

The visual homing problem: An example of robotics/biology cross fertilization [☆]

P. Gaussier ^{a,*}, C. Joulain ^{a,1}, J.P. Banquet ^b, S. Leprêtre ^a, A. Revel ^a

^a Neurocybernetic Team, ETIS, Cergy-Pontoise University, ENSEA, 6 Av. du Ponceau, 95014 Cergy-Pontoise Cedex, France

^b Institut Neurosciences et Modélisation, Jussieu, Paris, France

Abstract

In this paper, we describe how a mobile robot under simple visual control can retrieve a particular goal location in an open environment. Our model neither needs a precise map nor to learn all the possible positions in the environment. The system is a neural architecture inspired by neurobiological analysis of how visual patterns named landmarks are recognized. The robot merges these visual informations and their azimuth to build a plastic representation of its location. This representation is used to learn the best movement to reach the goal. A simple and fast on-line learning of a few places located near the goal allows this goal to be reached from anywhere in its neighborhood. The system uses only a very rough representation of the robot environment and presents very high generalization capabilities. We describe an efficient implementation of autonomous and motivated navigation tested on our robot in real indoor environments. We show the limitations of the model and its possible extensions. ©2000 Elsevier Science B.V. All rights reserved.

Keywords: Visual navigation; Homing; Hippocampus; Place recognition; On-line learning; Autonomous robot; Neural network architecture; Motivations; Action selection

1. Introduction

Surprisingly enough, nowadays, computers have sufficient memory and power to simulate insect brains or part of mammal brains. Nevertheless, we are still unable to build really autonomous robots with insect-like cognitive capabilities. Therefore, our

goal is to use neurobiological and psychological information to design neural architectures allowing an autonomous robot to “survive” in an a priori unknown environment (*animat* approach [40,41]). For this purpose, creating a generic autonomous control system could allow a real robot to learn and perform several complex tasks during the same “life” span. Moreover, these real robots could be used as simulation tools to validate the behavioral implications of the neurobiological models supporting the control systems.

In this paper, a parallel is drawn between insect and mammal strategies used to solve this class of “survival” problems and the underlying theoretical principles are applied to control autonomous robots. A simple and generic neural network architecture named PerAc, Perception–Action, inspired by

[☆] This paper is a synthesis of conference papers devoted to visual navigation and appeared in the Proceedings Control96, IROS97, ISIE97 and SAB98.

* Corresponding author. Tel.: 33-0-1-30-73-66-13; fax: 33-0-1-30-73-66-27.

E-mail addresses: gaussier@ensea.fr (P. Gaussier), revel@ensea.fr (A. Revel).

¹ This work has been supported by the French Cognitive Science Program: GIS Sciences de la Cognition.

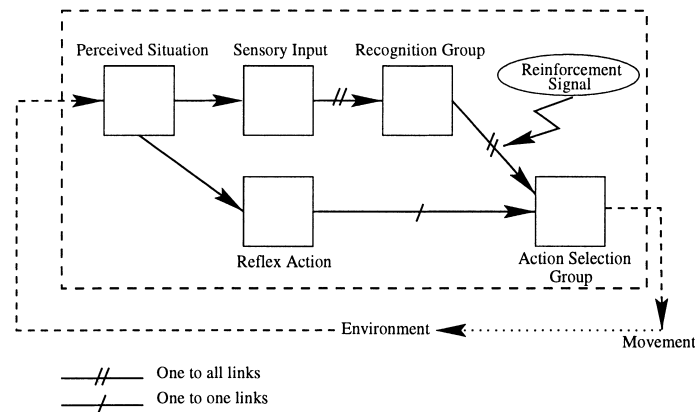


Fig. 1. Schematic representation of the PerAc block. From the perceived situation, the reflex system extracts information to control directly the actions. Concurrently, the recognition system learns sensory input patterns and how to link them to actions by associative or reinforcement learning. The system adapts itself dynamically to the environment.

Box 1. The PerAc architecture

The PerAc (Perception–Action) architecture is a neural computation architecture proposed to solve a wide variety of control problems requiring learning capabilities (in contrast to adaptation capabilities^a). The PerAc block is inspired by various neural networks and robot controllers [1,8,28,29,33]. It consists of an action level (a hardwired pathway able to play the role of a reflex mechanism) and a perception level to recognize particular situations and associate them with the correct action through an associative reinforcement based learning rule (see Fig. 1). The perception level allows the robot to react to a situation by generalization of previously learned situations (landmark configuration in the navigation task) even if it is too complex to allow the action pathway to propose an answer (the goal is not in sight for instance).

Moreover, if the action pathway induces a negative reward, the links between the recognition of the perceived situation and the current robot action can be inhibited and a link with an action avoiding a negative reward can be learned. This kind of architecture is not contradictory with Brooks' subsumption [9] and other robot controller architectures [54] but rather proceeds in the same direction, using a neural formalism (an obstacle overtaking reflex is directly implemented on the low level layer of our architecture for instance).

[11–13,19,32,33] and described in Box 1 and Fig. 1, is used to model place learning and retrieval mechanisms and is efficiently applied to control a real mobile robot.

^a Learning is a process that can be fast and involves a variation in the structure of the control architecture whereas adaptation consists in a slower variation of the structure parameter. Adaptation can induce learning if there is a nonlinear modification of the adapted parameters that induces a nonlinear variation of the system response.

A two stage cascade of PerAc modules operating at different levels of complexity is presented: first to learn patterns and associate them with specific actions, second by the same mechanism to learn locations (places) and also to associate them to a directional motion [31]. This paper emphasizes the second module which nevertheless requires the first pattern recognition module in order to function. This architecture allows an animat to choose and reach a particular goal according to its location and motivation even if the goal is not in

sight. Specifically, the implicit use of the topological properties of Euclidean space, makes our algorithm simple and robust. Successful transposition to robotic experiments reinforces this proposition. Along this, an explanation about “view cell” activity as reported by Rolls [51] in primate studies as a simple restriction of place-cell properties is also provided. In parallel, the relevance of different models is questioned in the light of our success or failure to implement them.

1.1. Visual guidance in robotics

Most of the existing navigation systems use Cartesian coordinates or map representations to locate the relevant and important objects (the robot, the goal, the areas to avoid,...). In well structured environments, the navigation problem consists in finding the best route to go from one place to another. These planning systems suppose the places are already known and usually use improved versions of the classical A* algorithm to find the best route to reach a goal. These systems require an important engineering work to select the pertinent information, to re-calibrate the robot position, to check its current state, or to wait for the recognition of the next state when a reactive planning mechanism is used. In the case of a real autonomous navigation, if the robot forgets to learn a place or learns several times the same physical place, it becomes unable to navigate correctly (cut or infinite loop in the graph of its cognitive map [10]).

In a less structured environment, when the robot does not move through corridors but must operate in a room or in any other “open” environment, potential field techniques [3,36] can be used. For each location, the strength of the goal attraction on the robot is computed. This implies at least to store the goal and the robot locations in a Cartesian referential frame (precise trigonometric computations are needed). Unfortunately, odometry currently used to measure distances is not precise in the long run and must be recalibrated by other sources of information such as particular visual patterns called landmarks [4,16]. Thus in both structured and “open” environments, the actual main problem in implementing really autonomous mobile robots concerns the choice of criteria for the selection of positions to be learned and the regulation of the learning level [18]. Further, the way information

is represented seems to be crucial to reduce algorithmic complexity. Indeed, if the robot had to learn each position in the environment before being able to navigate correctly, learning time would be prohibitive and the system would be unable to perform topological generalization [34].

Today, most of the robots performing visual homing tasks use a conic mirror associated with a classical CCD camera so that they can perceive directly a 360° panoramic view [15,22,23,43]. In the experiment proposed in this paper, the use of this device would avoid the rotation of our CCD camera. However, we do not plan to implement it since the versatility of the visual system of our robot is very important (in different applications, our robot has to be able to focus its gaze on a particular object to recognize them).

1.2. Role of the visual cues in animal navigation

Insects like ants, bees, and wasps are well known to use visual information, in conjunction with a compass,² to return to their nest [14,24,35,56]. It seems that these insects store multiple views taken from different places. The following strategy has been hypothesized: when they leave the nest, they “turn back and look” from different positions (possibly) in order to store new views of the nest and its associated landmarks. This information could be used on their way back to decide which direction could get them closer to a learned view, and to move in a learned direction associated with the best recognized view. How these views are recognized remains unclear since insects possibly use at least two different visual mechanisms to reach their nest (in addition to a path integration mechanism [45]). When they are far away, they seem to rely mainly on the whole surrounding landmarks (*I am here because there is a “tree” in that direction and a particular “rock” in that other direction*). Near the nest this information is not precise enough to allow to find directly the nest entrance (which can be very small: few millimeters). So it becomes necessary to navigate in the direction of an object or a scene directly associated with the goal. In both cases, how these insects succeed in matching learned view and current view remains unclear. They seem to be able to recognize an “object” only if it is perceived exactly under

² They know their orientation according to light polarization.

the same angle. The selection of the movement direction could then be explained for instance by a body rotation translating the image on the insect eye. Interestingly enough insects can learn to associate specific movements with particular sensory situations and thus learn to solve complex maze tasks [68]. Top down visual processes are also used in their brain to allow a kind of attentional focus [69].

Similarly, birds, rats and primates can return to a place by using visual cues and have the same problems as insects (if those cues are shifted for instance). In this case, both insects and mammals using vision to navigate will search the goal at a place which is also shifted from its usual place. Nevertheless, neurobiological studies of mammals reveal that they are able to perform ocular saccades or to use an internal “spot light” to focus their attention on particular local areas of their visual field.³ They can perform the recognition of different subparts of a panorama. The detailed identification of a complex scene is the result of a sequential process even if a global recognition of the scene can be performed more quickly (maybe using lower spatial frequency information). Mammal generalization capabilities could be higher than those of insects (if actually insects are unable to do the same [63], not completely proved today) but not so different in their basic principles. Similarly, mammals can either rely on surrounding landmarks to reach a particular place, or head in the direction of a particular object. A priori, the simple sensory-motor learning which seems to govern insect navigation can also be used by mammals for the same kind of tasks (even if the “hardware” is quite different).

Furthermore, in mammals like rats, a particular brain structure named hippocampus is involved during navigation tasks. Electro-physiological recordings of hippocampus cells performed in CA3 and/or CA1 regions have shown the existence of cells that maximally respond when the animal is at a specific place in its environment [47]. These cells have been named “place cells”. Moreover, it has been shown that the bilateral hippocampal ablation induces learning and navigation problems in new environments (lack of plasticity, i.e. Morris swimming pool [44,65]). More

interestingly, human bilateral ablation of the hippocampus induces an anterograde amnesia with a loss of declarative (in particular, episodic) memory [42] (subjects are unable to learn any new explicit information but preserve the possibility to acquire new implicit knowledge like manual skills, for instance). These results as well as anatomical and histological data lead to the idea that the hippocampus plays an important role not only in navigation tasks but also in a wide variety of spatio-temporal merging problems which are central to our “high level” cognitive ability. The understanding and modelization of such a structure can thus exceed the “simple” application to navigation problems.

In collaboration with a neurobiologist (one of the co-authors) we try to advance in parallel neurobiological modelization and robotic implementation that integrates these models (for a large review of existing models, see [20,59]). The robot can then be viewed as a simulator, at the behavioral level, of the dynamical effects of the model in an environment that we want to be the most realistic possible (environment not prepared for the experiment).

