School of Psychology and Clinical Language Sciences

University of Reading

Comparison of a reinforcement-learning and a biologically-motivated representation of 3D space

Andrew Glennerster | Alex Muryy

Introduction

If the brain does not use 3D coordinate frames to represent space (ego-centred or allocentric), what does it use instead[3]? Reinforcement learning provides a possible answer. Here, we analyse the representation generated by a recent RL algorithm that learned to navigate to rewarded images[1]. We compare it to a hand-crafted representation (also non-3D) that has more information about the distance of objects[3].

Methods

Zhu et al (2017)

- Zhu et al (2017)[1] trained a network to navigate to a rewarded image (5 views). In the example shown here, this gave rise to 900 stored feature vectors (45 training locations, 4 orientations, 5 goal images).
- It learns a policy (4 actions, moving forward, moving backward, turning left, and turning right) conditioned on both the target image and the current observation.
- We show a tSNE projection of the 900 $\mathbb{R}^{512}$ stored feature vectors
- We train a decoder to use the agent's internal representation to output (x, y) of a chosen observation
- We used this to estimate the midpoint between two learned locations

Relative visual direction (RVD)

- A hand-crafted representation of location is made up of a list (vector) of all the angles between points in a scene[2]
- This can be indexed by the amount of parallax (change) in each angle as the camera moves in different directions
- Using only the lowest parallax elements of the vector (30%) gives a smooth change in the vector with changes in camera location (see tSNE plot)
- It also allows accurate interpolation to recover midpoints between stored locations in the representation

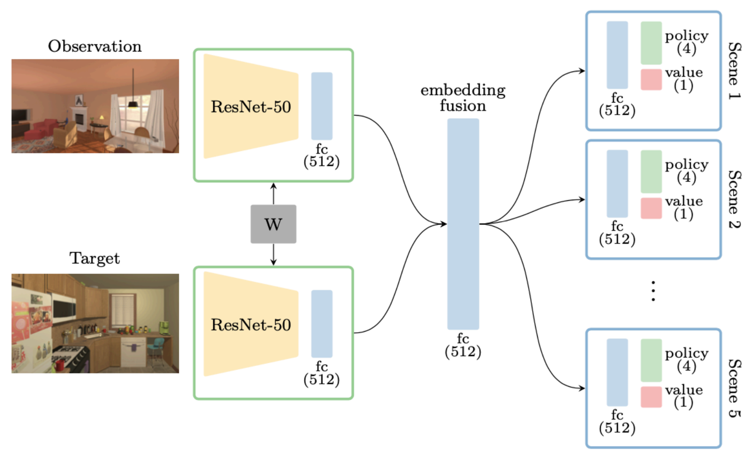
Read the full paper (accepted in *Vision Research*)

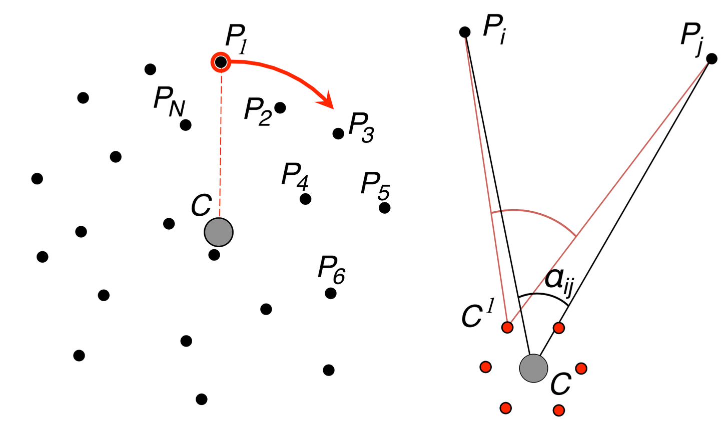
- Muryy, A, Siddharth, N., Nardelli, N., Glennerster, A and Torr, P.H.S. Lessons from reinforcement learning for biological representations of space. https://arxiv.org/abs/1912.06615

Zhu et al (2017)

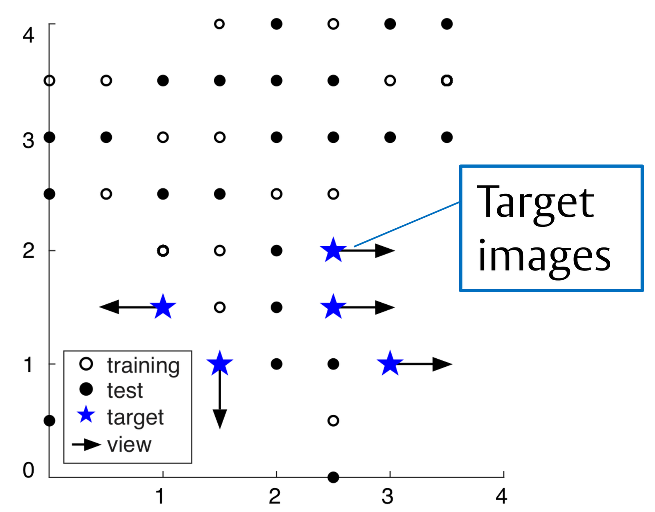
Relative visual direction (RVD)

