without having found a location where sEC activity is suitable to calibrate odometry). In this case, the robot resorts to a spiral searching behavior centered around the nest location. As soon as it finds a calibration location, the open-field exploring behavior is resumed.

3 Spatial Behavior: Learning Navigation Maps

The above hippocampal model allows the robot to self-localize itself within its environment (Fig. 7). In order to provide a cognitive support for spatial behavior, place cell activity has to be used to guide navigation. We derive *navigational maps* by applying *reinforcement learning* [37] to map CA3-CA1 ensemble activity into goal-oriented behavior. The navigation part has not been implemented on the robot yet, but done in simulation.

3.1 Reinforcement Learning in Continuous Space

The nucleus accumbens has been thought to play an important role in reward-based spatial learning [2, 31]. It receives place coding information from the hippocampal

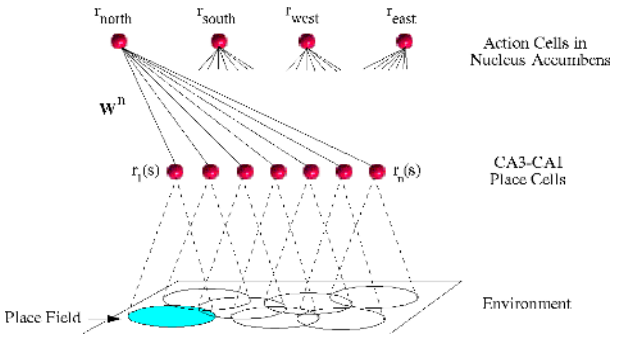


Fig. 10. CA3-CA1 place cells project to action cells (four for each target type) in the nucleus accumbens. Reinforcement learning is used to find the function which maps continuous spatial locations to locomotor actions.

formation (via the fornix) as well as rewarding stimulation from dopaminergic neurons (via the ventral tegmental area) [31].

We consider a population of *action cells* in the nucleus accumbens whose activity provides directional motor commands [2]. For each type of target (e.g., food or water), four action cells (coding for north, south, west, east allocentric actions) are driven by the population of CA3-CA1 place cells [4]. Synapses from hippocampal place cells to action cells are modified to learn the continuous *location-to-action* mapping function in goal-directed tasks. LTP occurs to associate spatial locations to rewarding actions, otherwise LTD takes place (Fig. 10).

Learning an action-value function over a continuous location space endows the system with spatial generalization capabilities. Thus, the robot may be able to associate appropriate actions to spatial positions that it has never seen before. Overlapping localized place fields in the CA3-CA1 layer, provide a natural set of basis functions that can be used to learn such a mapping function.

Let $\mathbf{s}$ be the robot's location (*state*), and let a be an action cell in the nucleus accumbens, with $a \in \mathcal{A} := \{north, south, west, east\}$. Let us denote the activation of a CA3-CA1 place cell i by r_i, and the activity of an action cell a by r_a. A robot position $\mathbf{s}$ is encoded by the place cell activity $\mathbf{r}(\mathbf{s}) = (r_1(\mathbf{s}), r_2(\mathbf{s}), \ldots, r_n(\mathbf{s}))$, where n is the number of CA3-CA1 place cells. Let $\mathbf{w}^a = (w_1^a, \ldots, w_n^a)$ be the synaptic projections from hippocampal place cells to the action cell a (Fig. 10). The activity r_a depends linearly on the robot's position $\mathbf{s}$ and on the synaptic weights $\mathbf{w}^a$:

$$r_a(\mathbf{s}) = (\mathbf{w}^a)^T \, \mathbf{r}(\mathbf{s}) = \sum_{i=1}^{n} w_i^a \, r_i(\mathbf{s}) \tag{13}$$

The learning task consists of updating $\mathbf{w}^a$ to approximate the optimal goal-oriented function which maps states $\mathbf{s}$ into action cell activity $r_a(\mathbf{s})$. To do this, we use the linear gradient-descent version of Watkins' Q-learning algorithm [37]. Given a robot position $\mathbf{s}$, we interpret the neural activity $r_a(\mathbf{s})$ as the "expected gain" when taking action a at location $\mathbf{s}$ of the environment.