2. Which “representation” for the visual information?

2.1. Visual pattern learning

The first functional step of a visual navigation mechanism is to measure the similarity between learned locations and the current location (e.g. a measure of the distortion between current and stored images). The recognition of an image can be simply performed by a global correlation measure between learned images and the current image. This kind of technique using directly 2D perceived images has the advantage to be very simple and relatively robust (qualitative navigation [38], visual homing [46], visual servoing [39]) but is not really robust to a contraction or dilatation of the image resulting from moving backward or forward, or to objects occlusion or displacement. In mammal and more specifically in primate vision, ocular saccade and/or attention focus play an important role in object recognition capabilities. Hence, by a bottom-up mechanism of “pop-out” attention, brain seems to focus on a particular part of the visual scene, to recog-

³ It would be really time and energy consuming for mammals to have rotate their head and their body in order to see objects always in the learned position on the retina!

nize it before focusing again on a new area, and so on. Moreover, brain imaging techniques as well as lesion studies have shown that after the first occipital visual processing steps involving gradient extraction, end-stop and corner detection among others, visual information flows through two distinct pathways devoted to “what” (inferior temporal cortex) and “where” (posterior parietal cortex) processing [60].

In our approach, the image is divided into sub-areas centered around focal features and the recognition level will only depend on the correct recognition of these features and the surrounding area, weighted by the displacement of these features from the learned configuration.

We argue that such a mechanism allows a more robust scene recognition than classical global correlation (without feature extraction) because the recognition level will only depend on the correct recognition of the learned sub-areas around the selected features and on their relative displacement with respect to the learned image. If the learned snapshots are taken around points that remain stable when the robot is moving then the recognition will be very robust. In our case, the visual system focuses on corners and/or end of lines (end-stop cells). The process consists in a gradient extraction followed by a convolution with an OFF-center difference of gaussians that allows to detect corners at a particular spatial resolution [25,31] (see Fig. 2).

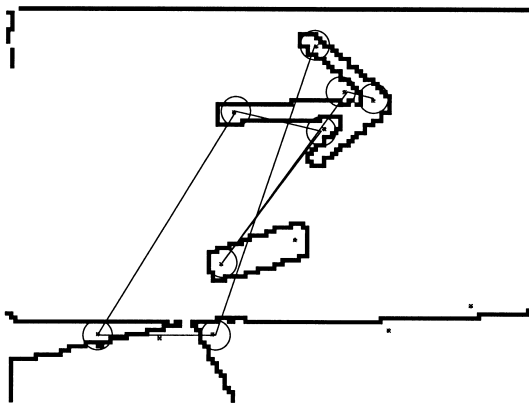


Fig. 2. This figure illustrates how the attentional focus of the robot visual system performs the exploration of a scene. The circles represent the center of the focalization area. The vision field spreads over about 50 pixels around the focal point.

For instance, if there is a door in the image, the system may focus on one of the door corners and will learn the local view around that corner. Then, if the robot moves forward in the direction of the door, it will find again the same feature points and the same local views (the local shape of the corner will not change⁴). The only difference will be the relative position of the different views (affecting the motor command that must be produced by the visual system to move from one focal point to the other). This kind of mechanism can be used to compute the apparent size of the objects as well as their azimuth.

In a first approach, azimuth appears more informative than size to locate objects. Indeed, information about distance obtained from apparent size varies in a very nonlinear manner and does not allow to estimate the distance with a good precision if the agent is far away from the object. The measure of the apparent size also requires two points on the same vertical structure at different heights. The chances of mistakes in performing the right associations make this kind of measure difficult to realize in a natural environment. Information about the apparent height of the floor in a plane environment could be easy to use but supposes no ambiguity in separating the floor from the walls (not granted if there is a texture on the floor or a lot of objects everywhere). Moreover, the only use of apparent size information does not allow to differentiate mirror situations [70]: if a place cell response is determined by the distances to two objects A and B, then the answer will be the same if A and B are swapped or if the scene is seen through a mirror. Of course, in a complete system, one will have to integrate both azimuth and apparent size information of the landmarks (optical flow and/or stereo-vision could also be an important information to estimate distances).

2.2. Robotic implementation of distant visual landmark learning

Most of our robotic experiments were done in a room of about 7.2 m × 5.4 m shown in Fig. 13 (the size of our Koala robot is about 30 cm × 30 cm). The visual input comes from a 384 × 288 gray-scale CCD

⁴ This is only true if the images are at the same resolution. If the second image is taken at a few centimeters from the door surface then completely new things will appear... .

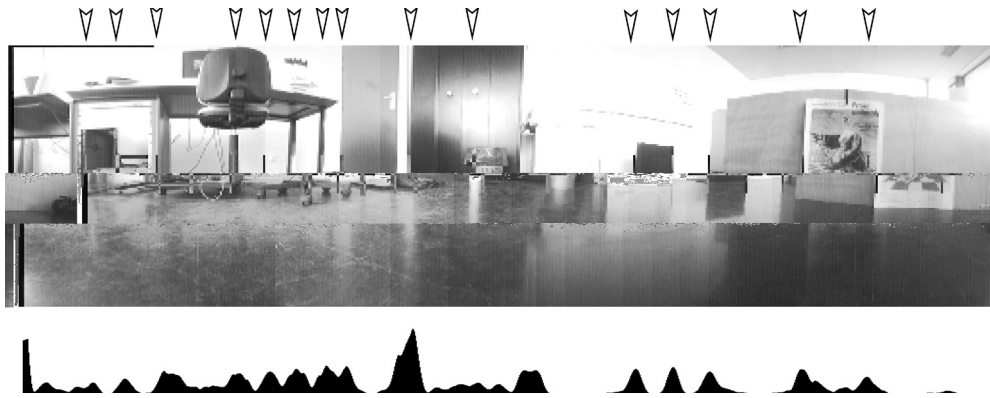


Fig. 3. Azimuth of the 15 local views stored and used to learn a place form a given panorama (environment Fig. 13, black circle S). The graph below the panorama represents the absolute value of the derivative of the signal obtained by averaging the image columns.

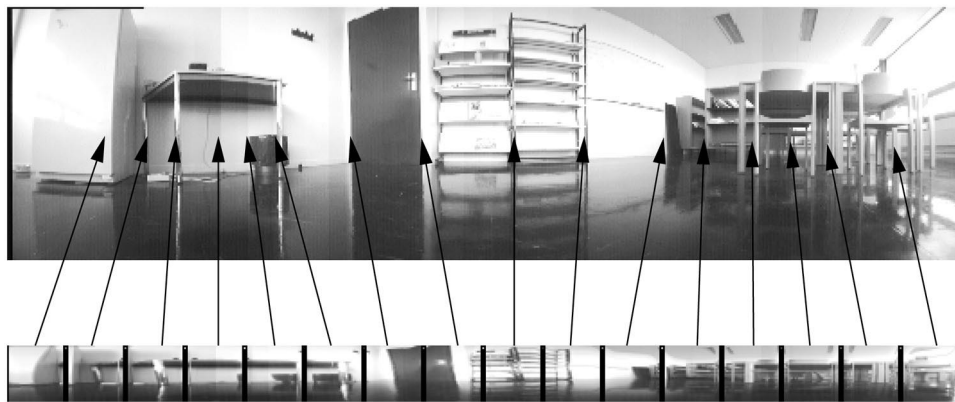


Fig. 4. 16 examples of 32×32 local views learned as landmarks and their azimuth in the panoramic image (panorama referenced as *image a*) in Fig. 13).

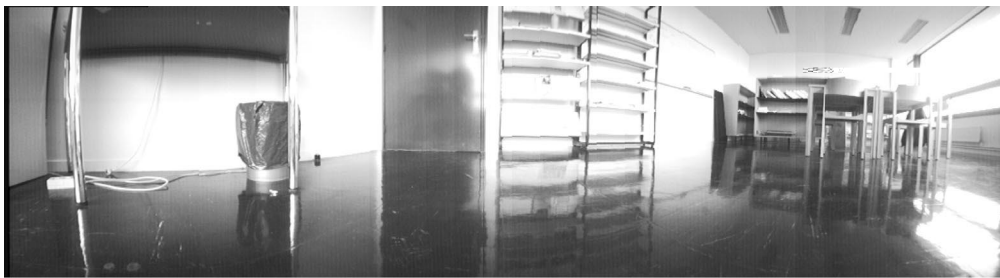


Fig. 5. Another panoramic image constructed by our system referenced as *image b*) in Fig. 13 that our system must be able to discriminate from the image Fig. 4, for instance.

camera. Its field of vision spans about 70° . To build a panoramic view a servo-motor is used to pan the camera. The robot takes 24 images with a 7.5° rotation between each image acquisition. The central vertical bands of each image (30×288 pixels) are merged to constitute the global panoramic view (only the central band is used because the camera distorts the image sides). A 1246×288 pixels panorama is obtained. The resulting field of view is about 250° . It is not a complete 360° image (there is a blind area) but it is enough in practice. As you can see on the different panoramic images (see Fig. 3, for instance) the images merger is not perfect but it works (which shows the robustness of our system).

The first step of the algorithm consists in locating the possible landmarks. To reduce the computation time, a very simplified version of the feature point detector described previously is used to control the focus of attention. All the image columns are averaged and weighted with larger weights for the points near the center of a column. The resulting one dimensional signal is differentiated and the local maxima and minima are used as centers of local views (an example is shown in Fig. 3). Each panorama projection contains, in average, 20 local maxima. In Fig. 4, 15 maxima are selected to be the centers of 15 local views. Another panoramic image constructed by our system is shown in Fig. 5.

For each selected focus point, a 32×32 pixels local view is built by averaging the 148×288 pixels of the corresponding panoramic image part. The y axis is just scaled whereas a logarithmic transformation is used for the x axis (no need of a complete log-polar transformation since there will not be object rotation problems in the camera plane). Then, each current local view is compared with each learned local view. This comparison is made by the norm of the difference between the pixels of the two local views. The best corresponding views are used as landmarks, i.e. their positions in the image are compared with the ones in the learned panorama. The sum of these absolute angles provides us the similarity measure between panoramas. Then, the movement associated during learning with the best corresponding panoramic image is performed and allows the robot to reach the goal (obviously all the angles are rotated according to the robot body orientation).

2.3. Building a robust representation of a place

In our system, the information about the landmark recognition and their azimuths are merged to produce a unified representation that can be easily learned and matched with previously learned representations. This merging process is implemented as the vector product of the information corresponding to “what” (Landmark) and “where” (Azimuth) the objects are (Fig. 10, Landmarks Azimuths group). A time integration process allows to suppress the sequential aspect of the scene exploration (spatio-temporal merging). In this representation, there is no need to “recognize” specifically what the landmarks are (a fridge, a chair, etc.). To discriminate them and to record their angular position only matters. Even if a landmark is missing, the other landmarks can allow a correct recognition. We will show in the section devoted to the experimental results how several landmarks can be moved, removed or hidden without altering the global recognition of the scene. In the mammal brain, this flexible merging by opposition to the static recognition of a multi-sensor configuration seems to be performed by the hippocampus⁵ [5]. An explicit neural network representing this kind of treatments is proposed in Fig. 6.

The activity of the “place cells” P_i when the robot is at the location (x, y) can be summarized by the following equation:

$$P_i(x, y) = 1 - \frac{\sum_{k=1}^{N_i} V_{i,k} \cdot f(|\Theta_{i,k} - \theta_k(x, y)|, v_k(x, y))}{\pi N_i} \quad (1)$$

An intuitive understanding of this measure can be seen in Fig. 7, where the negative term of the equation is plotted according to the robot position.

In this equation, N_i is the number of visible landmarks when the robot is at the learned place i (field corresponding to P_i). $\Theta_{i,k}$ represents the learned value

⁵ At least in primates, there is ample neurobiological evidence that this “what” and “where” merging takes place either in the perirhinal, parahippocampal, entorhinal cortices and/or in the hippocampus proper (CA3) [52, pp. 107–108]. In fact, the types of multimodal associations taking place in the perirhinal cortices and hippocampus proper could be different both in nature (rigid versus flexible) and extend (more local in the cortices, more global at CA3 level).

