Goal-Directed Navigation with A Hybrid Planning Strategy Through Learning from Vision

Xiaomao Zhou*, Bing Du, Tao Huang
Purple Mountain Laboratories, No.9, Mozhoudong Road, Nanjing, China

*xiaomaozhou26@gmail.com

Abstract. The existing hippocampal modeling approaches rarely span the wide functionality range from processing raw sensory signals to planning and action. This paper presents a goal-directed navigation system consisting of two planning strategies. The first one is a biologically inspired neural planning and navigation model that is related to learned representations of place and HD cells. It is responsible for generating spatial trajectories leading to the neighboring area of the target. The place and HD cells are trained unsupervisedly from visual images using a modified slow feature analysis (SFA) algorithm. To interpret their functional role in navigation, a planning network is trained to predict the neural activities of place and HD cell representations given selected action signals. Recursive prediction and optimization of the action signals generate goal-directed activation sequences, in which the continuous states and action spaces are represented by the population of place-, HD- and motor neuron activities. Furthermore, a second planning strategy relying on visual recognition is proposed and performs target-driven reaching on a local scale for finer accuracy. Experimental results show the effectiveness of the proposed system.

Keywords: Robot navigation, Neural networks, Place cells, Head direction cells.

1. Introduction

Numerous experiments have demonstrated the essential role of hippocampus cells in solving navigational tasks, such as route planning[1] and goal-directed behavior [2], etc. However, the exact mechanisms of how these cells contribute to navigation are still not fully understood. Previous research mainly focused on the function of the hippocampus in encoding spatial locations [3]. Recent research revealed that the hippocampal place cells could depict brief sequences encoding spatial trajectories from the current location to a remembered location [4], even to construct sequences through unexplored space based on previously acquired experiences [5]. This demonstrates the predictive role of place cells in navigation. Inspired by this, based on the learned hippocampal cell representations, we propose a navigation model performing look-ahead planning. The planning is based on a world model which is trained to approximate the world behavior, where the spatial positional and directional states are decoded by the generated different cell representations. The trajectory towards the target and corresponding action commands are generated by recursive prediction and optimization through looking ahead based on a world model chain. The action commands are used to control an agent's moving direction. While such model-based forward planning depends suffers from accumulated errors, in many
cases, this planning only leads the agent to the neighboring area of the target, instead of to the precise target position. To solve this problem, we propose a second planning strategy that works complementarily to the first one in the aim to reduce the distance between the robot's ending position and the desired target position. The second planning strategy starts executing after the look-ahead planning completes or when the target is within the robot's vision range. With these two planning strategies, a robot is able to reach a distant target position with a finer accuracy just based on learning representations from vision.

2. Model structure

The architecture of the proposed system is shown in Fig. 1. It consists of the front part (visual processing part), which has two parallel image-processing channels with two networks, and the latter part (route planning part) providing an illustration of how the proposed hybrid planning strategy works.

3. World modeling in an unsupervised way

In this work, we use SFA to extract spatial information by directly learning from a robot's visual images. For this part, we follow our previous work [6] in generating place and HD cells for spatial representation.

4. Model-based Look-ahead Planning

For this part, we first train a world model network that predicts state transitions, we use 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) to represent the world model. The parameters of the forward model are updated according to:

$$
\Delta W = -\mu \frac{\partial E_{\text{model}}}{\partial W}, \quad \text{where } E_{\text{model}} = \sum_{k=1}^{K} \left( S_k^{\text{truth}} - S_k^{\text{pred}} \right)^2
$$

(1)
The state vector $S$ decodes firing activity of place and HD cells, $\mu$ is a constant learning rate and $W$ represents the weight in the forward model.

The training objective is to obtain a world model that mimics the one-step state transition in the environment, where the subsequent state can be predicted given the current state and action command. $\text{Struth}$ represents the ground truth of the next state and $\text{Spred}$ represents the predicted state by the world model. The planning is based on the recursive use of the fully trained world model, where the calculation process is shown in Fig. 2.