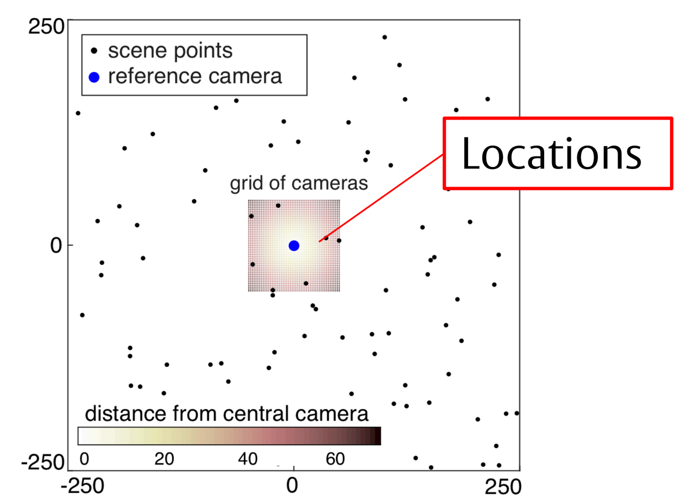
Generation of feature vectors



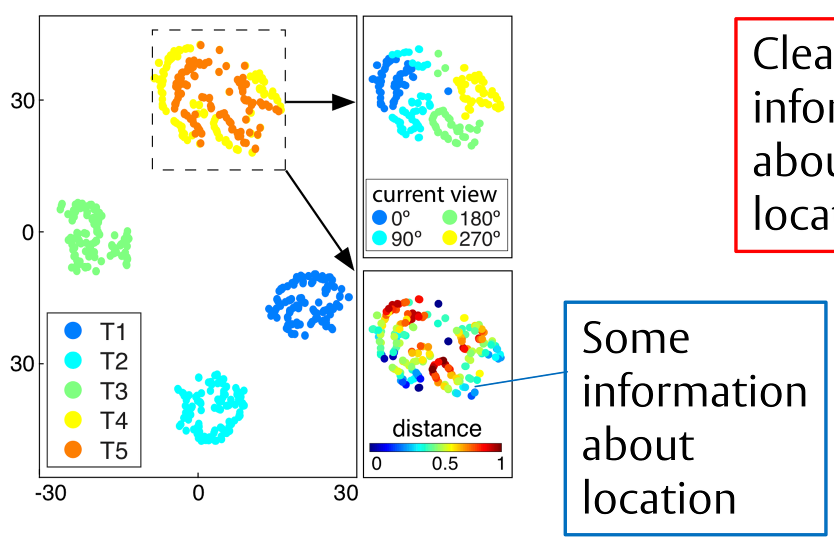


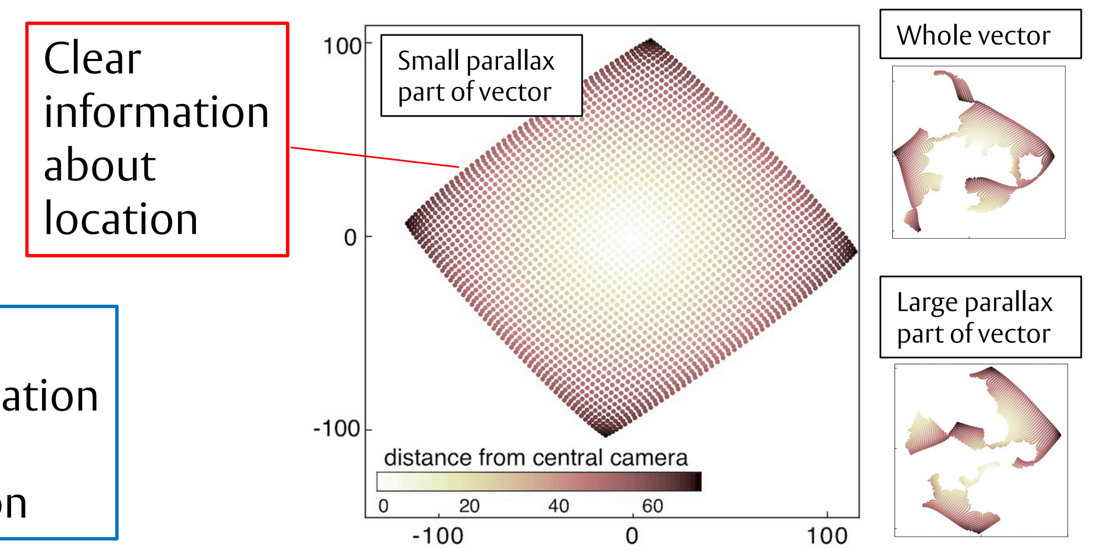
Plan view of scene



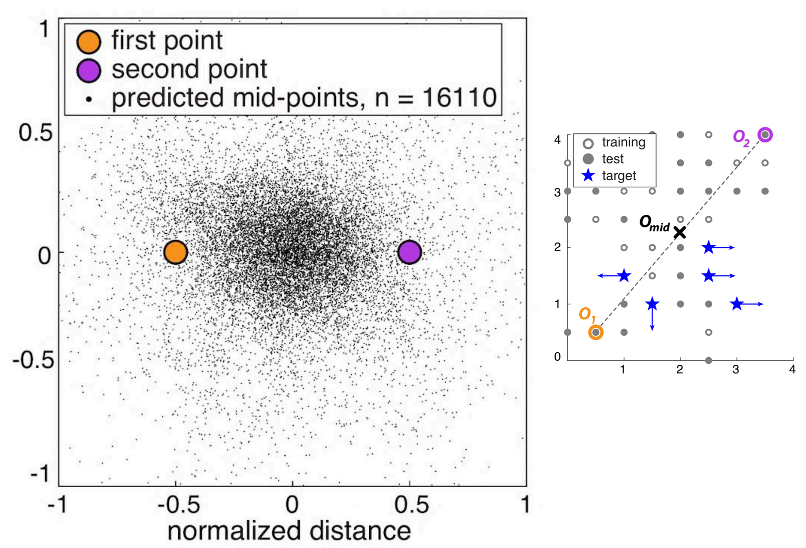


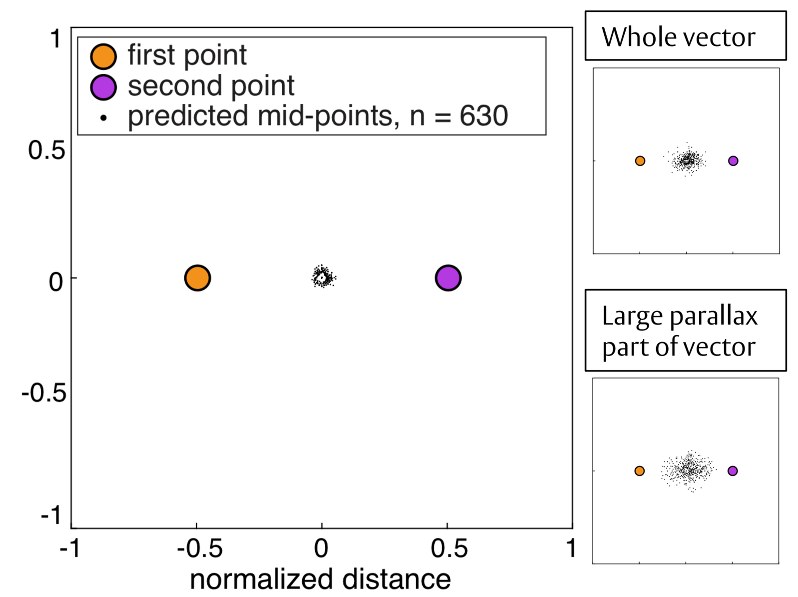
tSNE





Midpoint estimate





Results

- The tSNE plot (which preserves similarity between vectors) from the Zhu et al representation is dominated by task (targets T1 – T5), then camera orientation and finally camera location.
- The tSNE plot for the RVD representation shows a clear structural correspondence between the representation and the camera location (nearby locations have a similar representation)
- This applies only when parallax is used to pick out the elements of the vector that change slowly with camera translation
- As a result, without complex decoding, the RVD model can be used for geometric tasks such as interpolation between locations
- Interpolation using Zhu et al method requires decoding of the representation and it often fails

Conclusion

- **Learned navigation to a rewarded image is possible without a 2D map (either topological or metric)[1]**
- **The system has learned something about 2D location (decoding of location is possible, although not great)**
- **A hand-crafted representation that also uses high dimensional vectors instead of 2D coordinates shows greater geometric consistency (by design)[2]. Parallax is used to identify relevant elements in the vector for this task (and could be used to identify *different* elements for other tasks, e.g. visual control of balance).**
- **Reinforcement learning could, in future, incorporate information about parallax to identify 'persistent' and 'transient' features as the camera translates (providing some of the gains of a 2D map without building one)**

References

1. Zhu, R. Mottaghi, E. Kolve, J. J. Lim, A. Gupta, L. Fei-Fei, A. Farhadi (2017) Target-driven visual navigation …in: Robotics and Automation (ICRA), *IEEE*, 3357–3364
2. Glennerster, A., Hansard, M.E. and Fitzgibbon, A.W. (2001) Fixation could simplify, not complicate, the interpretation of retinal flow. *Vision Research*, **41**, 815-834
3. Glennerster, A. (2016) A moving observer in a three-dimensional world. *Philosophical Transactions of the Royal Society, B*, **371(1697)**, 20150265

Acknowledgement

- Funded by EPSRC/Dstl EP/N019423/1