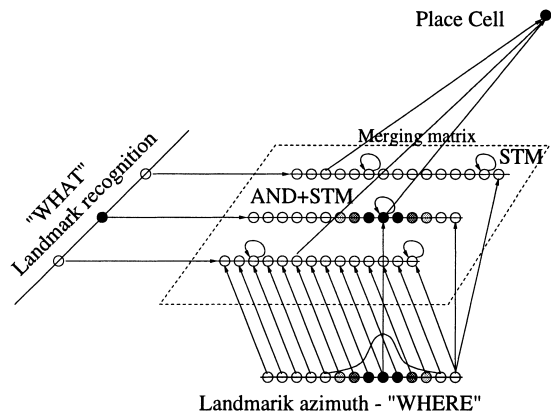


Fig. 6. Part of the neural network devoted to the merging of what and where information. The lateral diffusion allows to measure the difference between the learned azimuth and the current azimuth.

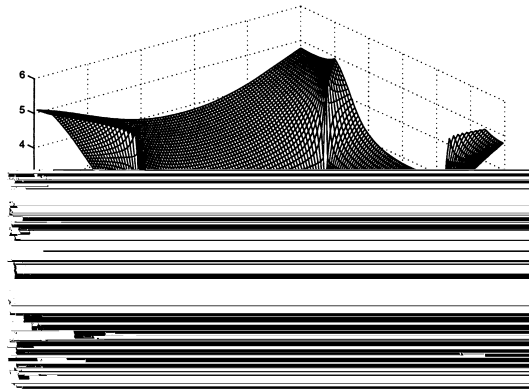


Fig. 7. Simulation of the error of a “place cell” response computed according to the negative part of Eq. (1). In that example, the learned location is associated with the point (50,50). Four recognizable landmarks are located at positions (20,20), (20,90), (90,20) and (90,90). The shape can be seen as an attraction basin centered on the learned location.

of the landmark k azimuth from the learned place i . θ_k is the value of the same landmark azimuth for the current robot location (x, y) . All the angles are expressed in radians and measured from an absolute direction (the north, for instance). $|\Theta_{i,k} - \theta_k(x, y)|$ is computed modulo π , $V_{i,k}$ is set to 1 when the landmark k is seen from the learned location i and 0 otherwise (the same rule applies to v_k for the current location) for the current robot location). When learned landmarks are not recognized, we can have $V_{i,k} = 1$

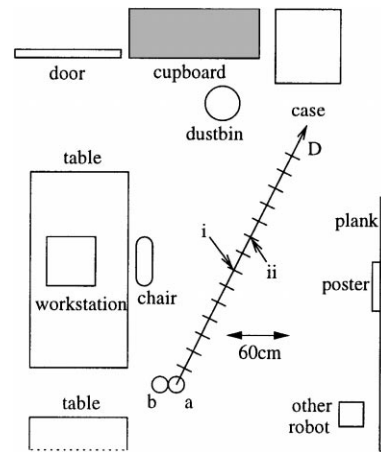


Fig. 8. A second room for the experiment, the white circles (a and b) and bars on line D represent several tested positions.

and $v_k = 0$. f is a nonlinear function to account for landmark recognition:

$$f(\theta, v_k) = \begin{cases} \theta & \text{if } v_k = 1, \\ \pi & \text{if } v_k = 0. \end{cases}$$

The error associated with a landmark azimuth is maximum when the landmark cannot be found (landmark not visible for instance: $f(\theta, 0) = \pi$). Eq. (1) gives a growing activity P_i that tends to 1 when the azimuths θ_k associated with the current location are close to the stored $\Theta_{i,k}$.

Experimentally, we first tested one robot “place cell” response for different orientations (36 tests with 10 degrees rotations) at a learned location and we verified the activation of the place cell was always the same up to a 1% error, in spite of landmarks lost in the blind area of the robot.

Next, the robot was moved along a rectilinear trajectory (line D in Fig. 8) in order to evaluate precisely the variations of the activities of two place cells learned at two neighbor locations. The maximum activity of the place cells is always obtained for the learned location and monotonously decreases with distance (see Fig. 9). These results emphasize the fact that even with large distances (compared to the robot size and to the environment size), the robot is able to perform an action in order to come closer to the goal.

In addition, the more complex the environment is, the more numerous landmarks can be, and thus, the more efficient our algorithm is.

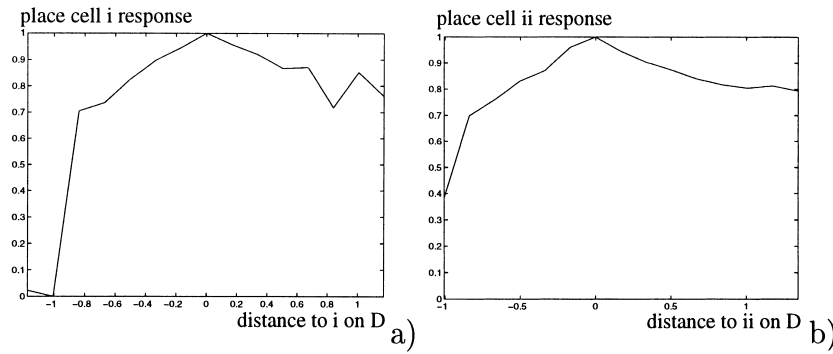


Fig. 9. Response of a place cell (learned view) according to their learned location on a rectiligne trajectory (see environment, Fig. 8). The asymmetry of the field results from the views differences before and after reaching the learned location and beyond some distance the views cannot be further used (generalization failure).

3. Learning of an homing behavior

In a formal perspective, returning to a given goal location X_i (homing task for instance) can be written as the minimization (or the maximization⁶) of a particular potential function $\psi_i(X)$ such as

$$\psi_i(X) = P(X_i, X) \quad (2)$$

with X the vector describing the current location and P the function used in the previous section to compute the place cell activity for instance.⁷ The potential equation must verify:

$$X_i = \text{Arg} \min_X \psi_i(X) \text{ or } \psi(X_i) = \min_X \psi_i(X). \quad (3)$$

Most of the navigation techniques use some kind of measure of their distance to a goal to decide of their movement. Gradient descent techniques are then a natural solution as attested by the success of potential field techniques proposed by Khatib [36] and Arkin [3]. In this paper, we will not consider these methods in relation to the use of an explicit field (i.e. the field has to be computed for each location; more importantly robot and goal position are supposed to be known in the same absolute coordinates). We will rather suppose the existence of an implicit potential field determined

by $P_i(x, y)$, the activity of the place cell i in Eq. (1) (see also Fig. 7). In this frame, the action vector A_i to reach the goal X_i can be expressed as the result of the spatial derivative of the potential function $\psi()$ for the current X location:

$$A_i(X) = \begin{pmatrix} \frac{\partial \psi_i(X)}{\partial x} \\ \frac{\partial \psi_i(X)}{\partial y} \end{pmatrix}.$$

Even without the requirement of building an explicit field, the gradient can be very flat when the system is far away from the goal. The system should therefore move a lot in one direction to be able to compute the value of the gradient in that direction. It would next have to return to the current position to try another direction and so on. Obviously, on a real robot, because odometry is imprecise and the measure of activity from all these locations is time consuming (a lot of unproductive movements must be performed), the direct implementation of this algorithm to approximate the gradient and to follow it to reach the goal is not straightforward. Nevertheless, this strategy could be used by rats in maze problems when they lack information (vicarious trial and error strategy proposed by Tolman [57,58] and used by Schmajuck for the modelization of maze learning [53]). Another solution inspired from bacteria strategy to reach a sugar area could be to keep on moving in the same direction until the gradient becomes negative. At this time, it is enough to choose randomly another direction and to repeat the

⁶ In the case of a “place cell” like measure of a place recognition, the activation of the cell is maximum when the animal is at the correct place. Returning to that place is equivalent to maximizing that activation or to minimize the error in the place recognition.

⁷ In that case $X_i = (\theta_{i,0}, \dots, \theta_{i,k}, \dots, \theta_{i,N_i}, V_{i,0}, \dots, V_{i,k}, \dots, V_{i,N_i})^T$ and $X = (\theta_0, \dots, \theta_k, \dots, \theta_{N_i}, v_0, \dots, v_k, \dots, v_{N_i})^T$.

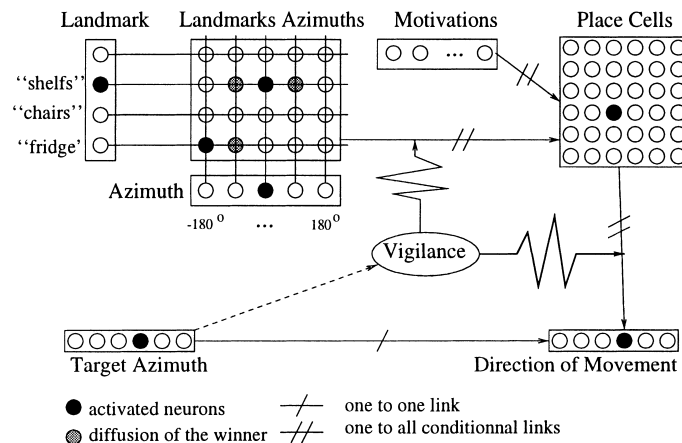


Fig. 10. The navigation neural network. Target azimuth and Direction of Movement groups are WTA. The Landmark-azimuths group emphasizes merging of visual and directional information.

process. At the end, the robot must reach the goal location but it is obviously not a very efficient strategy. Moreover this strategy may become very inefficient when obstacles appear. The robot could get stacked in local minima. A solution that looks like stochastic gradient descent with momentum could also be used and could be a good complement to the previous algorithm. An important number of other approaches also use specific algorithms that compute the best movement to perform in order to minimize the error of the landmark position in the visual field [14,22]. These algorithms have mainly been tested in simplified environments where there is no recognition problem (use of black cylinder landmarks in a white environment [43]) or where there are no segregation problem between the ground and the landmarks (toy houses on an homogeneous ground [22]). Recent works try to overcome these problems by using visual learning based on local correlation techniques and on the selection of the most relevant or stable landmarks when the robot is moving in the neighborhood of the goal [6,50]. Yet, because the movement direction is computed in an ad hoc manner only linked to the displacement of the perceived landmarks, those models do not explain how an animal could also use the recognition of a particular place to decide to move in a particular direction corresponding to a place different from the learned place. Moreover those algorithms do not allow to learn how to reach several goal locations (lack of versatility of the algorithm).

To avoid these problems, it looks simpler to learn not the goal location but rather places in its neighborhood and to associate these places with the direction of the movement to perform to reach the goal (straightforward application of the PerAc architecture — see Fig. 10 and Appendix A for a mathematical formulation).

These positions must be sufficiently remote from the goal to obtain different place cell activities according to the different learned locations. Moreover, they must be close enough to the goal to allow the robot to associate these locations with the correct movement direction to reach the goal. This can be done with the help of the odometry which does not need to be very precise or by visual tracking of the goal location (to maintain active the goal direction on the motor map — for a neural description see [31]). In practice, we have shown that learning 3–6 places at 30 cm from the goal is sufficient to join it from distances higher than 3 m (the panoramic image size is about 1246×288 pixels — a better resolution than insects). For these remote (not learned) locations the movement decision depends on the best recognized place or on a combination of the most active place cells [26] (see Fig. 11).

