

The Next Phase for Tracking and Predicting the Navigational Behavior Using Machine Learning

Susumu Takahashi
Graduate School of Brain Science
 Doshisha University
 Kyoto, Japan
 stakahas@mail.doshisha.ac.jp

Satoshi Hoshino
Graduate School of Brain Science
 Doshisha University
 Kyoto, Japan
 dddb1001@mail4.doshisha.ac.jp

Kaoru Ide
Graduate School of Brain Science
 Doshisha University
 Kyoto, Japan
 kide@mail.doshisha.ac.jp

Abstract—Animals navigate through their ways in various scales ranging from centimeters to thousands of kilometers. How does the brain realize such spatial navigation? Several lines of evidence have suggested that the hippocampal place cell activity in the brain has a potential to answer the question. Although statistical machine learning plays critical roles on precisely deciphering the nature of the place cell activity, some technical issues remain unsolved since the ground truth is missing. By virtue of progressive efforts together with the advance on machine learning, methodologies for tracking and predicting navigational behaviors have been improved. Here we review the popular methodologies based on statistical machine learning that track the animal behavior from images and predict it from place cell activity, and discuss about what the next phase of the analysis tools is to deeply understand the neuronal underpinning of spatial navigation.

Keywords—Place cell, Hippocampus, Deep learning, Bayesian decoder, Markerless tracking

I. INTRODUCTION

Spatial navigation including homing, migration, foraging and exploring is considered a common behavior widely observed across several species. In the field of neuroscience, to elucidate the neuronal underpinning of spatial navigation, the relationship between animal's trajectory and neuronal activity have been investigated. For instance, it is well known that neurons maximally firing at a specific location, called place cell [1], were discovered according to the procedure. In addition, the damage of the hippocampus in the brain leads to severe dysfunction on spatial navigational ability in animals including humans [2]. The hippocampus is therefore thought to be critically involved in spatial navigation.

How has the relationship between the place cell activity and animal's location been examined in practice? First, animal's behaviors were captured from a video camera mounted on the ceil of an experimental room. During the experiment, a few high visibility markers were attached on the animal's head to easily track the head's location and direction from the video image sequences. Second, the position of the markers were detected and used for reconstructing the animal's trajectory. Third, hippocampal neuronal activity was simultaneously recorded from extracellular electrodes. The timing between the trajectory and neuronal activity was synchronized through a common digital signal. Finally, the spatial correlation between neuronal activity and the animal's location was examined.

In the analyses, there are some technical issues. For the tracking, the markers hinder animal's natural behavior. Therefore, markerless tracking software showing near human-

level performance is anticipated. For the prediction, neuronal decoder [3] is widely used.

It is possible that together with the advances on machine learning, their performance will be improved especially in difficult situation. Here we review the recent advances and technical issues on the analyses based on statistical machine learning, and discuss about their future perspective.

II. TRACKING THE NAVIGATIONAL BEHAVIOR FROM IMAGES

A. Tracking the animal's head location with markers

The initial step for examining the place specificity of firings of neurons in the brain is to observe where and when the animal ran. Neuroscientists focus on the head's location and direction because animals are thought to recognize the environment in a head centered coordination system. Normally, to easily track the head location, high visibility markers such as LED were attached on the head. Those makers are easily detectable using simple image processing functions including thresholding at an arbitrary value even in a selected color channel. However, the markers could affect the animal behavior; for example, the reflected light from the LED could be visual cues dynamically changing according to the animal's head movement; those light would affect the place cell firings as the previous reports suggested that the receptive field covered by a place cell firings can be remapped in response with the visual cues [4].

B. Tracking using conventional image processing

Image processing highly customized to each animals could track the head location. However, some parameters for the processing need to be optimized every experimental conditions by a human operator. Nevertheless neuroscientists are typically not familiar with image processing, the expertized knowledge is often required for those optimizations. Thus, the simple object-based tracking without sophisticated image processing has been widely used in the place cell studies.

C. Markerless posture tracking based on deep learning

Recently, machine vision based on deep learning shows an outstanding performance on the object detection. In the object detection area, the performance is going away beyond the human expertise. However, neuroscientists have not fully gained the great benefit from the remarkable performance powered by deep learning. The primary reason is that any cutting-edge software for helping to understand the questions in the field of neuroscience using the deep learning have not been developed yet. The second reason is that almost all of

experimental neuroscientists are not familiar with the state-of-the-art techniques of the machine learning.

