Experiences and Open Questions on using Vector Symbolic Architectures for Mobile Robotics

Peer Neubert, Kenny Schlegel, Stefan Schubert, and Peter Protzel

Chemnitz University of Technology, Germany
`firstname.lastname@etit.tu-chemnitz.de`
https://www.tu-chemnitz.de/etit/proaut/en/

Abstract. We present a selection of our work in the area of mobile robotics that is related to hyperdimensional computing and Vector Symbolic Architecture and list open question for further discussion.

1 High-dimensional encodings of real world robot data

A central area of our work in the field of mobile robotics is state estimation and in particular localization problems in changing environments. For example, image-based recognition of a place that is known from a sunny summer day when returning at a cloudy winter evening. The essential source of information for this task is pairwise similarity of image descriptors. It turned out, that high-dimensional vectors from (multi-purpose) deep-learning systems can be surprisingly robust against such changing conditions. We used such descriptors in several of our works on visual place recognition, e.g. [4, 7].

Open questions include: •What are properties of such image descriptors from a VSA perspective? The distribution of pairwise distances is different from that of random high-dimensional vectors, nevertheless we have been able to successfully apply VSA algorithms on these descriptors (next section). •How can we encode other sensor or state information like odometry or 6D poses? So far, we conducted some experiments with grid-cell encodings [6]. •Can we address the particular setup of changing environments in a "vector symbolic" way? For example, in recent work, we describe how to remove part of the influence of the environmental condition from the high-dimensional vector based on statistical methods [7]. Is there a way to combine this with vector symbolic approaches?

2 Encoding context using hyperdimensional computing and Hierarchical Temporal Memory (HTM)

Mobile robot place recognition provides more structural information than plain image retrieval, in particular, database and query are often sequences of images (aka. videos). Exploiting the additional information of neighbored images can significantly improve place recognition performance. In [3] we present a simple VSA implementation of the SeqSLAM approach. Basically, we encode the sequential context of images in high-dimensional vectors based on bundling image descriptors which are bound to their position in the sequence.

To address more complex context patterns we adapted the higher order sequence memory of Hierarchical Temporal Memory (HTM) [1], a theory on working principles of the human neocortex that builds on high-dimensional sparse distributed representations. Again, the idea is to encode sequential context in a single descriptor [2]. With this system we have been able to surpass state-of-the art performance on several place recognition datasets [4]. We also extended this system with a grid-cell encoding to include odometry information [6]

Open questions include: •What are potential and limitations in using VSA for encoding temporal and spatial context of real world data? •We have promising preliminary results [3] on bundling descriptors of multiple views of an object from multiple perspectives in order to associate new views to the superposed representation. Is there a VSA way of extending this to object models which can be continuously updated with new views? •If superpositions are used to encode context, what are the potential and limits of evaluating context without resolving the superposition?

3 Experimental comparison of VSA implementations

In our experience, novices of this field find it challenging to recognize the differences in the VSA operators directly from the primary literature. What is missing is an overview of available VSAs and their properties. We partialy address this issue in [5] by comparing properties of 8 VSAs from the literature and presenting experimental results on the performance of their operators on different tasks.

Open questions include: •The available binding operators have significantly different properties. Are there any efforts towards a unifying (axiomatic?) definition of VSAs? •What are other benchmark tasks that reveal important properties of VSAs? •Which additional VSAs are possible (or exist already)? What would be additional beneficial properties of (novel) VSAs?

References

1. Hawkins, J., Ahmad, S.: Why neurons have thousands of synapses, a theory of sequence memory in neocortex. Frontiers in Neural Circuits **10**, 23 (2016). https://doi.org/10.3389/fncir.2016.00023
2. Neubert, P., Ahmad, S., Protzel, P.: A sequence-based neuronal model for mobile robot localization. In: Trollmann, F., Turhan, A.Y. (eds.) KI 2018: Advances in Artificial Intelligence. pp. 117–130. Springer International Publishing, Cham (2018)
3. Neubert, P., Schubert, S., Protzel, P.: An Introduction to High Dimensional Computing for Robotics. German Journal of Artificial Intelligence Special Issue: Reintegrating Artificial Intelligence and Robotics, Springer (2019)
4. Neubert, P., Schubert, S., Protzel, P.: A neurologically inspired sequence processing model for mobile robot place recognition. IEEE Robotics and Automation Letters (RA-L) and presentation at Intl. Conf. on Intel. Robots and Systems (IROS). (2019)
5. Schlegel, K., Neubert, P., Protzel, P.: A comparison of vector symbolic architectures. CoRR (2020)
6. Schubert, S., Neubert, P., Protzel, P.: Towards combining a neocortex model with entorhinal grid cells for mobile robot localization. In: 2019 European Conference on Mobile Robots (ECMR). pp. 1–8 (Sep 2019)
7. Schubert, S., Neubert, P., Protzel, P.: Unsupervised learning methods for visual place recognition in discretely and continuously changing environments. In: International Conference on Robotics and Automation (ICRA) (to appear) (2020), https://arxiv.org/abs/2001.08960