Interestingly, the simplest way to foster this kind of learning is to mimic the exploration strategy used by insects like ants or wasps when they leave their nest ("turn-back-and-look" in the direction of the nest from several locations [17,66,67]). At the beginning of the exploration phase, we suppose our robot (Prometheus)

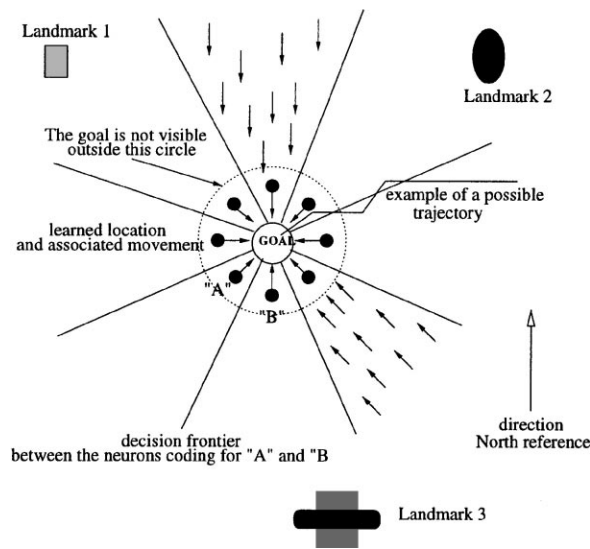


Fig. 11. Example of a goal retrieval with an absolute direction. Each “place cell” wins for a particular angular sector. The robot position during each place learning is represented by a black circle. The arrows associated to these circles represent the action to be performed from that place in order to reach the goal location. Competition between these place–action associations allows the robot to reach the goal even if the goal is not in sight.

moves randomly, looking for something interesting. The first time it finds a particular goal location, an appropriate sensor allows to detect this success. A positive reinforcement signal is emitted and the vigilance level for on-line, autonomous learning is increased. As a consequence, learning of “new” local views and places are facilitated. This signal also triggers the learning procedure which consists in moving backward for a particular distance (30 cm), taking from that place a panorama while the body orientation is in the goal direction. Then, the robot moves forward on the same distance than before and returns on the goal (see Fig. 12).

Classical Hebbian learning allows the robot to associate the recognition of that place with the fact that moving forward in the direction specified by the compass allows to reach the goal. When the robot is once again on the goal location, it can, for instance, rotate on itself of 60° , move backward again on 30 cm, take a new panorama, analyze it, return on the goal and repeat that until enough perception–action associations are learned. If there is an obstacle while moving back-

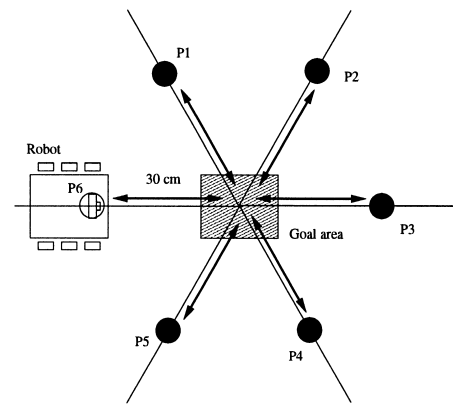


Fig. 12. Procedure to learn place–action associations around a new goal. The robot moves backward on 30 cm, takes a panorama, analyses it, returns on the goal, rotates 60° , moves backward again, takes a second panorama,....

ward (detected by the infra-red sensors in the back of the robot), the robot simply comes backs to the goal and tries to go to the next place to be learned.

The decision to learn a new place does not need to be directly controlled by the learning behavior (see [27] for a simulation of how our neural network can be used to build autonomously a cognitive map of the environment). If the activation of the learned place cells is too low (inferior to the vigilance threshold) the current place is considered as new and has to be learned. If the variations of the perceived situations are important on a small distance, several place–action couples can be learned whereas for the same learning trajectory in a poorer environment very few places will be learned (in that case, the detection of the number of learned places could be used to carry on the learning behavior on more important distances). After a while, an habituation mechanism allows the robot to cease to be interested in the goal and then to stop the learning behavior.

Because the robot may have an orientation different from that used during learning, all the azimuths are shifted of an angle corresponding to the angle between the robot and an absolute direction. In our case, an electronic compass⁸ mounted on the top of the robot allows to obtain an indication about the north

⁸ Pawatron sensor model 6070.

direction⁹ (precision about 3–4 degrees with a bias which depends on the proximity of metallic objects). The inverse shifting mechanism is applied to the motor output, by subtracting the same angle. Thus, the learned movements do not depend on the robot orientation and are correct whatever the robot orientation is (preservation of the sensory topology in the whole neural network architecture).

When the robot wants to return to the goal, there are two possibilities. If the robot recognizes the goal, it moves in its direction (reflex link between Target Azimuth and Direction of Movement in Fig. 10, for a detailed description, see [34]). Otherwise, it considers the information of the place cells (i.e. the cells which react to a specific set of local views associated with their azimuth) and moves in the direction related to the most active place cell (competitive mechanism) to reach the goal (we will only address this second issue in this paper). Thus at each step, the distance to the target is reduced (Fig. 11) and the robot returns inevitably to the learned position [31]. The PerAc architecture for place learning realizes an approximation of an attraction field [3] avoiding the cost of either learning what to do at each position in the environment or knowing goal and robot location in a Cartesian space. Moreover, it has been mathematically proved that no local minima were induced by the competition between the action neurons within the domain bounded by the set of landmarks [71].

3.1. Importance of using a soft competition mechanism

In the first robotic experiments, a strict Winner take all (WTA) mechanism (that had to associate a unique label to each perceived snapshot) was used. The results were not very good and sometimes the neural activity patterns associated with different views were very close because some of these learned views were very similar. This problem increases with the number of learned views. The explanation is quite simple.

⁹ A directional information linked to the robot odometry or the perception of the different robot accelerations (like in the mammal vestibular system) could also be used to build a dead-reckoning mechanism [21,63]. We have also shown in simulation the possibility to use a distant landmark as a reference direction (the algorithm works correctly if the robot remains in the area surrounded by the landmarks, see [31]).

For two different places to be learned, the robot can have the same object in its visual field. According to the vigilance threshold, the two different views of the same object A, can be just different enough to be linked to two different labels and coded on two different neurons. When the robot is between the two previously learned places, it is not sensible to attribute only one of the two labels to the view linked to the object A, because the choice of the associated label will introduce an important bias in the place recognition. For instance, the view can be more similar to the first place but can have an azimuth that matches better the recognition of the second place. The interest of the nonlinear effect of taking local snapshots to build a nonlinear measure of panorama similarity was becoming in these cases an important hindrance.

Yet, in biological competitive structures almost never a single winner emerges. The WTA is more like a contrast enhancement mechanism. This process was implemented in our system by using a smooth WTA mechanism. The activities of all the neurons which are below a given threshold are set to zero and the activities of the others are left unchanged (a lot of other solutions are, of course, possible — see the section about motor control for an example of a more efficient solution). In this system, the same “object” can have several interpretations. The system avoids to take decisions too early in the processing chain. It does not take a hard decision either at the view recognition level, or at the place recognition level. Only the competition between the neurons associated with the different places allows the robot to move correctly. It never “recognizes” correctly a given place (in the classical pattern recognition sense) but it succeeds to reach the center of the attractor created by learning the place–action associations. Indeed, in the navigation task, the system’s actual recognition of a given place does not matter. What only matters is the place field of the winner neurons in the place recognition group being the closest possible to the robot current location. Hence, when one or several landmarks are randomly moved, the system keeps on working correctly. The system only needs that, at least, two landmarks remain stable on the 20–30 landmarks found in the image. Of course, the activities of the place cells will decrease if landmarks are moved randomly or if distractors are introduced but the outcome of the competition will not be changed (all the place cell

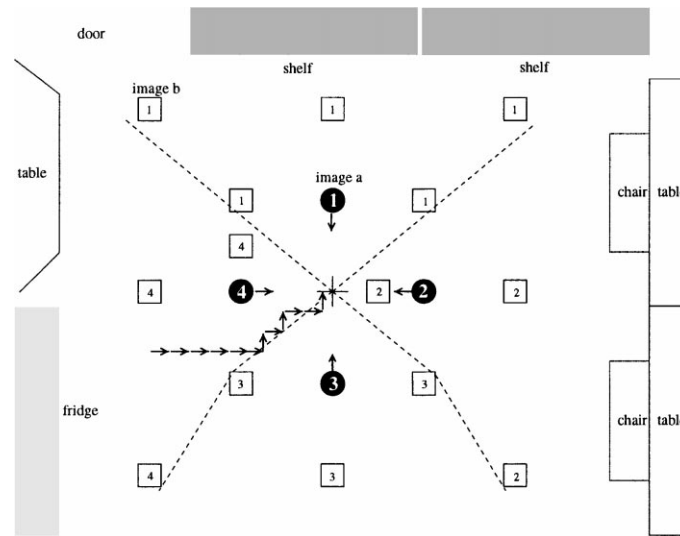


Fig. 13. Place cell-like segregation of space in a robotic experiment. Four views are learned (circles), others are associated with one of these learned views (boxes). As we can see, if the robot learns to reach the cross from each learned view it can reach the cross from any other view associated with the learned one (generalization). The set of arrows represents a possible path. The views are in a 1.2 m x 1.2 m area, the learned views are at 30 cm from the center. The scale is not respected for the position of the different furnitures (in fact they are about 1.5 m from the center).

activities will be reduced by almost the same amount). Problems can only arise if a majority of landmarks are moved in a coherent manner (say rotated around the room of the same angle) or if the stable landmarks for the recognition of a given place are in the blind area of the device used to take the panorama. In a normal case, just the system’s robustness to noise will be decreased (in new experiments we try to measure this kind of signal to noise ratio in order to obtain a better characterization of the system performances).

3.2. Robotic experiments of a simple visual homing

The test environment is first presented Fig. 13. The results of the robotics experiments on simple visual homing show the activities of the four learned place cells according to the robot location.

The variation of the place cell activity as a function of the robot location is emphasized (Fig. 14). A combination of the four responses of the place cells (which is the actual measure used to decide the direction to move) is presented in Fig. 15. These responses allow to test the correct generalization capabilities of our place cells to large open areas.

Fig. 15 presents a combination of the four responses of the place cells recorded during a first experiment (maximum of response at each location). It verifies that our “place cells” can generalize correctly to large open areas. The robot reaches the center of the goal area (attractor) and stays in its vicinity with a precision of 5 cm (each move was about 20 cm and the robot size is about 30 cm). Even if the starting point of the robot can be very far from the goal, the robot is able to reach the goal with great precision (the precision can be increased with smaller steps).

The correct goal reaching behavior (see Fig. 16) is induced by the action sequences linked to the recognition level of the different learned places. Fig. 16 also shows the movement precision can be better if the size of each move is reduced to about 4.5 cm. Here, the smallest distance of the robot to the goal is less than 2 cm. In theory, with landmarks at distance d , the precision p representing a 1 pixel shift in the image is $p = \tan \frac{1}{4}d$. It comes from the field of vision being about 270° , the x axis made of 1066 pixels, that each pixel represents about $\frac{1}{4}$ degree. Thus, with landmarks at 1.5 m, the maximum precision is 7 mm and with landmarks at 15 m it would be 7 cm. The starting

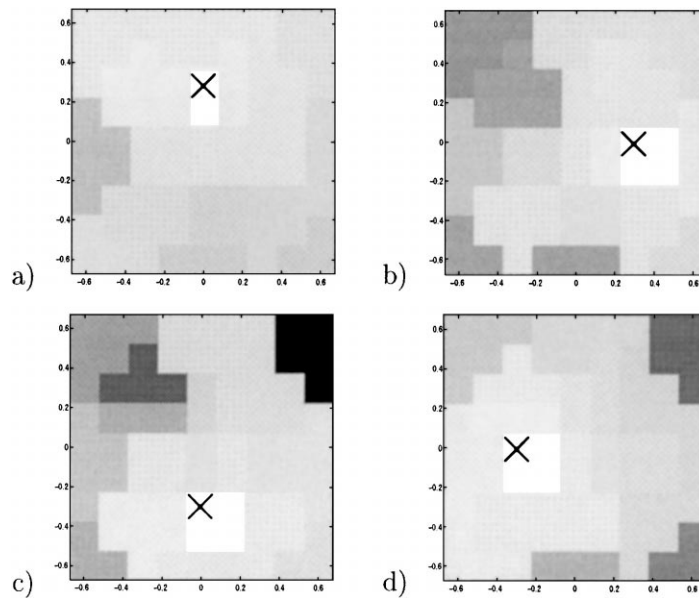


Fig. 14. Place field response of each place cell when covering a $1.2 \text{ m} \times 1.2 \text{ m}$ square (environment Fig. 13), white color represents high place recognition. The cross represents the position of the robot when learning the scene, the goal position is (0,0).