During training, the robot behaves in order to either consolidate goal-directed paths (*exploitation*) or find novel routes (*exploration*). This exploitation-exploration trade-off is determined by an ϵ-greedy action selection policy, with $0 \leq \epsilon \leq 1$ [37]. At each time t, the robot takes the "optimal" action a_t^* with probability $1-\epsilon$ (exploitation)

$$a_t^* = arg \max_a r_a(\mathbf{s}_t) \tag{14}$$

or, it might resort to uniform random action selection with probability equal to ϵ (exploration). At each time step Δt, the synaptic efficacy of projections $\mathbf{w}^a$ changes according to [37]

$$\Delta \mathbf{w}^a = \alpha \, \delta_t \, \mathbf{e}_t \tag{15}$$

The terms in Eq. 15 have the following interpretation:

(i) The factor α, $0 \leq \alpha \leq 1$, is a constant learning rate.

(ii) The term δ_t is the *prediction error* defined by

$$\delta_t = R_{t+1} + \gamma \max_a r_a(\mathbf{s}_{t+1}) - r_a(\mathbf{s}_t) \tag{16}$$

where R_{t+1} is the actual reward delivered by an internal brain signal, and γ, $0 \leq \gamma \leq 1$, is a constant discount factor. The temporal difference δ_t estimates the error between the expected and the actual reward when, given the location $\mathbf{s}$ at time t, the robot takes action a and reaches location $\mathbf{s}'$ at time $t+1$. Training trials allow the robot to minimize this error signal. Thus, asymptotically $\delta_t \approx 0$, which means that, given a state-action pair, the deviation between predicted and actual rewards tends to zero.
Neuro-physiological data show that the activity of dopamine neurons in mammalian midbrain encodes the difference between expected and actual occurrence of reward stimuli [35]. In particular, the more reliably a reward is predicted, the more silent a dopaminergic neuron. Thus, the temporal difference error δ_t used to update our synaptic weights $\mathbf{w}^a$ may be thought of as a dopamine-like teaching signal.

(iii) During training paths, Eq. 15 allows the robot to memorize action sequences. Since recently taken actions are more relevant than earlier ones, we need a memory trace mechanism to weight actions as a function of their occurrence time. The vector $\mathbf{e}_t$, called *eligibility trace*, provides such a mechanism [37]. The update of the eligibility trace depends on whether the robot selects an exploratory or an exploiting action. Specifically, the vector $\mathbf{e}_t$ changes according to

$$\mathbf{e}_t = \mathbf{r}(\mathbf{s}_t) + \begin{cases} \gamma\lambda\mathbf{e}_{t-1} & \text{if exploiting} \\ 0 & \text{if exploring} \end{cases} \tag{17}$$

where λ, $0 \leq \lambda \leq 1$, is a trace-decay parameter [37], and $\mathbf{r}(\mathbf{s}_t)$ is the CA3-CA1 vector activity. We start with $\mathbf{e}_0 = \mathbf{0}$.

3.2 Behavioral Experiments

Given the experimental setup shown in Fig. 6, we define a specific target region (e.g., a feeding location) within the environment. We apply the above reward-based learning scheme to build up a navigational strategy leading the robot toward the target from any location, while avoiding obstacles. In this work, we do not address the problem of consolidating and recalling hippocampal representations [31]. We simply assume that entering a familiar environment results in recalling the hippocampal

chart associated with this environment [21]. To study robot behavior, we adopt the same protocol as employed by neuro-ethologists with rats [31]. Navigational maps are learned through a training session consisting of a sequence of trials. Each trial begins at a random location and ends when the robot reaches the target. At the beginning of each trial the robot retrieves its starting location on the hippocampal chart based on the allothetic (visually-driven) representation (Sec. 2.2.2) [21, 20].

