To appear at: Proc. 27th International Conference on Artificial Neural Networks "ICANN 2018"

A Hybrid Planning Strategy through Learning from Vision for Target-directed Navigation

Xiaomao Zhou[1,2], Cornelius Weber[2], Chandrakant Bothe[2] and Stefan Wermter[2]

[1] Harbin Engineering University, College of Automation, Nantong Street 145, 150001, Harbin, China
[2] University of Hamburg, Department of Informatics, Vogt-Kölln-Strasse 30, 22527 Hamburg, Germany
{zhou,weber,bothe,wermter}@informatik.uni-hamburg.de
http://www.informatik.uni-hamburg.de/WTM

Abstract. In this paper, we propose a goal-directed navigation system consisting of two planning strategies that both rely on vision but work on different scales. The first one works on a global scale and is responsible for generating spatial trajectories leading to the neighboring area of the target. It is a biologically inspired neural planning and navigation model involving learned representations of place and head-direction (HD) cells, where a planning network is trained to predict the neural activities of these cell representations given selected action signals. Recursive prediction and optimization of the continuous action signals generates goal-directed activation sequences, in which states and action spaces are represented by the population of place-, HD- and motor neuron activities. To compensate the remaining error from this look-ahead model-based planning, a second planning strategy relies on visual recognition and performs target-driven reaching on a local scale so that the robot can reach the target with a finer accuracy. Experimental results show that through combining these two planning strategies the robot can precisely navigate to a distant target.

Keywords: Navigation, Place cell, Head-direction cell, Vision-recognition

1 Introduction

Studies in neuroscience have revealed that animals' spatial cognition and planning behaviors during navigation involve certain types of location- and direction-sensitive cells in the hippocampus, which support an animal's sense of place and direction [1][2]. More recent studies suggest that these spatially related firing activities also underlie animals' behavioral decisions [3].

Considering existing approaches for modeling hippocampal cells, most of them just focus on how to develop the location- or direction-related firing patterns while only few care about the computational principle underlying the formation of these firing activities [4]. Slow feature analysis (SFA) [5] tries to explain this problem by an unsupervised learning algorithm that extracts slowly varying features from fast-changing source signals based on the slowness principle. In our previous work, place- and HD cells were simultaneously learned from visual inputs using a modified SFA learning algorithm which can develop separated populations of place and HD cell types by restricting their learning to separate phases of spatial exploration [6]. However there remains a question of how to use the metric information hidden in these cell activities, which are obtained by unsupervised learning, to support a navigation task.

In this paper, based on the learned cell representations, we propose a navigation model that performs forward look-ahead planning and predicts a sequence of neural activities encoding intermediate waypoints from a starting position to a goal position, where the spatial positional state and directional state are represented by the learned place and HD cell representations, respectively. Furthermore, inspired by the biological finding that place cells are able to generate future sequences encoding spatial trajectories towards remembered goals, which demonstrates their predictive role in navigation [7], we propose a model of their functional role in directing spatial behaviors. Here, we mainly introduce the look-ahead planning whose architecture is shown in Fig. 1. The front part (visual processing part) consists of two parallel image-processing channels with a different network for the emergence of place and HD cells, respectively. For the unsupervised training and network parameters please refer to our previous work [6]. The latter (route planning part) is a world model that supports the imaginary planning in goal-directed navigation, where the world state is represented by the ensemble activity of place and HD cells.

However such model-based forward planning suffers from significant accumulation errors when dealing with long-range predictions. Furthermore, it takes into account only the place cell representations of the target, irrespective of specific visual properties of a target. In many cases, this planning can only lead the robot to the neighboring areas of a target, instead of to the precise target position. To solve this problem, we propose a second planning strategy that starts to perform after the look-ahead planning. Its aim is to recognize the target based on vision and to move directly towards it after recognizing it.

2 Hybrid Planning Strategy

Based on information learned from vision, the proposed hybrid planning strategy uses two different coordinate systems. The first one is based on space representations which are obtained in an unsupervised way. The second one is based directly on visual representations of the goal. The concept of switching between different planning strategies during navigation can be found in similar work [8][9].