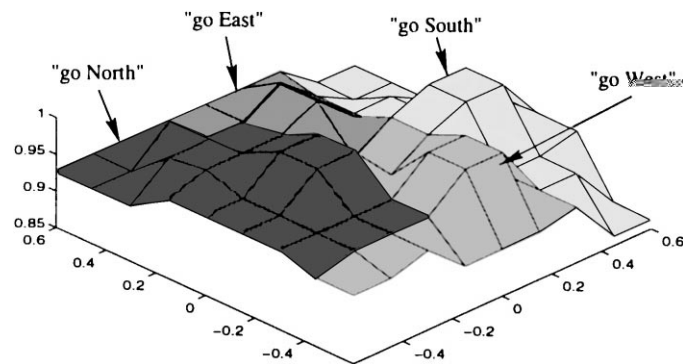


Fig. 15. Combination of the four place cells activities (see Fig. 14). For each point in space the maximum activity of the four cells has been taken. The associated experimental area is shown Fig. 13. Unit is meter and a gray level is associated with each direction (verification on Fig. 13).

point of the robot can be very remote from the goal. The robot is still able to reach the goal with great precision.

An important limitation in the first implementation of the neural network was the coarseness of the movements. So, we combine the different responses of each “place cell” as a function of their activity. We have to apply a transformation on each activity to enhance dif-

ferences because activities are very close. The contrast enhancement of the movement direction is performed as follows: $S'_i = (S_i/S_{\max})^n$ and $\vec{V} = \sum_i S'_i \cdot \vec{v}_i$ where S_i is the value of the place cell i , \vec{v}_i its associated movement, S_{\max} the value of the most activated place cell and \vec{V} the robot movement vector. n is chosen high enough to avoid that place cells with a low activation (place cells coding for a distant

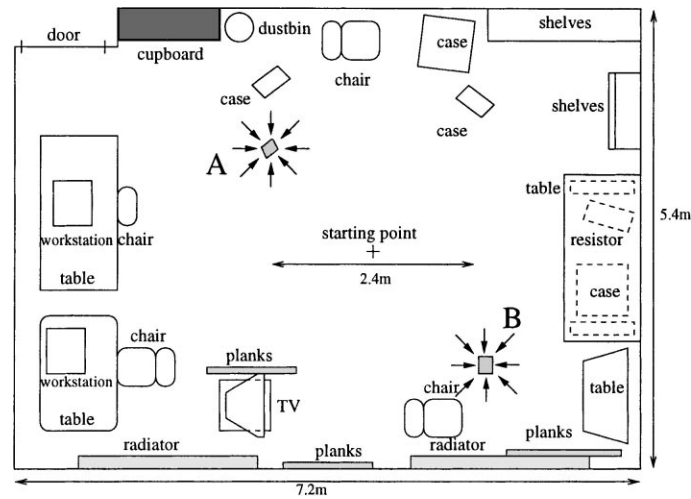


Fig. 18. Map of a room used for simple motivated navigation. We can see the two goals A and B. The 3 arrows around each goal represent the learned positions and the associated movements. The distance between each learned position and the center of the goal is only 30 cm.



Fig. 19. Panoramic image constructed by our system. We can see the experimental room with the two goals A and B (see plan Fig. 18).

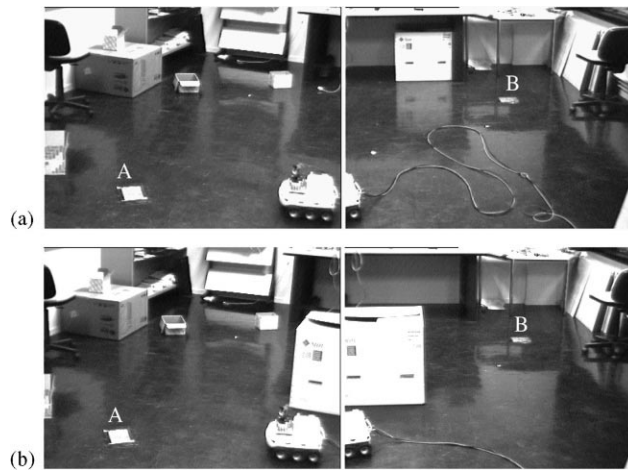


Fig. 20. The experimental room, we can see the robot and the two goals A and B. On the image (b) we can see the introduced obstacle which is a landmark moved in the center of the experimental area.

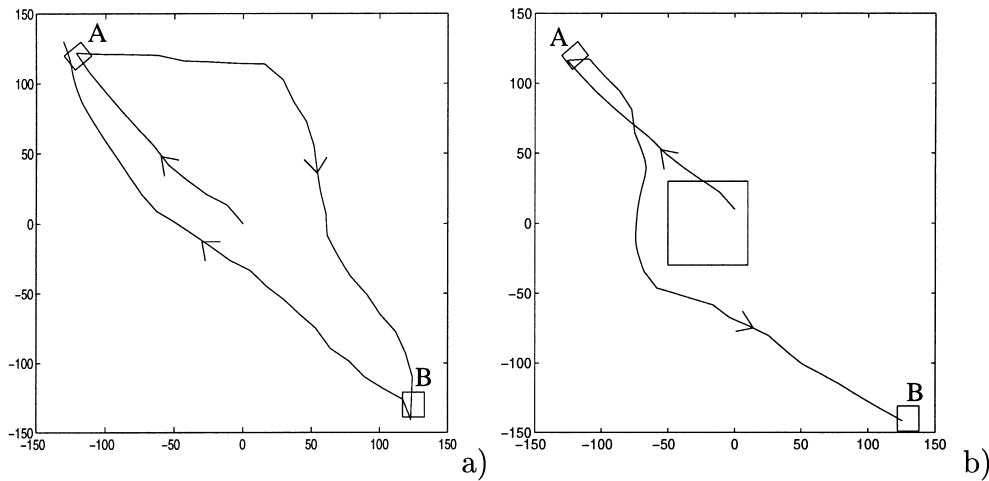


Fig. 21. (a) Path followed by the robot attracted by two kinds of supply locations (linked with two different motivations). The robot starts at the center of the room and is more motivated by goal A than B. When it reaches A the corresponding drive decreases and the robot is attracted by B. Finally at B the drive A becomes preponderant once again. (b) In this example, the robot starts from the center of the room without obstacle. Then, when it reaches the goal A associated with the most activated drive, an obstacle is put at the center (the large box) and the robot has to reach the other goal B.

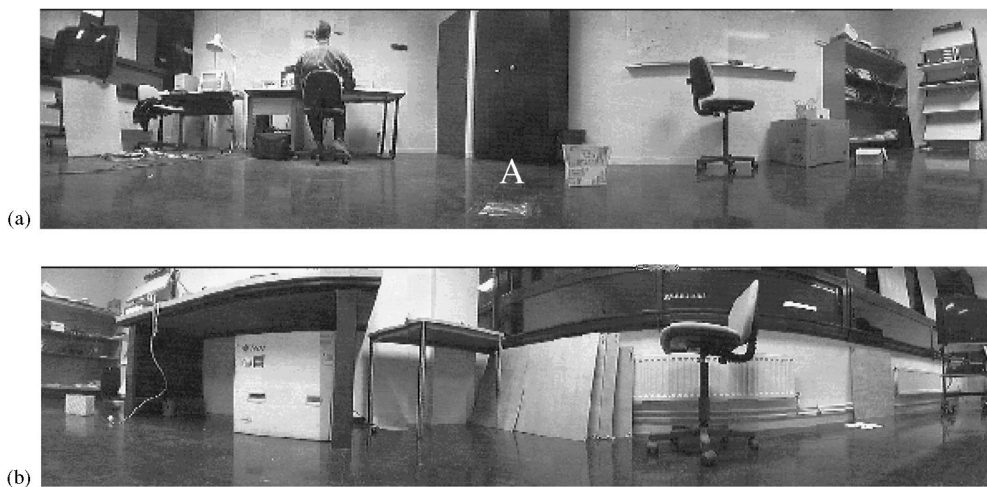


Fig. 22. The robot is close to a goal. In image (a) the robot is about 80cm from the goal A, we can see it on the panoramic view. In image (b) the robot is about 40cm from the goal B, it is too close, we can not see the goal.

the CCD camera of the robot (Fig. 22) show that those targets are too small to be used by our algorithm and allow us to say that our system is blind to the goal.

The first time the robot finds a particular goal location (detection of the color on the floor by a specific sensor) it triggers the learning procedure (see previ-

ous section). When the robot is on a goal, the corresponding drive is resolved (its intensity decreases) and the robot is attracted by the other goal. Several experiments were undertaken on the robot (see Fig. 21), with motivations evolving with time. As we can see the robot is always able to reach the goals, even if the

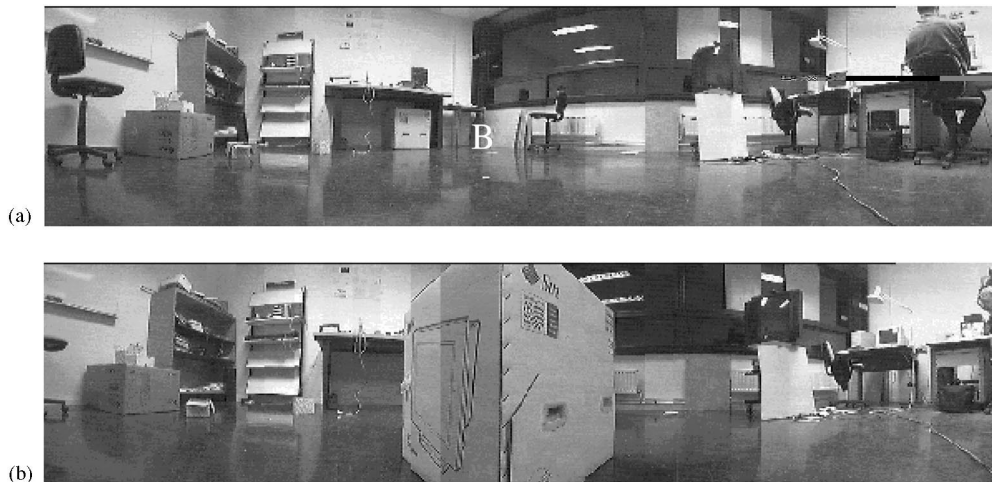


Fig. 23. The robot is close to the goal. In image (a) the robot is about 80 cm from the goal A, we can see it on the panoramic view. In image (b) the robot is near the introduced obstacle which occludes some of the learned landmarks.

size of the goal is small and if the robot is far from it. In fact, in these experiments, goals are two rectangles ($16\text{ cm} \times 14\text{ cm}$ and $16\text{ cm} \times 20\text{ cm}$) and the distance between goals is 3.5 m. It is important to compare this distance with the distance between learned positions and goals which is only 30 cm. Fig. 21(a) and Fig. 21(b) represent the robot trajectories, goal positions (A left-up and B right-down) and the position of an obstacle (not present at that place during learning). The vision and the compass are used to reach the goal while the infra-red sensors are used to avoid obstacles. Though the goal can be visible (the metallic paper on the black floor) the robot does not use this information. As we can see in Fig. 23(a) when the robot is far from the goal (3.5 m) this one is invisible. Moreover the shortest distance from where the goal can be perceived is 80 cm (Fig. 22(a)), closer (40 cm Fig. 22(b)) the goal is no more visible due to the camera position on the robot body (we have verified that it has not been learned as a landmark).