During learning we consider a discrete set of four actions $\mathcal{A} = \{north, south, west, east\}$. However, after learning, population vector coding is applied to map $\mathcal{A}$ into a continuous action space $\mathcal{A}'$ by averaging the ensemble action cell activity. Given a position $\mathbf{s}$ of the robot, the action $a'(\mathbf{s}) \propto \binom{\cos\phi}{\sin\phi}$ is a direction in the environment encoded by the action cell activity in the nucleus accumbens

$$a'(\mathbf{s}) = \frac{\sum_{a\in\mathcal{A}} a\, r_a(\mathbf{s})}{\sum_{a\in\mathcal{A}} r_a(\mathbf{s})} \tag{18}$$

where $a_n = \binom{0}{1}$, $a_s = \binom{0}{-1}$, $a_w = \binom{-1}{0}$, and $a_e = \binom{1}{0}$ are the four principal directions. Eq. 18 results in smooth trajectories.

The experiments have been carried out with a learning rate $\alpha = 0.1$, a discount factor $\gamma = 1.0$, and a decay factor $\lambda = 0.9$. The reward-signal function $R(\mathbf{s})$ is defined by

$$R(\mathbf{s}) = \begin{cases} 1 & \text{if } \mathbf{s} = \text{target state} \\ -0.5 & \text{if } \mathbf{s} = \text{collision state} \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

where collision means contact with walls or obstacles.

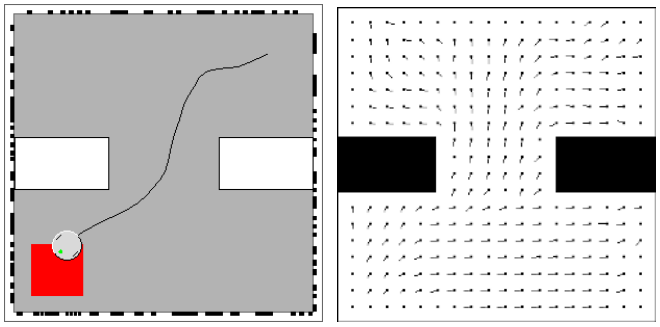
We adopt a dynamically changing ϵ-probability. The idea is to increase the probability of exploring novel routes as the time to reach the target increases. The ϵ parameter is defined by the exponential function

$$\epsilon(t) = \frac{\exp(\beta\, t) + k_1}{k_2} \tag{20}$$

where $\beta = 0.068$, $k_1 = 100$, and $k_2 = 1000$, where $t = 0, 1, 2, \ldots$ are discrete time steps. If we consider the dynamic of ϵ over a time window of 100 time steps, at $t = 0$ the robot behaves according to a value $\epsilon = 0.101$ (i.e., enhancing exploitation), and at $t = 100$ it behaves according to a value $\epsilon = 1.0$ (i.e., enhancing exploration). If at the end of the time, $t = 100$, the target is not reached yet, exploration is further enhanced by keeping a fixed $\epsilon = 1.0$ for another 100 time steps. Then, exploitation is resumed by setting $t = 0$ and $\epsilon = 0.101$. Moreover, every time the target is reached the time window is re-initialized as well, and ϵ is set equal to 0.101. These are heuristic methods to ensure a sufficient amount of exploration.

3.2.1 *Experiment with a single target type (e.g., food)*

Fig. 11(a) shows a two-dimensional view of the arena of Fig. 6(a). White objects are transparent obstacles. Only infrared sensors can detect obstacles, which are transparent with respect to the vision system. Since obstacles are not visible and have been added after learning, the place fields in the model are not affected. The dark square represents the feeder location. The target area is about 2.5 times the area occupied by the robot (grey circle). In Fig. 11(b) we show the navigational map learned by the robot in about 920 time steps, which correspond to 50 trials from random starting positions to the target. The vector field representation of Fig. 11(b) has been obtained by rastering uniformly over the whole environment. Dots represent sampled positions and pointers indicate the direction calculated from Eq. 18 at each position. Finally, the solid line shown in Fig. 11(a) is an example of a robot trajectory from a novel starting location using the learned navigational map.



(a) (b)

Fig. 11. (a) A two-dimensional view of the environment with a feeder location (dark grey square), and two obstacles (white rectangles). An example of robot trajectory induced by the action cell activity after learning. (b) Vector field representation of the learned navigational map.