2.1 Model-based Look-ahead Planning

For look-ahead planning, we first train a predictive world model network which predicts the subsequent state given the current state and action. The continuous

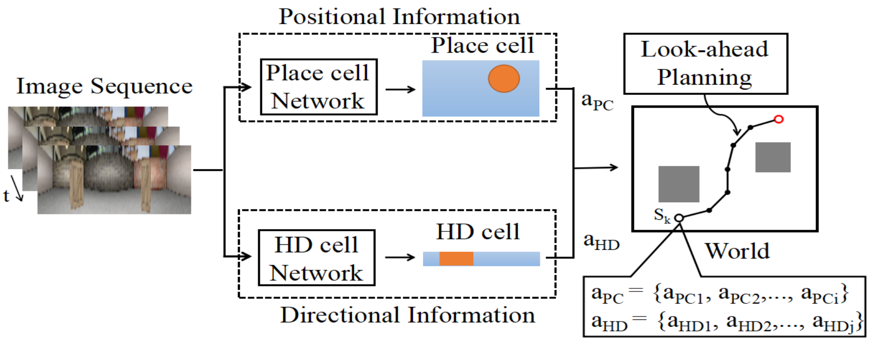


Fig. 1: An overview of the system architecture. The immediate response of the trained place or HD cell network to an image resembles the firing activity of place and HD cells at a certain position or to a certain direction where the image is captured. The world model trained based on the learned cell representations is used to support look-ahead planning.

spatial state is represented by the ensemble activity of place and HD cells and the continuous action determines the change of moving direction during a transition, assuming a forward movement of constant speed. The world model is represented by a multi-layer perceptron (MLP) with 81 inputs (30 place cells + 50 HD cells + 1 rotation angle) and 80 outputs (30 place cells + 50 HD cells).

The planning process is based on the recursive use of the fully trained world model which generates a sequence of neural activations encoding the spatial trajectory from an initial location to a given target location (represented in the same place- and HD space), together with corresponding action commands [10]. To generate an optimal route, the planner first constructs a multi-step forward look-ahead probe by sequentially simulating the execution of each command in a given action sequence on a world model chain, as shown in Fig. 2. Then it optimizes the actions recursively in the direction of the desired goal location. The planning trajectory is optimized by modifying the actions via gradient descent to minimize the distance to the goal location. With this approach, routes towards a desired goal are imaginatively explored prior to execution by activating the place cell activities, while corresponding moving directions along the route are encoded by HD cell activities. For each optimization iteration, the action is updated as follows:

$$\Delta a(t) = -\eta \frac{\partial E_{plan}}{\partial a(t)}, \text{ where } E_{plan} = \sum^{K} \frac{1}{2}(S_k^{goal} - S_k^{pred})^2 \tag{1}$$

The state vector S consists of an ensemble firing activity of place and HD cells (K in total), η is a constant learning rate. The training objective is to optimize the action sequence $a(t)$ such that the predicted ending state S^{pred} is close to the goal state S^{goal}, which is calculated by the SFA network given the image taken at the target position.

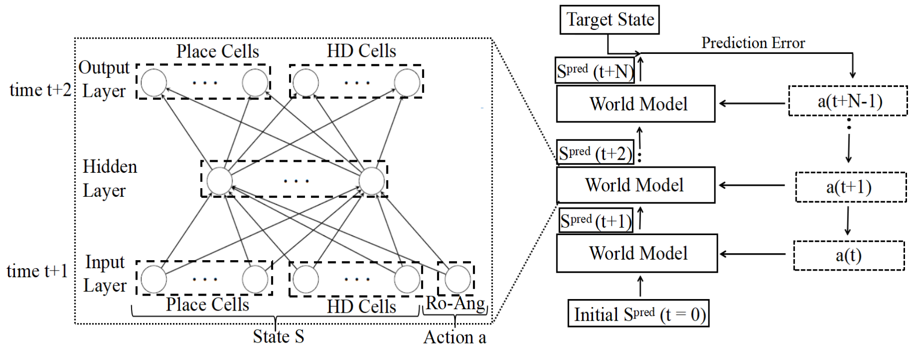


Fig. 2: An overview of the planning architecture. The world model which has been trained based on the learned cell representations is used to support look-ahead planning. Left (inset), the MLP used for one-step prediction. Right, multi-step prediction in the planning phase with feedback of the prediction error.

