



OPEN ACCESS

EDITED AND REVIEWED BY
Alois C. Knoll,
Technical University of Munich,
Germany

*CORRESPONDENCE
Akira Taniguchi
a.taniguchi@aem.ci.ritsumeai.ac.jp

RECEIVED 23 October 2022
ACCEPTED 02 November 2022
PUBLISHED 15 November 2022

CITATION
Taniguchi A, Spranger M, Yamakawa H
and Inamura T (2022) Editorial:
Constructive approach to spatial
cognition in intelligent robotics.
Front. Neurobot. 16:1077891.
doi: 10.3389/fnbot.2022.1077891

COPYRIGHT
© 2022 Taniguchi, Spranger,
Yamakawa and Inamura. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Editorial: Constructive approach to spatial cognition in intelligent robotics

Akira Taniguchi^{1*}, Michael Spranger², Hiroshi Yamakawa^{3,4,5}
and Tetsunari Inamura^{6,7}

¹College of Information Science and Engineering, Ritsumeikan University, Kyoto, Japan, ²Sony Computer Science Laboratories, Tokyo, Japan, ³The Whole Brain Architecture Initiative, Tokyo, Japan, ⁴School of Engineering, The University of Tokyo, Tokyo, Japan, ⁵Center for Biosystems Dynamics Research, Institute of Physical and Chemical Research (RIKEN), Osaka, Japan, ⁶Principles of Informatics Research Division, National Institute of Informatics, Tokyo, Japan, ⁷Department of Informatics, The Graduate University for Advanced Studies, SOKENDAI, Tokyo, Japan

KEYWORDS

simultaneous localization and mapping, spatial reasoning, place recognition and categorization, navigation and path planning, spatial language understanding

Editorial on the Research Topic

Constructive approach to spatial cognition in intelligent robotics

1. Introduction

For agents operating in the real world, spatial reasoning and understanding of the spatial properties of the environment are important abilities for executing tasks related to spatial movement (Kostavelis and Gasteratos, 2015; Garg et al., 2020). Agents can perform various tasks by acquiring and utilizing the semantic and linguistic knowledge related to place and object locations. Current research on spatial reasoning and semantic understanding in robots is important for realizing self-localization with uncertainty in the real world and planning with human–robot interactions. It is also closely related to a constructive approach to brain-inspired AI related to spatial cognition represented by the hippocampal formation (Tolman, 1948; O’keefe and Nadel, 1978).

In this Research Topic, we address the interdisciplinary fusion of knowledge of artificial intelligence, robotics, cognitive development, and neuroscience in spatial cognition and spatial reasoning. For example, in the fusion area of natural language processing and computer vision, research on vision-and-language navigation (VLN) has recently been implemented (Anderson et al., 2018; Chen et al., 2021).

However, few studies have applied VLN technology to real-world robots. In the future, a VLN that operates in a real environment will be a necessity. Additionally, in robotics AI, referring to the cognitive and neuroscientific findings of concept formation related to place and spatial language acquisition would be useful. To achieve this, a constructive approach using robots operating in the real world would be effective (Asada et al., 2009; Cangelosi and Schlesinger, 2015; Taniguchi et al., 2019).

This Research Topic was widely assembled from fundamental to applied research, which is related to spatial reasoning using robots and semantic understanding including language interaction and in the fusion area of artificial intelligence, such as machine learning, robotics, and computational neuroscience. These papers contribute to advancing the field on the technical side, for example, semantic mapping, place recognition, and navigation for performing tasks including spatial movement. Additionally, some studies contribute to computational models related to spatial reasoning, such as referring to hippocampal formation and spatial cognitive capabilities. Furthermore, the focus is on contributions to cutting-edge machine learning for use in the above aspects.

2. About the Research Topic

We are pleased to present five research articles related to spatial relation learning, spatial concept formation, bio-inspired models, localization, and navigation. In this section, we briefly introduce each paper.