To verify that our system actually does not need to see the goal and can integrate goal seeking and obstacle avoidance behaviors, we introduced a huge obstacle (a $60\text{ cm} \times 60\text{ cm} \times 60\text{ cm}$ box) in the center of the experimental area. Those experiments are shown in Figs. 21 and 23, we can see that the goal is totally occluded by the obstacle. The learned positions are always the same, the robot starts without obstacle and

when it reaches the first goal we move the obstacle at the center of the room. There, the robot uses an obstacle overcoming strategy combined with our navigation model (like a Braitenberg vehicle [7]). In fact the robot quits the goal reaching strategy and triggers an obstacle following strategy until it succeeds in retrieving the desired orientation, given by the goal reaching strategy. Hence we avoid the local minima problems that appear when two contradictory behaviors must be performed at the same time. Figs. 21(a) and (b) show that trajectories do not go to the goal directly but always reach it. Several other experiments with different starting points and starting orientations have been successfully completed. For the robot to succeed from any position in the room it just has to learn some positions around each goal (in fact 6 positions around each goal are enough to provide good performances but 3 positions can be sufficient).

The introduction of “motivation” neurons associated with the place cells allows the robot to bias competition between potential goals (Fig. 17) and then to reach the desired goal [72] if the goal belongs to the same visual environment than the current robot location. Conversely to classical potential field techniques [3] there is no need to know the location of the robot and the goal in an explicit and common referential.

The presented model does neither compute in which direction to go according to the perceived and the

learned panoramas nor learn what to do for each location in the environment. The information representation allows the system to use, at distance from the goal, the movement direction learned from places near the goal (a priori generalization). Our simulations and robot experiments show that if we can separately recognize at least two landmarks (local snapshots) in a panoramic image and know what their azimuths are (relative to an absolute direction given by a compass or by the odometry) then, this information can be recombined in a very robust manner to decide whether the robot is far or near the learned location.

Moreover, the complexity of the algorithm for a panorama analysis is very low. The system has to perform 14 million integer additions and 1 million floating point multiplications. More than 95% of the calculation time is dedicated to the creation of the local views, the remainder is spent in view comparison. The total calculation time is less than 1 second on a Pentium 133 and could be easily reduced. In fact, when the robot has learned 10 locations (2 learned goals), it performs a movement every 15 seconds. About 13 of those 15 seconds are spent in the acquisition of the panoramic image (camera rotation).

5. Synthesis and conclusion

The experiments presented in this paper show the robustness of our algorithm and its adaptability to a wide variety of tasks. It tolerates a lack of landmarks or a misinterpretation of a few of them. There is no need of a particular number of landmarks (more than 2) to be learned or present in the visual scene. The precision of place recognition will only grow with the number of landmarks. These properties support the idea that animals do not need to learn a large number of positions in an open environment to be able to return to a goal location. Moreover, there is no need of an internal allocentric representation of the world, like a Cartesian map. Instead of learning a map of the environment, it is possible in simple but frequent situations (free open environment) to learn to build an attractor around a desired location.

Whatever the interest of those visual navigation models, one must not forget they cannot account for navigation in a complex environment. Obviously, if the obstacles are convex, it is possible to build a

repulsive field with a limited distance of activity which functions exactly in an opposite manner to the place attraction behavior. In the same way, selection of a goal in such open environment is also easy to model and simulate [30,71,72]. But when, no visual information is available to reach a goal associated with a particular motivation, the system must plan its journey from a recognizable place to the other (need of latent learning [57]). During its exploration of the environment, it must learn places different enough from each other. It must also code the links between these places and so build a kind of “cognitive map”. In [27] we proposed a detailed simulation of such a system based on the results of our real robot experiments. Our first real world experiments show the PerAc architecture might be applied as a starting point to solve this new problem. Fig. 24 shows how the learning of few place–action associations is enough to create a complex shape of attractor basin so that the robot is able to go round huge obstacles to reach a goal that does not belongs to the visual environment of the starting point location.

This experiment demonstrates that place learning is a prerequisite and a first step before learning a cognitive map since if the robot had learned the places corresponding to the different parts of the trajectory a simple linking of those places could allow to create a cognitive map (see some simulations in [27]). The decision to learn intermediate places can be performed with the information about the goal recognition level. For instance, if the best recognized view is not correct, the robot can move in a bad direction but then at its new position the global activity of the place cells will decrease (see experimental results). It is easy to build a learning rule that is triggered when the sum of the place cells response decreases [30]. The robot would then find a movement that allows it to go in a direction associated with a global increase of the goal recognition (an efficient reinforcement learning rule is described in [28]). Our future work will consist in testing for real a planning level allowing the robot to pass from one subgoal to another in order to reach a particular goal.

Our robot is intended to be in interaction with its environment. It is just an agent that learns to agree with both its environment and its internal motivations. It has no global or complete representation of the world. The intrinsic meaning of recognized objects is not

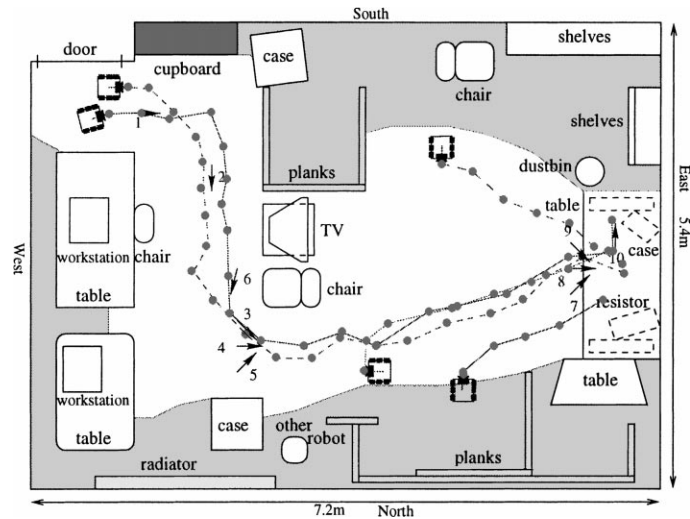


Fig. 24. The positions with figures and arrows represent the learned positions. The goal is under the table (near position 10). The gray area represents locations where the robot cannot succeed to reach the goal. The black points linked by lines represent some of the robot trajectory. The starting point of the trajectories is represented by the picture of the robot.

really understood by the robot (all the views associated with the same object are not explicitly linked to each other). It “keeps in memory” the link between a particular situation and the action that it has learned to be correct in that situation. Should the universe collapse, the robot’s memory would lose any meaning. The system functions correctly because of the dynamical interactions between the robot and its environment. In our point of view (following Piaget and others [48,55,61]), we think that new internal representations are built (in biological systems) and must be built (in robotics systems) autonomously when the interactions between the robot and the environment makes it necessary.

Concerning the understanding of natural cognition, our work shows the interest to carry analogical information all along the processing chain and to avoid to take hard decisions. The “weak” winner take all structures found in biological systems are not a limitation but one of their strongest properties. They allow them to be robust to the inherent contradictions and ambiguities of the perceived information. In related works [5] we proposed a global model of the hippocampus (Hs) which specifies the contributions of Hs in Long Term Memory (LTM) consolidation and transient Working Memory (WM) operations (spatio-temporal process-

ing). The WM function thought to operate both at cortical and Hs level is devoted to this *segregation* between information worthy to be “permanently” stored because of its relevance for survival, or its human social relevance for the personal history of the subject, and information which can be forgotten without major damage. WM provides the animal with the possibility to escape from a purely reflexive behavior, by taking into account its more or less immediate past and simulating the consequences of planned actions into the future. This temporo-spatial property is crucial to explain the short term storage of the information concerning the sequential exploration of the landmarks (that allows to separate angular information from the pure recognition of the landmarks). In our model, the structuration of space is considered as a degraded form of sequence learning and of the Hs capacity to perform cross correlations between multi-modal stimuli. In the robotic experiment, the use of a direct feedback from the motivations to the “place cells” is certainly not plausible (at least at the level of the Hs), since it is known that place cells activity does not depend on the animal motivation (does not depend on the current goal). But our model can be easily modified so that the motivations act on copies of the “place cells” that could be located in the nucleus accumbens for

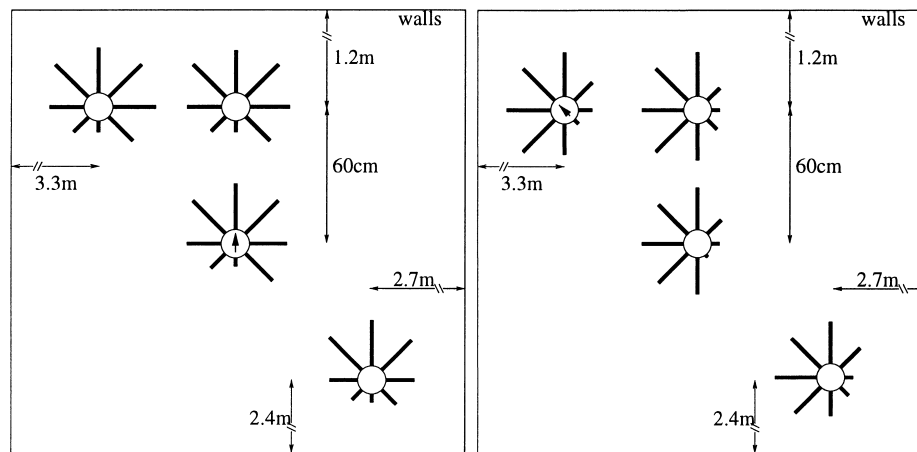


Fig. 25. When the robot visual field is reduced “place cell” like activity turns into “view cell”. (a), (b) Activities of 2 view cells according to the robot location represented by circles. The arrow in one of the circles represents the position and orientation used for the view learning. The directional vectors represent the neuron activity when the robot is facing a particular view.

instance where cells have been recorded that react to a combination of location and motivation.

Of course, our level of modelization cannot account today for the neurobiological details of each brain structure. Nevertheless, this kind of integrated modeling could allow to test for the coherence of global brain models and to verify if two functional boxes can really be connected [2].

Going back and forth between robotics and biology is full of promise. For instance, our robotic system will soon be improved by a better model of the hippocampus. Indeed, most of the robot time is spent in taking images to build a panoramic view (14 seconds to take the images and less than 1 second to compute the movement). This problem can be easily solved if the merged representation of “what” and “where” information was not reset after each move. The robot could then only turn its head to verify old information in order to update it. Moreover, the scalability problem which appears when merging thousands of landmarks recognition with hundreds of possible azimuths, can also be solved by using a more realistic hippocampus model. If the merging matrix is no more a hardwired “AND” operator but a limited set of neurons able to learn possible situations, the size of the matrix will no more directly depend on the number of possible landmark configurations. The computation complexity would then be reduced.

Finally, if places are well determined by their angular positions relative to a particular set of landmarks then the failure to find “place cells” in primates and more specially in monkeys [51] could have a very simple explanation. Indeed, the best condition to recognize a place surrounded by a lot of landmarks occurs when the animal is able to see landmarks all around the scene. The rat, because of the position of its eyes, has a large visual field similar to the angular size of the panoramic views of our robot. Thus, there is no reason the rat would be unable to recognize a place since the place can be characterized by its panoramic view. At the opposite, primates have a more limited visual field and the recording of their hippocampal cells predominantly show “view cells” activity. This view cell activity depends of the direction of gaze and the corresponding view. Conversely, it is independent of animal’s location. Our robotic experiments show that with a limited visual field (180° in our experiment) our robot exhibits the same kind of response (see Fig. 25). If the head is rotated beyond a certain angle, then there is no relationship between the extracted visual information even if the animal stays at the same place. The animal only recognizes the view. This recognition of a global view is very robust. The robot “view cells” exhibits important distance and orientation variations as observed in monkeys.