Note that planning assumes a predefined prediction depth according to the distance to a goal location, while prior information about the optimal depth is not always available. To overcome this assumption of the existing model [10], we propose an adaptive-depth approach where the planning starts with a 1-step prediction and incrementally increases the depth until adding one more prediction step would let the ending position of the current plan go beyond the goal location. During depth increase, the previous plan naturally provides a good proposal for the initialization of the next plan whose prediction increases in depth. Since the previous plan is already optimized but fails due to its small prediction depth, this enables the planner to find the best prediction depth towards a goal without any prior information to efficiently optimize the trajectory.

2.2 Vision-directed Reaching based on Target Recognition

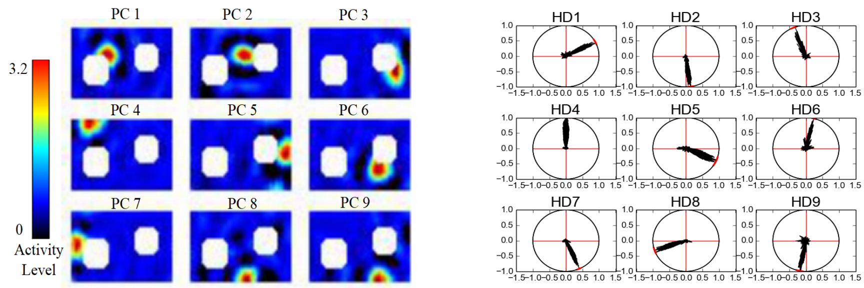
While the look-ahead planning can approximately navigate the robot towards the target position, the robot will either overstep or stop short of the target by about one step size and will rarely stop precisely on the target. To solve this problem, we adopt a second planning strategy that is based on object/scene recognition. The goal-directed planning will be activated after the robot has executed the plans optimized by the look-ahead planning, in which case the robot is supposed to be close to the target and will be able to see the target. Since a target always refers to particular objects (like chair, computer...) or specific scenes (like kitchen, corridor...), the robot can recognize the target. After perceiving the target, the robot will adjust its head direction to keep the target in the center of its view and move towards it.

3 Experiments and Results

3.1 Simulation Experiment for Look-ahead Planning

To test the look-ahead planning, we first used a simulated robot moving in a RatLab virtual-reality environment which also generated the visual data for training place- and HD cell networks [11]. RatLab is designed to simulate a virtual rat doing random explorations and allows to modify the environmental parameters and movement patterns according to the user's purposes.

We first trained place- and HD cell networks by learning from the visual input with SFA, where the images generated during turning movements are used to train the place cell network, while the HD cell network is mainly trained using images from forwarding movements [6]. We trained 30 place cells and 50 HD cells whose ensemble activity encodes the spatial position and direction, respectively. Training results are partly shown in Fig. 3.

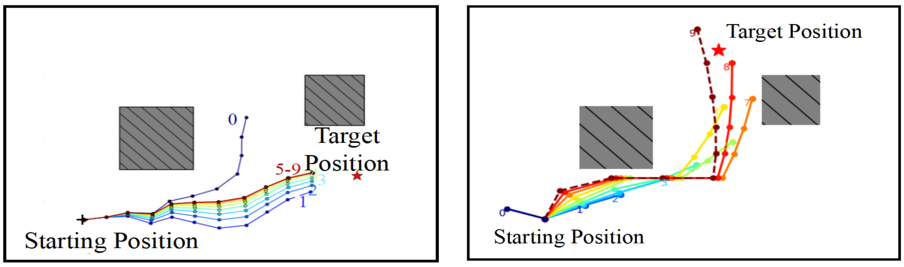


(a) (b)

Fig. 3: Firing patterns of learned place and HD cells to different positions or directions. (a) Firing patterns of 9 representative place cells. (b) Polar plots showing the firing patterns of 9 representative HD cells.

