Poster presentation

Reinforcement learning on complex visual stimuli

Niko Wilbert*[1,2], Robert Legenstein[3], Mathias Franzius[4] and
Laurenz Wiskott[1,2]

Address: [1]Institute for Theoretical Biology, Humboldt-Universität zu Berlin, 10115 Berlin, Germany, [2]Bernstein Center for Computational Neuroscience, Humboldt-Universität zu Berlin, 10115 Berlin, Germany, [3]Institute for Theoretical Computer Science, Graz University of Technology, 8010 Graz, Austria and [4]Honda Research Institute Europe, 63073 Offenbach, Germany

Email: Niko Wilbert* - mail@nikowilbert.de

* Corresponding author

Animals are confronted with the problem of initiating motor actions based on very complex sensory input. We have built a biologically plausible model that uses reinforcement learning on complex visual stimuli to direct an agent towards a target. This is made possible by first extracting a high-level representation of the scene with a hierarchical network and then applying a correlation based RL-learning rule.

The sensory input given to the model consists of grayscale images of size 155 × 155 pixels; see figure. Given this complex input, the model should extract the position and direction of the agent, and the position of the target. This estimation is successfully performed by a multi-layer hierarchical network modeled after the visual system [1]. In each layer, we use Slow Feature Analysis (SFA) [2,3] to efficiently extract higher-level features based on time structure. SFA has the advantage that learning is done unsupervised, just by feeding the model with image sequences. The high-level output of the hierarchical network is then used to learn corresponding motor commands with a reinforcement-learning algorithm. The reward signal is given by the distance to the target, which is the only supervision signal in the whole model (biologically it could be interpreted as a scent of the target). The motor command output is then used to update the scene, so the model runs in a feedback loop. The resulting trajectories (Figure 1) show how the model directs the agent towards its target. Our model demonstrates that by a division-of-labor strategy simple learning rules can solve a rather difficult problem.

References

1. Franzius M, Wilbert N, Wiskott L: **Invariant object recognition with slow feature analysis.** In *Proc 18th Int'l Conf on Artificial Neural Networks* Edited by: *Kurková V, Neruda R, Koutník J. Springer-Verlag*; 2008:961-970.
2. Wiskott L, Sejnowski TJ: **Slow feature analysis: Unsupervised learning of invariances.** *Neural Computation* 2002, **14**:715-770.
3. Zito T, Wilbert N, Wiskott L, Berkes P: **Modular toolkit for data processing (MDP): A Python data processing framework.** *Front Neuroinformatics* 2008, **2**:8.
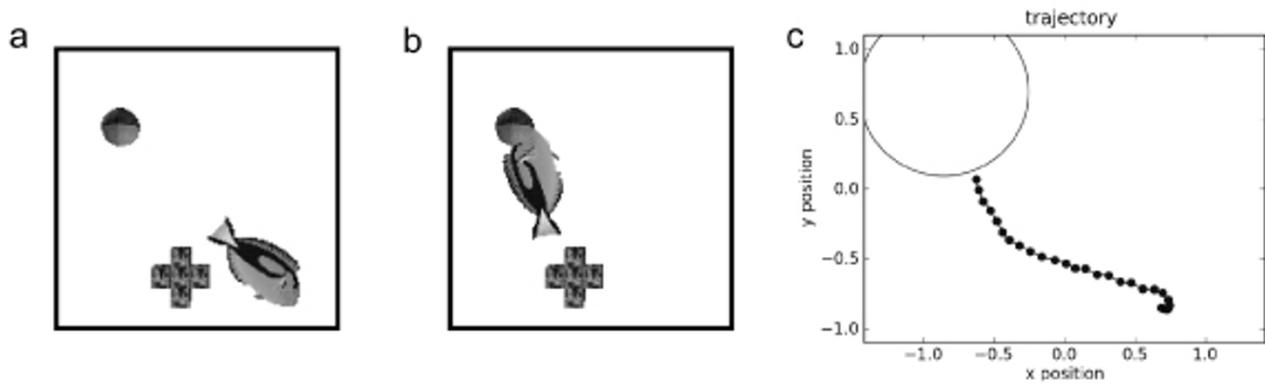
Figure 1
Stimulus images and agent trajectory. (a) shows the first stimulus image given to the model. In this setup the fish is the agent controlled by the model and the sphere is the target. The cube is a distractor and should be ignored. The model is then supposed to rotate the fish towards the target and move forward until the target is hit, as shown in (b). (c) shows the complete trajectory produced by the model (the circle marks the target radius).