3.2.2 *Moving the learned target*

This experiment consists of changing the location of a previously learned target, and of allowing the robot to adapt its navigational behavior consequently. The idea is to endow the system with an internal *reward-expectation mechanism*.

During training trials, the robot learns to correlate the CA3-CA1 place cell activity to the positive reward signal, $R = 1$, which it receives at the food location. This is achieved by considering a neuron d, that we call the *reward-expectation cell*, one synapse downstream the place cell layer (Fig. 12(a)). Let i be an index over the CA3-CA1 cell population. Connections w_{di} from place cells to the reward predicting cell are inhibitory synapses, and are initialized to random values within the interval $[-0.1, 0]$. The cell d receives as input the external rewarding stimulus R as well. The activity r_d of cell d is non linear and it is defined by

$$r_d = \begin{cases} f(\sum_i w_{di} r_i) + R & \text{if } R \geq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{21}$$

where $f(x) = \tanh(x)$. Thus, the activity of cell d depends on both the external reward R and the CA3-CA1 network activity.

In order to learn the desired correlation between the event "positive reward" and the place cell activity, we apply Hebbian learning and modify the inhibitory weights w_{di} by an amount

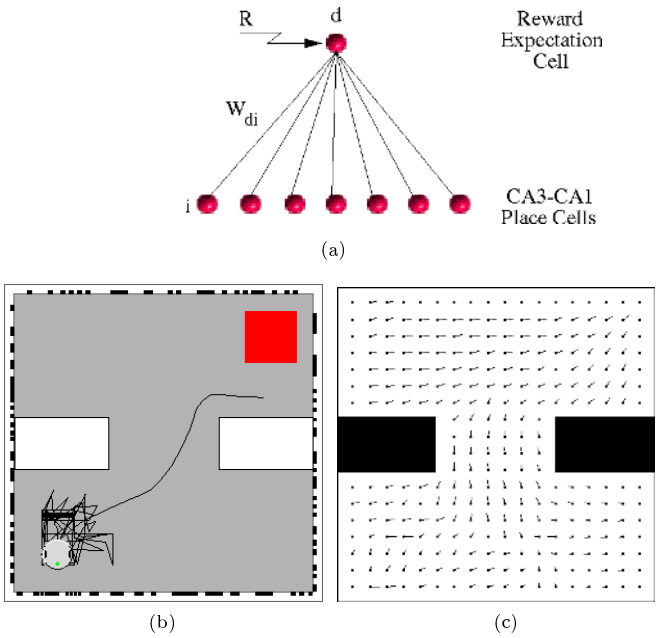


Fig. 12. (a) The internal *reward-expectation mechanism*. The activity of cell d depends on the CA3-CA1 place cell activity and on the external reward signal R. (b) The arena and the previously learned target (dark square) which has been moved to a new location. Solid lines represent trajectories of the robot searching for the previously learned food location. (c) The re-adapted navigational map corresponding to the new rewarding location.

$$\Delta w_{di} = r_i r_d (w_{di} - 1) \tag{22}$$

The more correlated the activity $r_i r_d$, the more inhibitory the synapses w_{di}.

As a consequence, before correlating the external reward signal with internal spatial representation, cell d responds maximally when the robot receives a positive $R = 1$. Indeed, since weights w_{di} are initially close to zero, the activity $r_d \approx R = 1$ (according to Eq. 21). As training proceeds, the robot starts predicting the external stimulus R by learning synapses w_{di}. Then, every time the robot is near the target location, the cell d receives a strong inhibitory input $\sum_i w_{di} r_i$ which compensates for the excitatory reward R. Thus, when R is fully predicted, even if the robot receives the $R = 1$ signal the cell d remains silent. On the other hand, if the fully predicted reward signal fails to occur (i.e., the learned target has been moved away), the activity of cell d is strongly depressed ($r_d \approx -1$), and an internal negative reward is generated. When the number of collected negative internal rewards exceeds a fixed threshold D (e.g., $D = 10$), the robot "forgets" the previous target location and starts looking for a new goal. Fig. 12(b) shows the same environment of Fig. 11(a) where the previously learned target has been moved to another location. The robot is attracted by the previous feeder position and it accumulates internal negative rewards. Fig. 12(c) presents the navigational map re-adapted to the new food location.