A novel open source software recently developed for marker-less posture estimation based on deep learning, called *DeepLabCut* [5], may be a game-changer. Typically, a huge amount of labelled images are a prerequisite for training deep convolution networks. The *DeepLabCut* overcame the problem using transfer learning. It uses Resnet-50 [6], a deep residual network with 50 layers trained from ImageNet as a pretrained initial network. Thus, it achieved that using only 200 labelled images randomly selected, it is capable of accurately tracking the target body joints such as neck, head, and leg. It can also track the posture of multiple objects by virtue of pair-wise probability maps inspired from the *DeeperCut* [7], a multi-person position estimation model. Because the outstanding performance on the object detection, post-processing such as Kalman filter frequently used in machine vision is not necessarily required.

Indeed, using the *DeepLabCut*, we could accurately estimate postures from several species including rat, mouse, seabird and salmonid fish. Figure 1 demonstrates an example of head, center of body and tail of a rat tracked using the *DeepLabCut*.

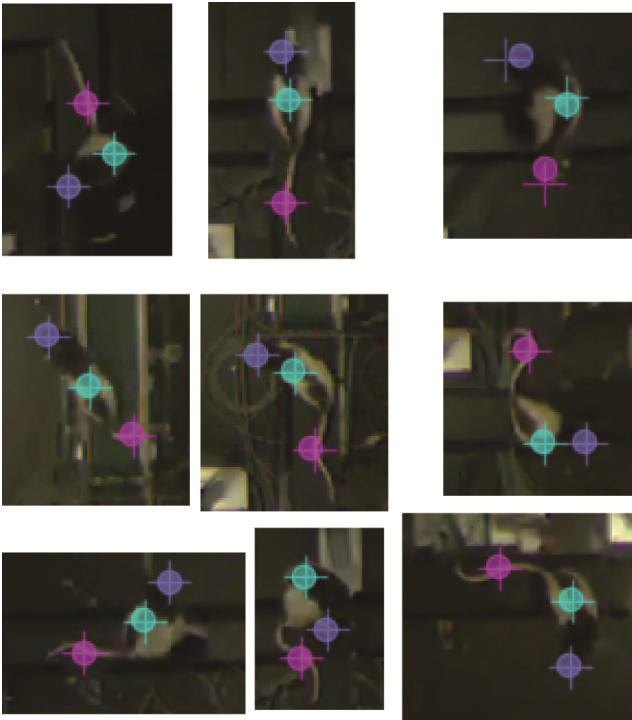


Fig. 1. Nine example images of the head (purple), center of body (blue) and tail (pink) of a rat tracked using the *DeepLabCut*.

D. Beyond the *DeepLabCut*

Whereas the *DeepLabCut* must be the first candidate for tracking the animal behavior, there is still room for improvement as the developer mentioned in the article.

DeepLabCut based on the deep convolution network for image detection does not consider the temporal dynamics of the posture. It is well known that Kalman filter, conventional image processing utilizes transition probability between pre and post images in the video. The temporal information must

help to track the animal behavior under situations where objects were occasionally hidden for a short period. In the field of deep learning, long short term memory (LSTM) can accurately capture such temporal sequences at a high precision. *DeepLabCut* in conjunction with LSTM would be the next phase of evolution on tracking the navigational behavior.

III. PREDICTING THE NAVIGATIONAL BEHAVIOR

How does the brain realize flexible spatial navigational behavior? The place cell in the hippocampus and its relatives including head-direction cell [8], grid cell [9], border cell and speed cell [10] in the parahippocampal areas and entorhinal cortex are thought to play important roles in the spatial navigation because they are the key components for path-integration [11]. In addition, the place cell is an excitatory pyramidal neuron in the hippocampus, which fires only at a specific location in an environment. Thus, provided that we are capable of simultaneously monitoring over 100 place cells, the animal's running trajectory can be precisely decoded from the place cell activity sequences using statistical machine learning [12].

A. Bayesian decoder

One of the most popular methodologies for predicting animal's trajectory from the place cell activity is called Bayesian decoder [3]. It is an algorithm that reconstructs the position of the animal from the place cell activities by calculating $P(\mathbf{x}|\mathbf{n})$, the conditional probability for the animal to be at discrete position \mathbf{x} given by the number of spikes, \mathbf{n} . Since the $P(\mathbf{x}|\mathbf{n})$ cannot be directly calculated, it must be estimated using probabilities derived from experimentally measurable values according to the Bayes theorem as show in (1).

$$P(\mathbf{x}|\mathbf{n}) = \frac{P(\mathbf{n}|\mathbf{x})P(\mathbf{x})}{P(\mathbf{n})} \quad (1)$$

where $\mathbf{x} = (x, y)$, $\mathbf{n} = (n_1, n_2, \dots, n_N)$ denote the position of the animal and the number of spikes fired by recorded cells within a time window, respectively. $P(\mathbf{x})$ represents a probability that the animal occupied a position \mathbf{x} in the environment, called occupancy map. $P(\mathbf{n})$ indicates the firing rate.