Autonomous mobile robots and self-driving vehicles require accurate and reliable self-localization to handle dynamic environments. Colomer et al. address the problem of visual location awareness using a neuro-cybernetic approach. The proposed method is a biologically-inspired neural network called the log-polar max-Pi (LPMP) model. In particular, the visual-spatial processing associated with the hippocampus and entorhinal cortex is referenced. A mechanism is constructed to integrate information from two parallel pathways, the “what” and “where” pathways of the visual cortex. The localization performance is evaluated in a road environment, demonstrating its usefulness compared to conventional methods.

Grid cells in the medial entorhinal cortex of the mammalian brain are essential for the path integration and representation of the external world (Hafting et al., 2005; McNaughton et al., 2006). However, it has been argued that few models explain the formation of various grid-field structures in recent relevant studies. To fill this gap, Gong and Yu propose an interpretative plane-dependent model of 3D grid cells to represent both 2D and 3D spaces. The proposed method comprises a spatial transformation using 6-DOF motion and a recurrent neural network (RNN) for 3D grid cells. In simulation experiments, a representation similar to that of hexagonal grid cells is reported to be obtained. The results validate the hypothesis that “grid fields gradually lose global but not local order as the refresh interval decreases.”

For agents to perceive and act in a physical space, learning relevant concepts about space and its environment (objects, colors, shapes, etc.) is essential. Lee et al. propose

three approaches to enable cognitive agents to learn spatial relations. The proposed approach integrates (i) the learning of language-spatial relations through embodied experience, (ii) the learning of directional relations using large-scale image data, and (iii) the inference of spatial relations using a knowledge base. For learning, an upper-body humanoid robot and a neural network-based model are used. Partial experiments for each component of the proposed approach demonstrate its applicability. The authors present the concept of an integrated architecture, which still needs to be implemented and validated.

Humans are assumed to recognize continuous high-dimensional observations by segmenting and classifying them into significant segments. Nagano et al. propose HcVGH, a method for learning spatio-temporal categories by segmenting first-person-view videos captured by a mobile robot. HcVGH combines a convolutional variational auto-encoder (cVAE) with the authors' previous model, a hierarchical Dirichlet process-variational autoencoder-Gaussian process-hidden semi-Markov model (HVGH) (Nagano et al., 2019), based on a probabilistic generative model. HVGH is an unsupervised segmentation method for high-dimensional time series. The experimental results show that the proposed method is capable of classifying and segmenting sequences of robot perspective videos with high accuracy in a simulation environment. The transition probabilities estimated by HcVGH can be used for global path planning, potentially enabling the planning of different paths for different purposes.

Sentences containing spatial instructions from the user to the robot are not only based on a coordinate system that is absolute concerning the environment, such as “kitchen.” Some instructions relate to relative positions, such as “next to the chair” or “in front of the TV.” Relative spatial concepts are based on a coordinate system relative to an object or agent. In a previous study on spatial concept formation (Taniguchi et al., 2017), the main focus was only on absolute concepts. In contrast, the method proposed by Sagara et al. enables the robot to simultaneously estimate the coordinate system and spatial concepts (absolute and relative). The relative and absolute spatial concept acquisition method (RASCAM) is based on the model of the authors' previous work ReSCAM+O (Sagara et al., 2021) and the model of absolute spatial concept formation (Taniguchi et al., 2017), which is composed of a probabilistic generative model. Experiments in a simulated home environment show that the proposed approach can learn relative and absolute spatial concepts while accurately selecting the coordinate system. This approach will help the service robot flexibly understand new environments through human interaction.

We hope the above articles will interest the readers in the recent efforts in spatial cognition in

robots and highlight the importance of this research field.

3. Next step

Many challenges remain in further developing spatial cognition and spatial semantic understanding using robots.

The first problem is that models are still often tested by simulators. To achieve intelligent models that perform robustly over the long term, they should be tested in dynamic real-world environments with various observational noises and uncertainties. In particular, real-world applications of VLNs will rapidly advance.

The second issue is spatial linguistic semantic understanding. As discussed in a paper on our topic, learning the relationship between spatial concepts and language is an interesting field for further development. Moreover, applying large-scale language models to real-world environments is an issue to be addressed in the future.

The third challenge is the integration of advanced machine learning theory with neuroscientific findings. For example, we expect brain-inspired models such as HF-PGM (Taniguchi et al., 2022) to be demonstrated on real robots. The relationship between the free-energy principle (Friston et al., 2012) and spatial cognition in robots is also very interesting. As discussed in several papers on our topic, it would be useful, from an engineering point of view, to construct brain-referenced autonomous intelligence.