Our reward-expectation cell d finds its neuro-physiological counterpart in dopaminergic neurons observed in mammalian midbrain. The response of these neurons is a function of the unpredictability of incoming stimuli [35]. In particular, they respond positively to external rewards which occur unpredictably. Instead, they remain silent if a fully predicted stimulus arrives. By contrast, when a fully expected reward fails to occur, dopamine neurons respond negatively exactly at the time at which the reward is expected [35]. Instead of Eq. 21, we could have also used the prediction error δ_t defined in Eq. 16 to monitor an unexpected target location.

3.2.3 *Experiment with multiple target types (e.g., food and water)*

The reward-based learning scheme described in Sec. 3.1, Fig. 10, can also be applied to multiple target types. Let $\mathbf{T} = \{T_1, \ldots, T_m\}$ be a set of distinct target types (e.g., T_1 could be a food location, T_2 a water location, and so on). For each given target T_i we consider a set of location-to-action mapping function $r_a^{T_i}(\mathbf{s})$, and a set of synaptic weights $\mathbf{w}^{a,T_i}$. We also consider distinct rewarding signals $\mathbf{R} = \{R^{T_1}, \ldots, R^{T_m}\}$. Then, we adopt the above Q-learning algorithm to approximate the $r_a^{T_i}(\mathbf{s})$ functions.

In this experiment we consider two distinct types of rewarding stimulations T_1 (food) and T_2 (water). Fig. 13(a) shows the two target locations (left and right bottom squares) within the environment. The learning session starts by focusing on the feeder location T_1. Thus the primary task for the robot is to approximate the $r_a^{T_1}(\mathbf{s})$ functions. The navigational map learned during about 1300 time steps is shown in Fig. 13(b).

Notice that when searching for food it might happen that the robot encounters the water location and receives a positive reward signal with respect to T_2, $R^{T_2} = 1$. This information can be exploited by the robot by adjusting $\mathbf{w}^{T_2}$ weights. That is, even if T_2 is not the current target, the robot can partially learn a navigational map leading to it. Fig. 13(c) shows the knowledge about the water location T_2 acquired by the robot while learning the optimal policy to reach the food T_1. Thus, when the robot decides to focus on the water target (i.e., to approximate the $r_a^{T_2}(\mathbf{s})$ action cell activity), it does not start from zero knowledge. This results in a shorter learning time for T_2, and accelerates the robot's progress. Fig. 13(d) presents the navigational map learned by the robot after about 440 time steps when looking for water.

4 Discussion

We have presented a computational model of the hippocampus to study its role in spatial cognition and navigation. Even if it relies on neuro-physiological experimental data, the proposed neural architecture is highly simplified with respect to biological hippocampal circuitry.

In particular, we have stressed the importance of integrating external and internal stimuli to drive place cell activity in CA3-CA1 regions [28, 32]. An allothetic vision-based representation is formed in a model of the superficial entorhinal cortex. Spatial properties of the environment are extracted from visual inputs in order to