For the planning result, Fig. 4 (a) and (b) show separately plans with a fixed depth of 10 and adaptive-depth plans, where the planning in the place cell space is mapped to the 2D space through finding the position that yields the most similar firing pattern. The prediction depth of 10 for Fig. 4 (a) is obtained empirically and the initial route 0 is gradually optimized towards the desired goal location. The given example shows plans with a quite good initialization, while if given a starting route 0 that extends into a very different direction from the desired one, the planning may not be successful. This is because a long prediction makes the planning optimization based on back-propagating through a long chain of world models very difficult. Due to a vanishing gradient, initial segments receive too little correction. While the adaptive-depth planning could start with a bad initialization, as route 0 shown in Fig. 4 (b), the planning starts with a 1-step prediction and is optimized immediately to a better direction through the world model chain which currently contains only one model step. This optimized plan then works as a good basis for initializing the next plan with one step more. This explains why the initial part of each route in Fig. 4 (b) clusters in a narrow area. The planning depth increases incrementally until finding an appropriate plan (route 8) to the goal location.

To evaluate the look-ahead planning performance over the global area, we fixed the starting position and uniformly sampled 120 positions from the environment as the target. As shown in Fig. 5, the planning performance deteriorates as the distance between the target and the starting position increases. Especially when the target lies in the areas behind the second obstacle, which is far away, planning becomes very difficult and may fail. This might be due to the accumulation error in the long world model chain and also the optimization based on backpropagation is difficult for a long-step planning.

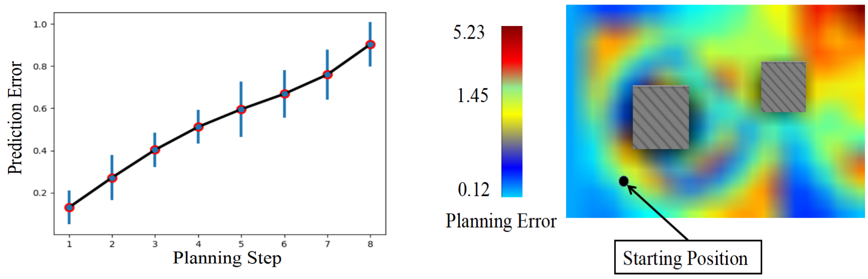


(a) (b)

Fig. 4: The proposed look-ahead trajectories with (a) a fixed depth of 10 steps and (b) an adaptive depth. The solid dots represent the intermediate locations from the starting position to the target position (red star). The dashed line (route 9) represents a route that exceeds the goal. Planning is performed in place- and HD cell representation space and the trajectory based on actions of the plan is shown in x, y- space for visualisation.

3.2 Real-world Experiment for Target Object Approaching

As a second step in our hybrid model, we test the vision-based target approaching in a real-world environment with a Turtlebot3 robot in a simple goal reaching task. The robot is placed at a position where the target is in the range of its vision (which refers to the state after executing the look-ahead planning) and its goal is to find the target object and move close to it. For detecting and recognizing the target, we used the YOLO network which is fast and can accurately recognize, classify and localize objects [12]. If the robot cannot see the target object at the initial state, it will rotate locally with a constant speed until perceiving and recognizing the object with a certain probability. While trying to keep the



(a) (b)

Fig. 5: (a) The prediction error of the world model increases with the number of the planning steps. (b) The planning error over the whole environment, where the starting position is fixed (the black dot) and the target is sampled uniformly from the rectangular environment which has a size of 14×10 units and 120 positions are sampled from it. The error value is represented by the color.

target object in the center of the view, the robot moves directly towards it until reaching the threshold distance to the target (Fig. 6).