To evaluate the key probability, $P(\mathbf{n}|\mathbf{x})$, we assume that the spikes have a Poisson distribution and that place cells are statistically independent each other. Then, the final formula is described in (2).

$$P(\mathbf{x}|\mathbf{n}) = C(\tau, \mathbf{n}) P(\mathbf{x}) \left(\prod_{i=1}^N f_i(\mathbf{x})^{n_i} \right) \exp \left(-\tau \sum_{i=1}^N f_i(\mathbf{x}) \right) \quad (2)$$

where τ denotes the length of a time window, $C(\tau, \mathbf{n})$ represents a normalization constant that depend on the time window, τ and the number of spikes, \mathbf{n} , $f_i(\mathbf{x})$ represents firing rates of place cells at the position \mathbf{x} , and n_i represents the number of spikes of cell i . Since the probability $P(\mathbf{x}|\mathbf{n})$ has a two-dimensional distribution, the reconstructed position is defined as a peak of the probability map. Similar to the *DeepLabCut*, the reconstructed positions can be improved with history of past positions using Kalman filter. However, the place cell community does not recommend to use it because it is possible that such artificial smoothing distorts the results obtained from the nature of place cell activity.

B. Animal's trajectory predicted by the Bayesian decoder

The memoryless Bayesian decoder is widely used to demonstrate and evaluate the relationship between spatial navigational behavior and hippocampal place cell activity. For instance, some sequential activity patterns of place cell are reactivated at a ten-fold compression scale during slow-wave sleeps, one of the stages of sleep [13], and a brief immobility period [14]. Since the decoder enables us to visually assess the reactivated spatial trajectory patterns encoded by a place cell activity sequence, it helps us to understand the neuronal mechanism of spatial navigation [15] (Fig. 2).

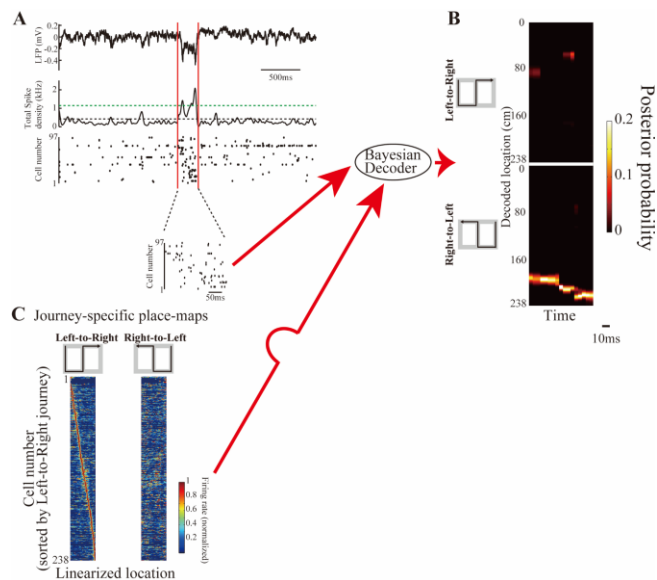


Fig. 2. A procedure for predicting animal's trajectory using the Bayesian decoder. **A**, Candidate reactivation were defined as periods when the smoothed (Gaussian kernel; SD: 10 ms) population activity (middle) was higher than the mean (black dotted line) and the peak was above the defined threshold (mean + 3SD, green dotted line). During candidate replays, the local field potential (LFP) (top) clearly showed sharpwave ripples. Location were estimated based on the spiking activity (bottom, spike raster) enclosed by the red lines in the reactivation. **B**, The memoryless Bayesian decoder was used to decode the posterior probability of the linearized location from spiking activity and journey-specific place maps, **C**, for every time window. Adapted from Takahashi, *eLife*, 2015 [15].

C. Beyond the Bayesian decoder

The reactivation of place cell activity sequences is temporally compressed. Although the Bayesian decoder employs a fixed time window, the compression scale might be dynamically modulated in a natural situation. To address the issue, template matching have been utilized [16]. Dynamic time warping—a sort of template matching—may handle such dynamic changes of the temporal compression rate. However, the computation load is heavy. The combination of the Bayesian decoder and the template matching or hidden Markov models may be a solution.