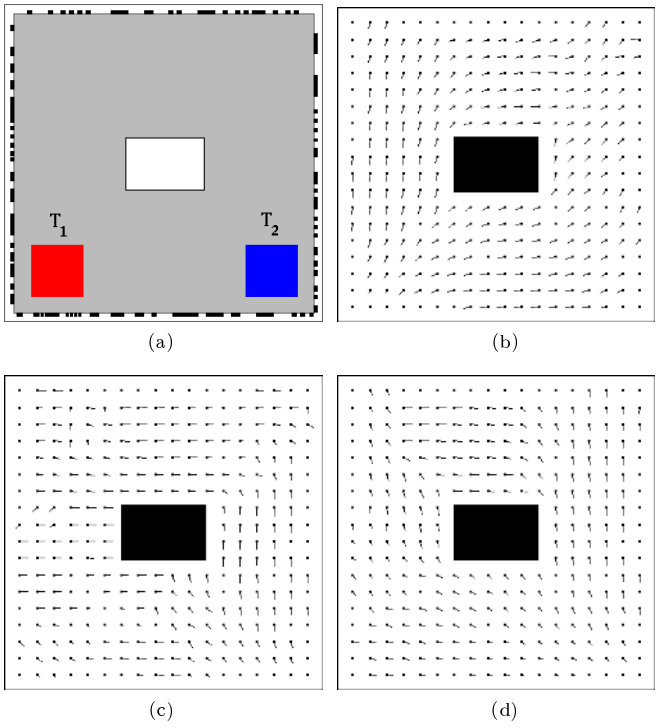


Fig. 13. (a) The arena with two distinct target types T_1 (e.g., food) and T_2 (e.g., water). The white rectangle is an obstacle. (b) The navigation map corresponding to the food rewarding location T_1. (c) The partial navigation map corresponding to the water location T_2 learned by the robot when focusing on food T_1. (d) The final map acquired by the robot when focusing on water T_2.

characterize distinct regions of the environment by combinations of visual cues. On the other hand, an idiothetic representation takes place in our model of the medial entorhinal cortex, integrating the internal movement-related information provided by proprioception. Allothetic and idiothetic representations converge onto CA3-CA1 areas of the hippocampus and form a rather stable place fields representation. Allothetic and idiothetic charts are correlated by associative learning. This induces a mutual benefit in the sense that path integration may disambiguate visual singularities [9] and, conversely, visual information may be used for resetting the path integration [21]. This process is done on-line during the development of the hippocampal space representation (i.e., exploration). A threshold mechanism is used to evaluate the reliability of the visual input being used for dead-reckoning calibration.

Unsupervised Hebbian learning is applied to build the neural system incrementally and on-line. Redundancy in the place cell activity is considered as a crucial property to yield robustness. After learning, the model has developed a spatial representation consisting of a large population of overlapping place fields covering the environment uniformly and densely. To interpret the ensemble place cell activity as spatial locations we apply population vector coding [14, 41].

The hippocampus projects to the nucleus accumbens, a subcortical structure involved in spatial behavior [2, 31]. We consider a population of locomotor action neurons [4] in the nucleus accumbens and we apply reward-based learning to adjust synapses from CA3-CA1 cells to action cells [2]. For a given target location, this results in learning a mapping function from the continuous space of physical locations to the activity space of action cells. This allows us to accomplish goal-oriented navigation based on the neural activity in the nucleus accumbens. Navigation maps are derived by interpreting the ensemble action cell activity by means of population coding [4]. Note, however, that while population vector decoding allows us an interpretation of the place cell activity, this interpretation is not necessary for action learning by reinforcement: For Q-learning, place cells are simply a set of basis functions in the high-dimensional input space. Burgess *et al.* [4] have previously postulated a goal-memory system consisting of a population of goal cells (GC) driven by hippocampal place cells. The goal cell activity encodes the animal's position with respect to the goal (i.e., north, east, south, west). In his model, however, only the activity of hippocampal cells whose place field contains the target is correlated to the GC activity by Hebbian learning. This results in GC of limited attraction radius which impairs the animal's navigation at large distances from the target and does not allow for detours around obstacles. In addition, Burgess *et al.* [4] do not propose any re-learning mechanism to cope with targets whose location might change.

A robotic platform has been used to validate our computational model in real task-environment contexts. There is, of course, a whole body of work on robot navigation with neural networks (e.g., [9, 27, 7]), but only few authors have previously implemented hippocampal models on real robots [4, 13, 19]. Understanding the underlying mechanisms of hippocampal place cell activity offers the attractive prospect of developing control algorithms that directly emulate mammalian navigational abilities. On the other hand, the simplicity and the transparency of artificial agents make them suitable for studying and understanding neuro-physiological processes.

In the future, data analysis will be focused on the dynamics of the robot behavior using the same methodology as employed by ethologists for living animals. In particular, we will evaluate our hippocampal model through experiments concerning environment manipulations (e.g., shrinking and stretching the arena, changing light conditions). We are interested in studying the potential conflicts which might occur between allothetic and idiothetic information [10], and in modeling the mutual relationships between path integration and visual stimuli. For example, a system which is dominated by vision-based information will show stretched place fields in a stretched environment, whereas a system which mainly relies on path integration will not. Hopefully, a systematic study of these effects will allow us to make neuro-ethological predictions concerning animals trained in controlled environments [10].