References

- Anderson, P., Wu, Q., Teney, D., Bruce, J., Johnson, M., Sünderhauf, N., et al. (2018). "Vision-and-language navigation: interpreting visually-grounded navigation instructions in real environments," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Salt Lake City, UT: IEEE), 3674–3683.
- Asada, M., Hosoda, K., Kuniyoshi, Y., Ishiguro, H., Inui, T., Yoshikawa, Y., et al. (2009). Cognitive developmental robotics: a survey. *IEEE Trans. Auton. Mental Dev.* 1, 12–34. doi: 10.1109/TAMD.2009.2021702
- Cangelosi, A., and Schlesinger, M. (2015). "Developmental robotics: from babies to robots," in *Intelligent Robotics and Autonomous Agents Series* (Massachusetts: MIT Press).
- Chen, K., Chen, J. K., Chuang, J., Vázquez, M., and Savarese, S. (2021). "Topological planning with transformers for vision-and-language navigation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Nashville, TN: IEEE), 11271–11281.
- Friston, K. J., Adams, R. A., Perrinet, L., and Breakspear, M. (2012). Perceptions as hypotheses: saccades as experiments. *Front. Psychol.* 3, 151. doi: 10.3389/fpsyg.2012.00151
- Garg, S., Sünderhauf, N., Dayoub, F., Morrison, D., Cosgun, A., Carneiro, G., et al. (2020). Semantics for robotic mapping, perception and interaction: a survey. *Foundat. Trends® Rob.* 8, 1–224. doi: 10.1561/9781680837698
- Hafting, T., Fyhn, M., Molden, S., Moser, M. B., and Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature* 436, 801–806. doi: 10.1038/nature03721
- Kostavelis, I., and Gasteratos, A. (2015). Semantic mapping for mobile robotics tasks: a survey. *Rob. Auton. Syst.* 66:86–103. doi: 10.1016/j.robot.2014.12.006
- McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I., and Moser, M. B. (2006). Path integration and the neural basis of the 'cognitive map'. *Nature Rev. Neurosci.* 7, 663–678. doi: 10.1038/nrn1932
- Nagano, M., Nakamura, T., Nagai, T., Mochihashi, D., Kobayashi, I., and Takano, W. (2019). HVGH: unsupervised segmentation for high-dimensional time series using deep neural compression and statistical generative model. *Front. Rob. AI* 6, 115. doi: 10.3389/frobt.2019.00115
- O'Keefe, J., and Nadel, L. (1978). *The Hippocampus as a Cognitive Map, Vol. 27*. Cambridge: Cambridge University Press.
- Sagara, R., Taguchi, R., Taniguchi, A., Taniguchi, T., Hattori, K., Hoguro, M., et al. (2021). Unsupervised lexical acquisition of relative spatial concepts using spoken user utterances. *Adv. Rob.* 36, 54–70. doi: 10.1080/01691864.2021.2007168
- Taniguchi, A., Fukawa, A., and Yamakawa, H. (2022). Hippocampal formation-inspired probabilistic generative

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Acknowledgments

The authors gratefully acknowledge the contributions of participants in this Research Topic.

Conflict of interest

Author MS was employed by Sony Computer Science Laboratories. Author HY was employed by The Whole Brain Architecture Initiative.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

model. *Neural Netw.* 151:317–335. doi: 10.1016/j.neunet.2022.04.001

Taniguchi, A., Hagiwara, Y., Taniguchi, T., and Inamura, T. (2017). “Online spatial concept and lexical acquisition with simultaneous localization and mapping,” in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Vancouver, BC: IEEE), 811–818.

Taniguchi, T., Piater, J., Worgotter, F., Ugur, E., Hoffmann, M., Jamone, L., et al. (2019). Symbol emergence in cognitive developmental systems: a survey. *IEEE Trans. Cogn. Dev. Syst.* 11, 494–516. doi: 10.1109/TCDS.2018.2867772

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208. doi: 10.1037/h0061626