IV. CONCLUDING REMARKS

In summary, we reviewed the significance and technical issues of methodologies widely used in the field of systems neuroscience for tracking and predicting the navigational behaviors based on statistical machine learning.

Besides, machine learning also plays an important role on different analysis stages. For instance, a preprocessing called

spike sorting that sorts single neuronal activity from multiple ones monitored from extracellular recordings, is degraded by spike overlapping and non-stationary shape. Whereas several spike sorting algorithms have been developed to overcome the issues using independent component analysis [17], EM algorithm [18], a standard tool has not been established yet [19].

We believe that cutting-edge technologies based on state-of-the-art statistical machine learning will address the issues in the near future. To accomplish that, data scientist and neuroscientist must work together toward a common goal.

ACKNOWLEDGMENT

This work is supported by JSPS KAKENHI JP16H06543.

REFERENCES

- [1] J. O'Keefe and J. Dostrovsky, "The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat," *Brain Res*, vol. 34, no. 1, pp. 171–175, 1971.
- [2] W. B. Scoville and B. Milner, "Loss of recent memory after bilateral hippocampal lesions," *J Neurol Neurosurg Psychiatry*, vol. 20, no. 1, pp. 11–21, 1957.
- [3] K. Zhang, I. Ginzburg, B. L. McNaughton, and T. J. Sejnowski, "Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells," *J. Neurophysiol.*, vol. 79, pp. 1017–1044, 1998.
- [4] S. Leutgeb and J. K. Leutgeb, "Pattern separation, pattern completion, and new neuronal codes within a continuous CA3 map," *Learn. Mem.*, vol. 14, pp. 745–757, 2007.
- [5] A. Mathis *et al.*, "DeepLabCut: markerless pose estimation of user-defined body parts with deep learning," *Nat. Neurosci.*, vol. 21, no. 9, pp. 1281–1289, 2018.
- [6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [7] E. Insafutdinov, L. Pishchulin, B. Andres, M. Andriluka, and B. Schiele, "Deepercut: A deeper, stronger, and faster multi-person pose estimation model," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2016.
- [8] 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," *J. Neurosci.*, vol. 10, pp. 420–435, 1990.
- [9] M. Fyhn, S. Molden, M. P. Witter, E. I. Moser, and M. B. Moser, "Spatial representation in the entorhinal cortex," *Science (80-.)*, vol. 305, no. 5688, pp. 1258–1264, 2004.
- [10] E. Kropff, J. E. Carmichael, M.-B. Moser, and E. I. Moser, "Speed cells in the medial entorhinal cortex," *Nature*, vol. 523, pp. 419–424, 2015.
- [11] B. L. McNaughton, F. P. Battaglia, O. Jensen, E. I. Moser, and M. B. Moser, "Path integration and the neural basis of the 'cognitive map,'" *Nat Rev Neurosci*, vol. 7, no. 8, pp. 663–78, 2006.
- [12] M. A. Wilson and B. L. McNaughton, "Dynamics of the hippocampal ensemble code for space," *Science (80-.)*, vol. 261, no. 5124, pp. 1055–1058, 1993.
- [13] M. A. Wilson and B. L. McNaughton, "Reactivation of hippocampal ensemble memories during sleep," *Science (80-.)*, vol. 265, no. 5172, pp. 676–679, 1994.
- [14] D. J. Foster and M. A. Wilson, "Reverse replay of behavioural sequences in hippocampal place cells during the awake state," *Nature*, vol. 440, pp. 680–683, 2006.
- [15] S. Takahashi, "Episodic-like memory trace in awake replay of hippocampal place cell activity sequences," *Elife*, vol. 4, 2015.
- [16] Z. Nádasdy, H. Hirase, A. Czurró, J. Csicsvari, and G. Buzsáki, "Replay and time compression of recurring spike sequences in the hippocampus," *J. Neurosci.*, vol. 19, no. 21, pp. 9497–9507, 1999.
- [17] S. Takahashi, Y. Anzai, and Y. Sakurai, "Automatic sorting for multi-neuronal activity recorded with tetrodes in the presence of overlapping spikes," *J. Neurophysiol.*, vol. 89, no. 4, pp. 2245–58, Apr. 2003.

- [18] K. D. Harris, D. A. Henze, J. Csicsvari, H. Hirase, and G. Buzsaki, "Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements," *J. Neurophysiol.*, vol. 84, no. 1, pp. 401–414, 2000.
- [19] M. S. Lewicki, "A review of methods for spike sorting: the detection and classification of neural action potentials," *Network-computation Neural Syst.*, vol. 9, no. 4, pp. R53–R78, 1998.