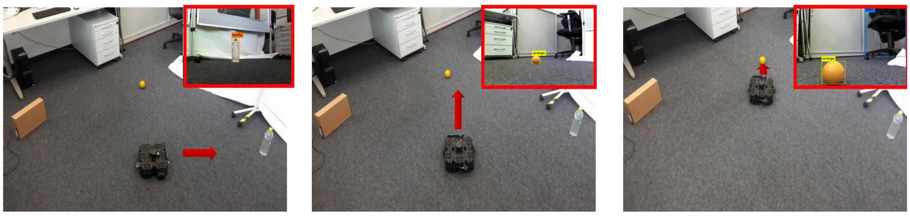


Fig. 6: Test of the object recognition and target approaching. The robot starts from a neighboring area and needs to reach the target orange. Left: The robot starts without the target in the current view (shown in the red box) and starts rotating. Middle: The robot perceives and recognizes the target and starts moving towards it. Right: The robot reaches the orange and stops just next to it.

4 Conclusion and Future Work

We have proposed a navigation system that relies on a hybrid navigation strategy in order to precisely reach a target location, which consists of two planning strategies that work on different distance scales but both rely on vision. The first one is look-ahead planning that works on a global coordinate system and proposes a spatial trajectory close to the desired goal location. The spatial state is represented by the ensemble activity of place and HD cells, which are modeled by learning directly from visual input based on an unsupervised SFA learning algorithm. The planning network allows looking into the future based on a chain of world model predictions and adaptively proposes optimized prediction steps to the goal location. The second part is a target approaching strategy working on a local scale, which enables object recognition and goal-directed reaching. Through combining these two complementary strategies, the robot can move from a random position to a target position with a high accuracy using just its vision system. As future work, we will extend the simulated scenario to a physical world where place and HD cells are modeled on a real robot using its vision sensor and the planning is validated in a challenging dynamic environment.

Acknowledgments. We acknowledge support from the German Research Foundation DFG, project CML (TRR 169) and the EU, project SECURE (No 642667).

References

1. O'Keefe, L. Nadel. The hippocampus as a cognitive map. Oxford: Clarendon Press, 1978.
2. J. S. Taube, R. U. Muller and J. B. Ranck, Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. Journal of Neuroscience, 10(2), pages 420-435, 1990.
3. Wills, T.J., Muessig, L. and F. Cacucci, The development of spatial behaviour and the hippocampal neural representation of space. Phil. Trans. R. Soc. B, DOI: 10.1098/rstb.2013.0409, 2014.
4. P. J. Zeno, S. Patel and T.M. Sobh, Review of neurobiologically based mobile robot navigation system research performed since 2000. Journal of Robotics, DOI: 10.1155/2016/8637251, 2016.
5. M. Franzius, H. Sprekeler and L. Wiskott, Slowness and sparseness lead to place, head-direction, and spatial-view cells. PLoS Computational Biology, 3(8), 2007.
6. X. Zhou, C. Weber and S. Wermter, Robot Localization and Orientation Detection Based on Place Cells and Head-Direction Cells. In International Conference on Artificial Neural Networks (ICANN 2017), pages 137-145, Springer, 2017.
7. Pfeiffer, E. Brad and D. J. Foster, Hippocampal place-cell sequences depict future paths to remembered goals. Nature, 497 (7447), pages 74-79. 2013.
8. L. Dollé, D. Sheynikhovich, B. Girard, R. Chavarriaga and A. Guillot, Path planning versus cue responding: a bio-inspired model of switching between navigation strategies. Biological Cybernetics, 103(4), pages: 299-317, 2010.
9. T. Oess, J. L. Krichmar and F. Röhrbein, A Computational Model for Spatial Navigation Based on Reference Frames in the Hippocampus, Retrosplenial Cortex, and Posterior Parietal Cortex. Frontiers in Neurorobotics, DOI: 10.3389/fnbot.2017.00004, 2017.
10. S. Thrun, K. Möller and A. Linden, Planning with an adaptive world model. In Advances in Neural Information Processing Systems, pages 450-456, 1991.
11. F. Schönfeld and L. Wiskott, RatLab: an easy to use tool for place code simulations. Frontiers in Computational Neuroscience, DOI: 10.3389/fncom.2013.00104, 2013.
12. J. Redmon, S. Divvala, R. Girshick and A. Farhadi, You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages: 779-788, 2016.