Acknowledgments
Supported by the Swiss National Science Foundation, project nr. 21-49174.96. The authors thank Dario Floreano for useful discussions.

References

1. H. Blair and P. Sharp. Anticipatory head direction signals in anterior thalamus: Evidence for a thalamocortical circuit that integrates angular head motion to compute head direction. *J. of Neuroscience*, 15(9):6260–6270, 1995.
2. M. Brown and P. Sharp. Simulation of spatial-learning in the Morris water maze by a neural network model of the hippocampal-formation and nucleus accumbens. *Hippocampus*, 5:171–188, 1995.
3. N. Burgess, K. Jeffery, and J. O'Keefe. Integrating hippocampal and parietal functions: A spatial point of view. In K. J. N. Burgess and J. O'Keefe, editors, *The Hippocampal and Parietal Foundations of Spatial Cognition*, chapter 1, pages 3–29. Oxford University Press, 1999.
4. N. Burgess, M. Recce, and J. O'Keefe. A model of hippocampal function. *Neural Networks*, 7:1065–1081, 1994.
5. T. Collett and J. Zeil. Places and landmarks: An arthropod perspective. In S. Healy, editor, *Spatial Representation in Animals*, chapter 2, pages 18–53. Oxford University Press, 1998.
6. P. Dayan. Navigating through temporal difference. In R. Lippmann, J. Moody, and D. Touretzky, editors, *Neural Information Processing Systems 3*, pages 464–470. Morgan Kaufmann, San Mateo, CA, 1991.
7. J. del R. Millán. Rapid, safe, and incremental learning of navigation strategies. *IEEE Trans. on Systems, Man and Cybernetics—Part B*, 26:408–420, 1996.
8. J. Droulez and A. Berthoz. The concept of dynamic memory in sensorimotor control. In D. Humphrey and H.-J. Freund, editors, *Motor Control: Concepts and Issues*, pages 137–161. John Wiley Sons, 1991.
9. T. Duckett and U. Nehmzow. Mobile robot self-localization and measurement of performance in middle scale environments. *J. of Robotics and Autonomous Systems*, 24(1-2):57–69, 1998.
10. A. Etienne, J. Berlie, J. Georgakopoulos, and R. Maurer. Role of dead reckoning in navigation. In S. Healy, editor, *Spatial Representation in Animals*, chapter 3, pages 54–68. Oxford University Press, 1998.
11. B. Fritzke. Growing cell structures –A self-organizing network for unsupervised and supervised learning. *Neural Networks*, 7(9):1441–1460, 1994.
12. D. Gabor. Theory of communication. *J. of the IEE*, 93:429–457, 1946.
13. P. Gaussier, C. Joulain, A. Revel, S. Zrehen, and J. Banquet. Building grounded symbols for localization using motivation. In *Fourth European Conf. on Artificial Life*, pages 299–308, 1997.
14. A. Georgopoulos, A. Schwartz, and R. Kettner. Neuronal population coding of movement direction. *Science*, 233:1416–1419, 1986.
15. A. Hill and P. Best. Effects of deafness and blindness on the spatial correlates of hippocampal unit activity in the rat. *Exp. Neurology*, 74:204–217, 1981.
16. M. Jung and B. McNaughton. Spatial selectivity of unit activity in the hippocampal granular layer. *Hippocampus*, 3(2):165–182, 1993.
17. J. Knierim, H. Kudrimoti, and B. McNaughton. Place cells, head direction cells, and the learning of landmark stability. *J. of Neuroscience*, 15:1648–1659, 1995.
18. C. Leonhard, R. Stackman, and J. Taube. Head direction cells recorded from the lateral mammillary nuclei in rats. *Soc. for Neuroscience Abstr.*, 22:1873, 1996.
19. H. Mallot, M. Franz, B. Schölkopf, and H. Bülthoff. The view–graph approach to visual navigation and spatial memory. In W. Gerstner, A. Germond, M. Hasler, and J. Nicoud, editors, *Artificial Neural Networks - ICANN'97. 7th Int. Conf.*, pages 751–756, Lausanne, Switzerland, 1997. Springer Verlag.