![Fig. 2 An overview of the look-ahead planning architecture.](image)

The final output represents the planning's ending state, which is compared with the target state. Then actions and planning trajectory are recursively optimized with the aim to minimizing the deviation of the final state with respect to the goal state. During learning, the action is calculated as follows:

$$\Delta a(t) = -\theta \frac{\partial E_{\text{plan}}}{\partial a(t)} \text{ where } E_{\text{plan}} = \sum_{t} \frac{1}{2} (S_{\text{goal}} - S_{\text{pred}})^2$$

(2)

5. Vision-Directed Reaching

With the look-ahead planning, the robot will either overstep or stop one step earlier and will rarely stop precisely on the target. Also, it suffers from accumulated errors during the long-depth planning. To solve this, a second planning strategy that is based on object recognition is adopted. It will be activated after the execution of the look-ahead planning, in which case the robot is supposed to be in the neighboring area the target. After perceiving the target, the robot adjusts its head direction to keep the target in the center of its view and moves towards it. A flowchart of the process is provided in Fig. 3.
6. Experiment results

To test the look-ahead planning, we used RatLab to perform simulation experiments, RatLab also generates images for training place- and HD cell networks. In this work, we test the first planning strategy of our navigation system in RatLab. Fig.4(a) shows an overview of the simulated rectangle environment and we also present an image captured by the virtual robot from a given position with a random head direction in the environment.

For the world spatial representation, we trained 30 place cells and 50 HD cells, which are shown in Figure 4(b). For the training data collection, we recorded the data from the simulator. In this work, we let the robot to move forward with a constant step size and controlled the robot by changing its moving direction. We collected 25000 state transitions in total, which were split randomly into 4:1 ratio for training and testing. With this data, we can train a world model to represent state transitions. Planning based on such a model only aims to generate trajectories to the target neighboring areas, rather than the precise position. Moreover, its planning result will be further improved by a second planning strategy.

![Diagram](image)

**Fig. 4** (a) An overhead view of the Ratlab environment. (b) The learned Place and HD cells.

To evaluate the performance of the look-ahead planning, a fixed starting position is first chosen and 120 positions are also sampled from the environment as the target. Fig.5(a) shows that the prediction error increases with the planning step. Fig.5(b) shows that the planning performance with respect to the moving distance. As we can see, when the target lies behind the second obstacle (the top-right part of the environment), planning almost fails. This is due to the accumulation error in the long world model chain.
Furthermore, we test the vision-based target approaching with a Turtlebot3 robot performing goal reaching tasks. During testing, the robot's starting position is about 2m away from the target. For recognizing the target, we used the YOLO-V3 network to recognize, classify and localize objects. The experiment result is shown in Fig.6. To quantitatively evaluate the performance of the goal-directed reaching, we select several different objects from the environment and conduct ten tests for each object. The results of the experiments are summarized in Fig.6. As we expect, the success rate relies on the performance of the object detection module. For example, some target objects like the book and the cup can be reached with a high success rate because these objects can be easily recognized by the YOLO detector. The orange has a smaller success rate because the detector sometimes recognizes it as a "sports ball", considering their similar appearance. While the success rate for the box is 0 because the detector fails to detect it.

7. Conclusion

In this paper, we have proposed a navigation system that involves a hybrid navigation strategy in reaching a target location, where two planning strategies are both based on vision but work on different distance scales. The first one performs look-ahead planning on a global coordinate system and the second one is a target approaching strategy working on a local scale, which performs goal-directed reaching based on object recognition. With these two strategies, the robot is able to perform precise navigation using just its vision system.
References

[1] Moser E.I, Kropff E, Moser M.B 2008 Place cells, grid cells, and the brain-spatial representation system Annual Review of Neuroscience 31
[2] Solstad T, Boccara C.N, Kropff E, Moser M.B and Moser E.I 2008 Representation of geometric borders in the entorhinal cortex Science 322 pp1865-1868.
[3] Chen L.L, Lin L.H, Green E.J, Barnes C.A and McNaughton B.L 1994 Head-direction cells in the rat posterior cortex Experimental Brain Research 101
[4] Markus E.J, Qin Y.L, Leonard B, Skaggs W.E, McNaughton B.L and Barnes C.A 1995 Interactions between location and task affect the spatial and directional firing of hippocampal neurons. Journal of Neuroscience 15 pp 7079-7094.
[5] Robitsek R.J, White J.A and Eichenbaum H 2013 Place cell activation predicts subsequent memory. Behavioural Brain Research 254 pp 65-72.
[6] Pfeiffer B.E, Foster D.J 2013 Hippocampal place-cell sequences depict future paths to remembered goals Nature 497 p 74