However, this interpretation leads to an important question. It is well known that hippocampal place cells of the rats for instance, continue to exhibit place field activities even in the dark. In that case the animal relies on tactile and smell stimuli as well as odometry (integration of the information linked to the animal motion — perception of the accelerations by the vestibular system for instance [62]). So if rat place cells receive both visual and odometric information why is this integration difficult in the case of primates? An explanation is first that view cells are not sufficient to build a place. Even if they are associated with odometric information they will allow the animal to predict the presence of a particular view but there is no reason they become place cells. Second, we can deduce that location segregation (building place cells) must be quite difficult from the sole odometric information (learning an environment in the dark and without proximity information that could be used to settle place cells near the borders of the area). Hence, the visual contribution to place cells would be preponderant in the process of building an association with other types of information and in particular odometric information. Indeed, odometric information is only relative and has to be linked to two places (departure place, integrated movement vector and arrival place) whereas visual information give an absolute result (a value representing the direction and distance to the learned place). So it is certainly more difficult to build a cognitive maps linking places to each other with only odometric information.

In conclusion, biological inspiration can be used to design very efficient robot controllers and pattern recognition systems. At the same time, robotic experiments appear to be a very promising tool to test biological models in more real life conditions and to propose new biological experiments.

Appendix A. Formalization of the PerAc architecture for the homing task

The PerAc architecture is formalized as follows. The visual input (perception) is a vector X . The activity of the neurons linked to the “recognition” of a learned situation X can be written as follows:

$$\begin{aligned} Rec(X) &= \Lambda(\psi_0(X), \psi_1(X), \dots, \psi_i(X), \dots, \psi_N(X))^T, \\ & \end{aligned} \quad (\text{A.1})$$

where Λ is an operator representing the competitive mechanism, Λ can be a strict winner take all (WTA) mechanism or a contrast enhancement of the most activated components in the input vector. $\psi_i(X)$ represents the recognition level of X by the neuron i (by definition $\psi_i(X) \in [0, 1]$). The size of the input and output vectors are the same. In the case of a strict WTA mechanism, the output V' of $\Lambda(V)$ is defined as a Kronecker vector $\delta_i = (0, \dots, 0, 1, 0, \dots, 0)^T$ with its only nonzero component being at i th position and V a vector.

$$\Lambda(V) = V' = \delta_j \quad \text{with } j = \text{Arg} \max_i \delta_i^T \cdot V. \quad (\text{A.2})$$

The vector \mathbf{A} associated with the output of the action selection group is

$$\mathbf{A} = \Lambda(\text{Act}(X)) \quad (\text{A.3})$$

with

$$\text{Act}(X) = [W] \cdot \text{Rec}(X) + A_R(X) + \text{noise},$$

where $[W]$ is the matrix of weights corresponding to the learned associations between perception and action. All the noise components are supposed to be generated by independent white noise sources. A_R is the proposed “reflex” action (unconditional signal or supervisor for action learning). A_R is defined as follows:

$$A_R = \delta_{[N \cdot \theta / (2\pi)]}, \quad (\text{A.4})$$

where

$$\begin{aligned} N &= \dim(\mathbf{A}) = \dim(\text{noise}) \\ &= \dim([W] \cdot \Lambda(\text{Rec}(X))) = \dim(A_R(X)) \end{aligned}$$

and θ is the direction of the proposed motion expressed in radians. $[x]$ is the closest integer to x . In the homing problem, A_R represents the discrete possibilities of rotation movements (with a constant translation speed). In a more complex case not studied here A_R is a one dimensional topological map using population coding [37].

The selected action resulting from Eq. (A.3) can be a rotation movement in the direction θ of the most activated component of the vector \mathbf{A} :

$$\theta = \frac{2\pi \cdot \arg \max_i^N (\delta_i^T \cdot \mathbf{A})}{N}. \quad (\text{A.5})$$

The learning of the perception–action associations is performed through the modification of the $[W]$ matrix so that the recognition $Rec(X)$ of the current state X alone allows to propose the correct action even if the reflex vector is not present. That means, we would like to have

$$\mathbf{A}(X) = [W] \cdot Rec(X). \quad (\text{A.6})$$

In order to get closer to this ideal situation, the following mean square error criterium¹⁰ is minimized:

$$E = \sum_{n=1}^N R(n) (A_R(n) - \mathbf{A}(n))^2, \quad (\text{A.7})$$

where $R(n)$ is the reinforcement signal at time n (the values can be positive, negative or null).

The matrix $[W]$ that minimizes the mean square error of E is

$$[W] = \sum_{n=1}^N R(n) \left((A_R(n) - \mathbf{A}(n)) \cdot Rec^T(X(n)) \right). \quad (\text{A.8})$$

In the one dimensional case, the learning procedure will consist in moving backhand forth to the left and to the right of the goal to learn two particular situations or states X_i . These states have to be associated with a so-called “reflex vector” that allows to reach the goal. The intensity of the two learned actions have only to be equal to each other (an unitary vector in the goal direction is enough) in the case of a symmetrical potential function. The learned states are X_L^- and X_L^+ such as

$$X_L^+ = X_0 + \Delta X, \quad X_L^- = X_0 - \Delta X \quad (\text{A.9})$$

and their associated actions are, respectively, $-\Delta X$ and $+\Delta X$. The different vectors can be written as follows:

$$Rec(X) = \Lambda \left(P(X_L^+, X), P(X_L^-, X) \right)^T \quad (\text{A.10})$$

with $P(X_0, X) = (X - X_0)^2$ for instance. The learned connections between the action group and the group

¹⁰ It has been shown that the Least Mean Square algorithm (LMS) [64] and the Rescorla–Wagner equation [49] modelization simple conditioning mechanisms are formally equivalent.

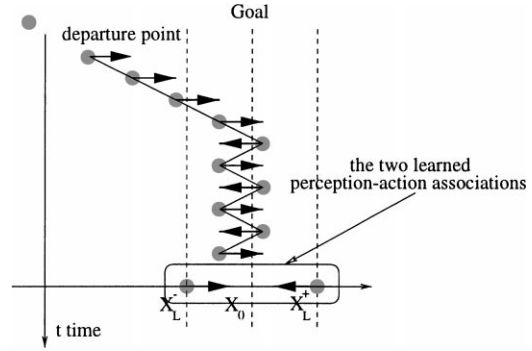


Fig. A.1. Example of a robot behavior using the PerAc architecture moving in a one dimensional space. Two Perception–Actions have been learned to define the homing behavior (learned places around X_L^+ and X_L^-). Note that the goal location X_0 has not been learned.

devoted to the recognition of the two learned situations X_L^+ and X_L^- will be $[W] = (-\Delta X, +\Delta X)$. Fig. A.1 represents the time trajectory of such a system in the case of a constant action speed. It is clear that the performances and the stability could be improved with a speed modulated by the distance to the goals (the values of the component of the vector P in the case an analogical competition is used). This case will not be studied here since the experimental results reported here were limited to that very simple case of constant speed (the constant step results from the need to move step by step because of computation time).

References

- [1] J.S. Albus, Outline for a theory of intelligence, IEEE Transactions on Systems and Cybernetics 21 (3) (1991) 473–509.
- [2] N. Almásson, G. Edelman, O. Sporn, Behavioral constraints in the development of neuronal properties: A cortical model embedded in a real-world device, Cerebral Cortex 8 (4) (1998) 346–361.
- [3] R.C. Arkin, Motor schema-based mobile robot navigation, International Journal of Robotics Research (1987) 92–112.
- [4] I.A. Bachelder, A.M. Waxman, Mobile robot visual mapping and localization: A view-based neurocomputational architecture that emulates hippocampal place learning, Neural Networks 7 (6/7) (1994) 1083–1099.
- [5] J.P. Banquet, P. Gaussier, J.C. Dreher, C. Joulain, A. Revel, Space-time, order and hierarchy in fronto-hippocampal system: A neural basis of personality, in: G. Matthews (Ed.), Cognitive Science Perspectives on Personality and Emotion, Elsevier, Amsterdam, 1997, pp. 123–189.

- [6] G. Bianco, R. Cassinis, A. Rizzi, N. Adami, P. Mosna, A bee-inspired robot visual homing method, in: Proceedings of the Second Euromicro Workshop on Advanced Mobile Robots (EUROBOT '97), Brescia, Italy, October 1997, pp. 123–189.
- [7] V. Braitenberg, *Vehicles: Experiments in Synthetic Psychology*, MIT Press/Bradford Books, Cambridge, MA, 1984.
- [8] R.A. Brooks, A robust layered control system for a mobile robot, *IEEE Journal of Robotics and Automation* 2 (1) (1986) 14–23.
- [9] R.A. Brooks, L.A. Stein, Building brains for bodies, *Autonomous Robots* 1 (1994) 7–25.
- [10] G. Bugmann, A connectionist approach to spatial memory and planning: Perspectives in neural networks, *Basic Concepts in Neural Networks: A Survey*, Springer, Berlin, 1997, Chapter 5.
- [11] Y. Burnod, Architecture par niveaux du cortex cerebral: Un mécanisme possible, in: *Cognitiva* 1987.
- [12] Y. Burnod, *An Adaptive Neural Network: The Cerebral Cortex*, Masson, Paris, 1989.
- [13] G.A. Carpenter, S. Grossberg, Invariant pattern recognition and recall by an attentive self-organizing art architecture in a nonstationary world, in: Proceedings of the International Joint Conference on Neural Networks, 1987, pp. 737–745.
- [14] B.A. Cartwright, T.S. Collett, Landmark learning in bees, *Journal of Computational Physiology* 151 (1983) 521–543.
- [15] J.S. Chahl, M.V. Srinivasan, Reflective surfaces for panoramic imaging, *Applied Optics* 36 (31) (1997) 8275–8285.
- [16] R. Chatila, Deliberation and reactivity in autonomous mobile robots, *Robotics and Autonomous Systems* 16 (2–4) (1995) 197–211.
- [17] T.S. Collett, J. Zeil, The selection and use of landmarks by insects, 1997, pp. 41–65.
- [18] M.J. Denham, J. Boitano, A model of the interaction between prefrontal cortex, septum and the hippocampal system in the learning and recall of goal-directed sensory-motor behaviours, Technical Report NRG-96-01, University of Plymouth, School of Computing, 1996.
- [19] G. Edelman, *Neural Darwinism: The Theory of Neuronal Group Selection*, Basic Books, New York, 1987.
- [20] A. Etienne, Mammalian navigation, neural models and biorobotics, *Connection Science* 10 (3–4) (1998) 271–289.
- [21] A. Etienne, R. Maurer, J. Berlie, B. Reverdin, T. Rowe, J. Georgakopoulos, V. Séguinot, Navigation through vector addition, *Nature* 396 (1998) 161–164.
- [22] M.O. Franz, B. Schölkopf, H.H. Bülthoff, Homing by parameterized scene matching, in: Proceedings of the 4th European Conference on Artificial Life, MIT Press, Cambridge, MA, 1997, pp. 236–245.
- [23] M.O. Franz, B. Schölkopf, H.A. Mallot, H.H. Bülthoff, Where did I take that snapshot? Scene-based homing by image matching, *Biological Cybernetics* 79 (1998) 191–202.
- [24] C.R. Gallistel, *The Organization of Learning*, MIT Press, Cambridge, MA, 1993.
- [25] P. Gaussier, J.P. Cocquerez, Neural networks for complex scene recognition: Simulation of a visual system with several cortical areas, in: Proceedings of the International Joint Conference on Neural Networks, Baltimore, MD, 1992, pp. 233–259.
- [26] P. Gaussier, C. Joulain, S. Moga, M. Quoy, A. Revel, Autonomous robot learning: What can we take for free? in: Proceedings of the International Symposium on Industrial Electronics, ISIE'97, IEEE, Guimarães, Portugal, July 1997.
- [27] P. Gaussier, S. Leprêtre, C. Joulain, A. Revel, J.P. Banquet, Animal and robot learning: Experiments and models about visual navigation, in: Proceedings of the Seventh European Workshop on Learning Robots, Edinburgh, UK, 1998.
- [28] P. Gaussier, A. Revel, C. Joulain, S. Zrehen, Living in a partially structured environment: How to bypass the limitation of classical reinforcement techniques, *Robotics and Autonomous Systems* 20 (1997) 225–250.
- [29] P. Gaussier, S. Zrehen, Avoiding the world model trap: An acting robot does not need to be so smart!, *Journal of Robotics and Computer-Integrated Manufacturing* 11 (4) (1994) 279–286.
- [30] P. Gaussier, S. Zrehen, Navigating with an animal brain: A neural network for landmark identification and navigation, in: Proceedings of Intelligent Vehicles, Paris, 1994, pp. 399–404.
- [31] P. Gaussier, S. Zrehen, PerAc: A neural architecture to control artificial animals, *Robotics and Autonomous Systems* 16 (2–4) (1995) 291–320.
- [32] S. Grossberg, Nonlinear neural networks: Principles, mechanisms, and architectures, *Neural Networks* 1 (1988) 17–61.
- [33] R. Hecht-Nielsen, Counterpropagation networks, *Applied Optics* 26 (23) (1987) 4979–4984.
- [34] C. Joulain, P. Gaussier, A. Revel, Learning to build categories from perception–action associations, in: Proceedings of the International Conference on Intelligent Robots and Systems, IROS'97, IEEE/RSJ, Grenoble, France, September 1997.
- [35] S.P.D. Judd, T.S. Collett, Multiple stored views and landmark guidance in ants, *Nature* 392 (1998) 710–712.
- [36] O. Khatib, Real-time obstacle avoidance for robot manipulator and mobile robots, *International Journal of Robotics Research* 5 (1) (1986) 90–98.
- [37] T. Kohonen, *Self-Organization and Associative Memory*, Springer, New York, 1984.
- [38] T. Levitt, D. Lawton, D. Chelberg, K. Koitzsch, J.W. Dye, Qualitative Navigation 2, in: Proceedings of the DARPA Image Understanding Workshop, Los Altos, CA, 1988, pp. 319–326.
- [39] Y. Matsumoto, M. Inaba, H. Inoue, Memory-based navigation using omni-view sequence, in: A. Zelinsky (Ed.), *International Conference on Field and Service Robotics*, Camberra, Panther Publishing, Camberra, 1997, pp. 184–191.
- [40] D. McFarland, Animal robotics — from self-sufficiency to autonomy, in: P. Gaussier, J.D. Nicoud (Eds.), *From Perception to Action*, Lausanne, Switzerland, IEEE Computer Society Press, Silver Spring, MD, 1994.
- [41] J.A. Meyer, S.W. Wilson (Eds.), *First International Conference on Simulation of Adaptive Behavior: From Animals to Animats*, MIT Press/Bradford Books, Cambridge, MA, 1991.