20. B. McNaughton. Neural mechanisms for spatial computation and information storage. In L. Nadel, L. Cooper, P. Culicover, and R. Harnish, editors, *Neural Connections, Mental Computation*, chapter 9, pages 285–350. MIT Press, Cambridge, MA, 1989.
21. B. McNaughton, C. Barnes, J. Gerrard, K. Gothard, M. Jung, J. Knierim, H. Kudrimoti, Y. Qin, W. Skaggs, M. Suster, and K. Weaver. Deciphering the hippocampal polyglot: the hippocampus as a path integration system. *J. of Experimental Biology*, 199:173–185, 1996.
22. B. McNaughton, C. Barnes, and J. O'Keefe. The contributions of position, direction, and velocity to single unit activity in the hippocampus of freely-moving rats. *Exp. Brain Research*, 52:41–49, 1983.
23. R. Muller, E. Bostock, J. Taube, and J. Kubie. On the directional ring properties of hippocampal place cells. *J. of Neuroscience*, 14(12):7235–7251, 1994.
24. J. O'Keefe and J. Dostrovsky. The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely moving rat. *Brain Research*, 34:171–175, 1971.
25. J. O'Keefe and L. Nadel. *The Hippocampus as a cognitive map*. Clarendon Press, Oxford, 1978.
26. J. O'Keefe and M. Recce. Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, 3:317–330, 1993.
27. R. Pfeifer and C. Scheier. *Understanding Intelligence*. MIT Press, Cambridge, MA, 1999.
28. G. Quirk, R. Muller, and J. Kubie. The firing of hippocampal place cells in the dark depends on the rat's recent experience. *J. of Neuroscience*, 10(6):2008–2017, 1990.
29. G. Quirk, R. Muller, J. Kubie, and J. Ranck. The positional firing properties of medial entorhinal neurons: Description and comparison with hippocampal place cells. *J. of Neuroscience*, 12(5):1945–1963, 1992.
30. M. Recce, A. Speakman, and J. O'Keefe. Place fields of single hippocampal cells are smaller and more spatially localized than you thought. *Soc. of Neuroscience Abstr.*, 17:484, 1991.
31. A. Redish. *Beyond the cognitive map*. PhD thesis, Dep. of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 1997.
32. A. Redish and D. Touretzky. Cognitive maps beyond the hippocampus. *Hippocampus*, 7(1):15–35, 1997.
33. E. Rolls and M. Tovée. Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *J. of Neurophysiology*, 73:713–726, 1995.
34. E. Salinas and L. Abbott. Vector reconstruction from firing rates. *J. of Comp. Science*, 1:89–107, 1994.
35. W. Schultz, P. Dayan, and R. Montague. A neural substrate of prediction and reward. *Science*, 275:1593–1599, 1997.
36. P. Sharp and C. Green. Spatial correlates of firing patterns of single cells in the subiculum of freely moving rat. *J. of Neuroscience*, 14(4):2339–2356, 1994.
37. R. Sutton and A. Barto. *Reinforcement learning, an introduction*. MIT Press-Bradford Books, Cambridge, Massachusetts, 1998.
38. J. Taube. Place cells recorded in the parasubiculum of freely moving rats. *Hippocampus*, 5(6):569–583, 1996.
39. J. Taube, R. Muller, and J. Ranck. Head direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *J. of Neuroscience*, 10:420–435, 1990.
40. O. Trullier and J. Meyer. Place sequence learning for navigation. In W. Gerstner, A. Germond, M. Hasler, and J. Nicoud, editors, *Artificial Neural Networks - ICANN'97. 7th Int. Conf.*, pages 757–762, Lausanne, Switzerland, 1997. Springer Verlag.
41. M. Wilson and B. McNaughton. Dynamics of the hippocampal ensemble code for space. *Science*, 261:1055–1058, 1993.

This article was processed by the author using the LaTeX style file *cljour2* from Springer-Verlag.