- [42] B. Milner, S. Corkin, H.L. Teuber, Further analysis of the hippocampal amnesia syndrome: 14-year follow-up study of h.m., *Neuropsychologia* 6 (1968) 215–234.
- [43] R. Möller, D. Lambrinos, R. Pfeifer, T. Labhart, R. Wehner, Modeling ant navigation with an autonomous agent, in: Proceedings of the fifth International Conference on Simulation of Adaptive Behavior: From Animals to Animats, MIT Press/Bradford Books, Cambridge, MA, 1998, pp. 185–194.
- [44] R.G. Morris, Spatial localization does not require the presence of local cues, *Learning and Motivation* 12 (1981) 239–260.
- [45] M. Muller, R. Wehner, Path integration in desert ants, *Cataglyphis fortis*, *National Academy of Science*, 1988, pp. 5287–5290.
- [46] R.C. Nelson, Visual homing using an associative memory, *Biological Cybernetics* 65 (1991) 281–291.
- [47] J. O’Keefe, *Neural Connections, Mental Computation*, MIT Press, Cambridge, MA, 1989, pp. 225–284.
- [48] J. Piaget, *La naissance de l’intelligence chez l’enfant*, Delachaux et Niestle Editions, Geneve, 1936.
- [49] R.A. Rescorla, A.R. Wagner, A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement, in: *Classical Conditioning II: Current Research and Theory*, Appleton-Century-Crofts, New York, 1972.
- [50] A. Rizzi, R. Cassinis, G. Bianco, N. Adami, P. Mosna, A biologically-inspired visual homing method for robots, in: Proceedings of the Seventh Workshop of AI*IA on Cybernetics and Machine Learning, Ferrara, Italy, April 1998.
- [51] E.T. Rolls, S.M. O’Mara, View-responsive neurons in the primate hippocampal complex, *Hippocampus* 5 (1995) 409–424.
- [52] E.T. Rolls, A. Treves, *Neural Networks and Brain Function*, Oxford University Press, New York, 1998.
- [53] N.A. Schmajuk, A.D. Thieme, Purposive behavior and cognitive mapping: A neural network model, *Biological Cybernetics* 67 (1992) 165–174.
- [54] L. Steels, A selectionist mechanism for autonomous behavior acquisition, *Robotics and Autonomous Systems* 20 (1997) 117–131.
- [55] J. Stewart, The implication for understanding high-level cognition of a grounding in elementary adaptive systems, *Robotics and Autonomous Systems* 16 (2–4) (1995) 107–116.
- [56] N. Tinbergen, *The Study of Instinct*, Oxford University Press, London, 1951.
- [57] E.C. Tolman, Cognitive maps in rats and men, *The Psychological Review* 55 (4) (1948) 189–208.
- [58] E.C. Tolman, C.H. Honzik, “Insight” in Rats, *California Publications in Psychology* 4 (1930) 215–232.
- [59] O. Trullier, S.I. Wiener, A. Bethoz, J.A. Meyer, Biologically based artificial navigation systems: Review and prospects, *Progress in Neurobiology* 51 (1997) 483–544.
- [60] L.G. Ungerleider, Functional brain imaging studies of cortical mechanisms for memory, *Science* 270 (5237) (1995) 769.
- [61] F. Varela, E. Thompson, E. Rosch, *The Embodied Mind*, MIT Press, Cambridge, MA, 1993.
- [62] H.S. Wan, D.S. Touretzky, A.D. Redish, Towards a computational theory of rat navigation, in: M. Mozer, P. Smolensky, D.S. Touretzky, J.L. Elman, A. Weigend (Eds.), *Proceedings of the 1993 Connectionist Models Summer School*, Lawrence Erlbaum, Hillsdale, NJ, 1994, pp. 11–19.
- [63] R. Wehner, B. Michel, P. Antonsen, Visual navigation in insects: Coupling of egocentric and geocentric information, *Journal of Experimental Biology* 199 (1996) 129–140.
- [64] B. Widrow, M.E. Hoff, Adaptive switching circuits, in: *IRE WESCON, Convention Record*, New York, 1960, pp. 96–104.
- [65] I.Q. Wishaw, Latent learning in a swimming pool place task by rats: Evidence for the use of associative and not cognitive mapping processes, *The Quarterly Journal of Experimental Psychology B* 43 (1) (1991) 83–103.
- [66] J. Zeil, Orientation flights of solitary wasps: I. Description of flight, *Journal of Comparative Physiology A* 172 (1993) 189–205.
- [67] J. Zeil, A. Kelber, R. Voss, Structure and function of learning flights in bees and wasps, *Journal of Experimental Biology* 199 (1996) 245–252.
- [68] S.W. Zhang, K. Bartsch, M.V. Srinivasan, Maze learning by honeybees, *Neurobiology of Learning and Memory* 66 (1996) 267–282.
- [69] S.W. Zhang, M.V. Srinivasan, Prior experience enhances pattern discrimination in insect vision, *Nature* 368 (1994) 330–332.
- [70] D. Zipser, A computational model of hippocampal place fields, *Behavioral Neuroscience* 99(5) (1985) 1006–1018.
- [71] S. Zrehen, *Elements of brain design for autonomous agents*, PhD thesis, EPFL, 1995.
- [72] S. Zrehen, P. Gaussier, Building grounded symbols for localization using motivations, in: P. Husbands, I. Harvey (Eds.), *Proceedings of the Fourth European Conference on Artificial Life, ECAL97*, Brighton, UK, July 1997, pp. 299–308.

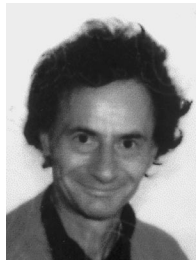


Philippe Gaussier was born in 1967 in Marseille, France. He received the M.S. degree in Electronics from Aix-Marseille University in 1989. In 1992, he received a Ph.D. degree in Computer Science from the University of Paris XI (Orsay) for a work on the modelization and simulation of a visual system inspired by mammal vision. From 1992 to 1994, he conducted research in neural network applications and in control of autonomous mobile robots at the Swiss Federal Institute of Technology. He was the Program Chairman of the “From Perception to Action Conference: PerAc’ 94” and has edited a special issue of the journal *Robotics and Autonomous Systems* on “Moving the Frontier between Robotics and Biology”. He is now Professor at the Cergy Pontoise University in France and leads the neurocybernetic team of the Image and Signal Processing Lab. He has worked on various problems related to industrial applications of neural networks, visual recognition, robot navigation, and neural modeling. Currently his research interests are focused on the

one hand on the modelization of the cognitive mechanisms involved in the visual perception and on the other hand on the modelization of the hippocampus and its relations with cortical structures like parietal, temporal and prefrontal areas. Current robotic applications include autonomous and on-line learning for motivated visual navigation (such as place learning, visual homing, and object discrimination).



Cedric Joulain was born in 1972 in Poitiers, France. He is a Graduate Engineer of the Ecole Internationale des Sciences du Traitement de l'Informations (EISTI) de Cergy-Pontoise (graduate in Computer Science). He is preparing a French Ph.D. in the Neurocybernetic team of the Image and Signal Processing Lab of ENSEA. He is currently working at the Banque Nationale de Paris (BNP) in the Fixed Income Engineering team.



Jean Paul Banquet holds an MD with a residency in Neuropsychiatry and a Ph.D. in Applied Mathematics from Paris VI (Pierre et Marie Curie University), 1981. During two three-year periods, he was appointed Fullbright Research Fellow at Stanley Cobb Laboratories for Neurophysiological Research (Harvard Medical School), and at the Center for Cognitive and Neural Systems chaired by Professor Grosberg at Boston University. After doing electrophysiological research on memory and learning, he is presently working at an INSERM unit: Neuroscience et Modelisation at Pierre et Marie Curie, where he conducts research on neurophysiology and neural networking modelling of the hippocampus, basal ganglia, and

their relations with the cortex. His collaboration with the Neurocybernetic group lead by Philippe Gaussier takes a significant amount of his time.



Sacha Leprêtre received an Engineer Degree in Electronic and Computer Science Engineering from the University of Montpellier, France, in 1996, the Masters Degree (D.E.A.) in Signal and Image Processing from the University of Cergy, France, in 1997. He is currently a Ph.D. student at the Equipe de Traitement des Images et du Signal (ETIS) within the Neuro-Cybernetics Group. His research interests include object recognition, visual cognition, active vision, image analysis and understanding, and most generally neural network architectures inspired by biological data.



Arnaud Revel was born in 1970 in Angers, France. He is a Graduate Engineer of the Ecole Nationale Supérieure de l'Electronique et de ses Applications (ENSEA) de Cergy-Pontoise (graduate in Computer Science) and is Doctor of the University of Cergy-Pontoise. He is currently Assistant Professor in the Neurocybernetic team of the Image and Signal Processing Lab of ENSEA. His research interests are developing neural architectures and learning algorithms inspired by biology and psychology in order to control an animal-like mobile robot.