

Poster presentation

Open Access

A model of cell specialization using a Hebbian policy-gradient approach with "slow" noise

Emmanuel Dauce^{1,2}

Address: ¹Institute of Movement Sciences University of the Mediterranean, Marseille, France and ²INRIA Lille Nord-Europe, Villeneuve d'Ascq, France

Email: Emmanuel Dauce - edauce@ec-marseille.fr

from Eighteenth Annual Computational Neuroscience Meeting: CNS*2009
Berlin, Germany. 18–23 July 2009

Published: 13 July 2009

BMC Neuroscience 2009, **10**(Suppl 1):P136 doi:10.1186/1471-2202-10-S1-P136

This abstract is available from: <http://www.biomedcentral.com/1471-2202/10/S1/P136>

© 2009 Dauce; licensee BioMed Central Ltd.

We study a model of neuronal specialization using a policy gradient reinforcement approach. (1) The neurons stochastically fire according to their synaptic input plus a noise term; (2) The environment is a closed-loop system composed of a rotating eye and a visual punctual target; (3) The network is composed of a foveated retina directly connected to a motoneuron layer; (4) The reward depends on the distance between the subjective target position and the fovea and (5) the weight update depends on the Hebb-like product $r(t)Z_{ij}(t)$ where $r(t)$ is the reward and $Z_{ij}(t)$ is a Hebbian trace updated according to the product $[S_i(t)-F_i(t)]e_j(t)$, where $S_i(t)$ is the post-synaptic spike, $F_i(t)$ is the firing probability and $e_j(t)$ is the pre-synaptic activity [1,2].

Several temporal scales are to be considered when modeling such neuromimetic controller systems. First, the typical integration time of the neurons is of the order of few milliseconds. Second, the motor commands have a duration on the order of 100 ms. In the design of an adaptive controller, this temporal mismatch must be taken into account.

For that, we consider that the firing probability is monitored by a "pink noise" term whose autocorrelation is of the order of 100 ms, so that the firing probability is overestimated (or underestimated) for about 100 ms periods. The rewards occurring meanwhile assess the "quality" of those elementary shifts, and modify the firing probability accordingly.

Every motoneuron being associated to a particular angular direction, we test at the end of the learning process the preferred output of the visual cells. We find that accordingly with the observed final behavior, the visual cells preferentially excite the motoneurons heading in the opposite angular direction (see Figures 1 and 2).

Acknowledgements

The author thanks the INRIA Lille-Nord europe for 1-year delegation in the SEQUEL team.

This work is supported by the french ANR MAPS (ANR-07-BLAN-0335-02).

References

1. Bartlett P, Baxter J: **Synaptic modifications in spiking neurons that learn**. *Technical report, Australian National University*; 1999.
2. Florian R: **A reinforcement learning algorithm for spiking neural networks**. *Proc of Seventh International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC'05) 2005:299-306.*

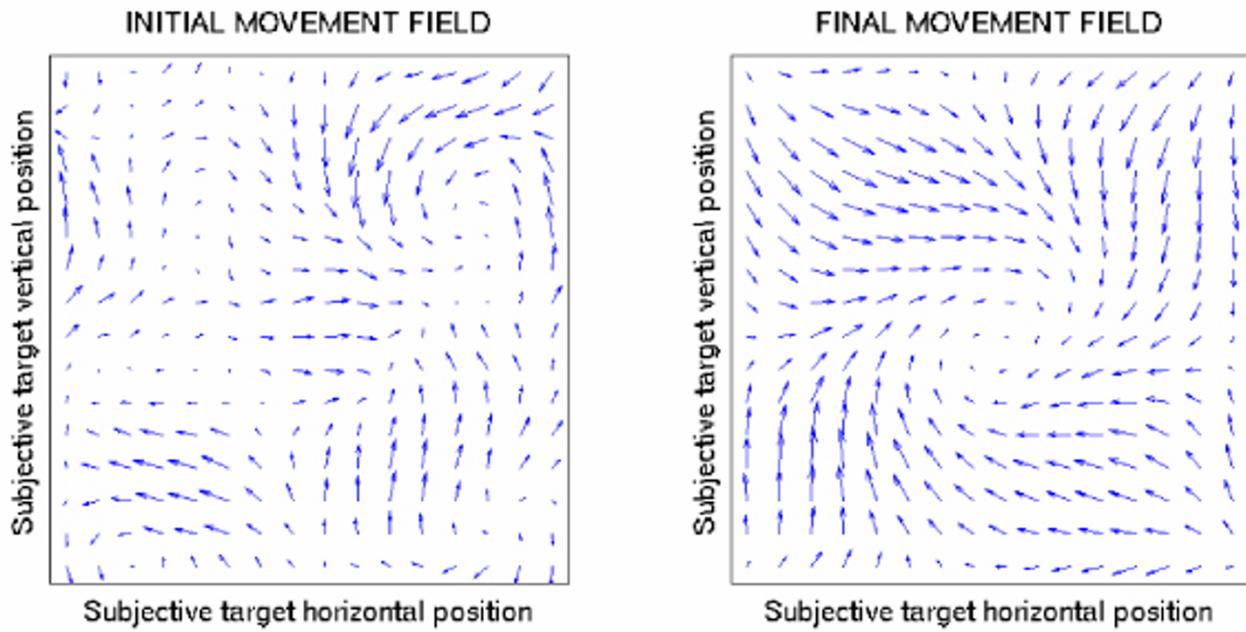


Figure 1

(Left) Network scheme. The visual layer is composed of 256 neurons sending excitatory axons toward a motor layer composed of 32 neurons. The motor neurons inhibit each other. A slow noise $b(t)$ is added to every neuron of the motor layer. (Right) Initial average motor output for a target appearing at the considered subjective position.

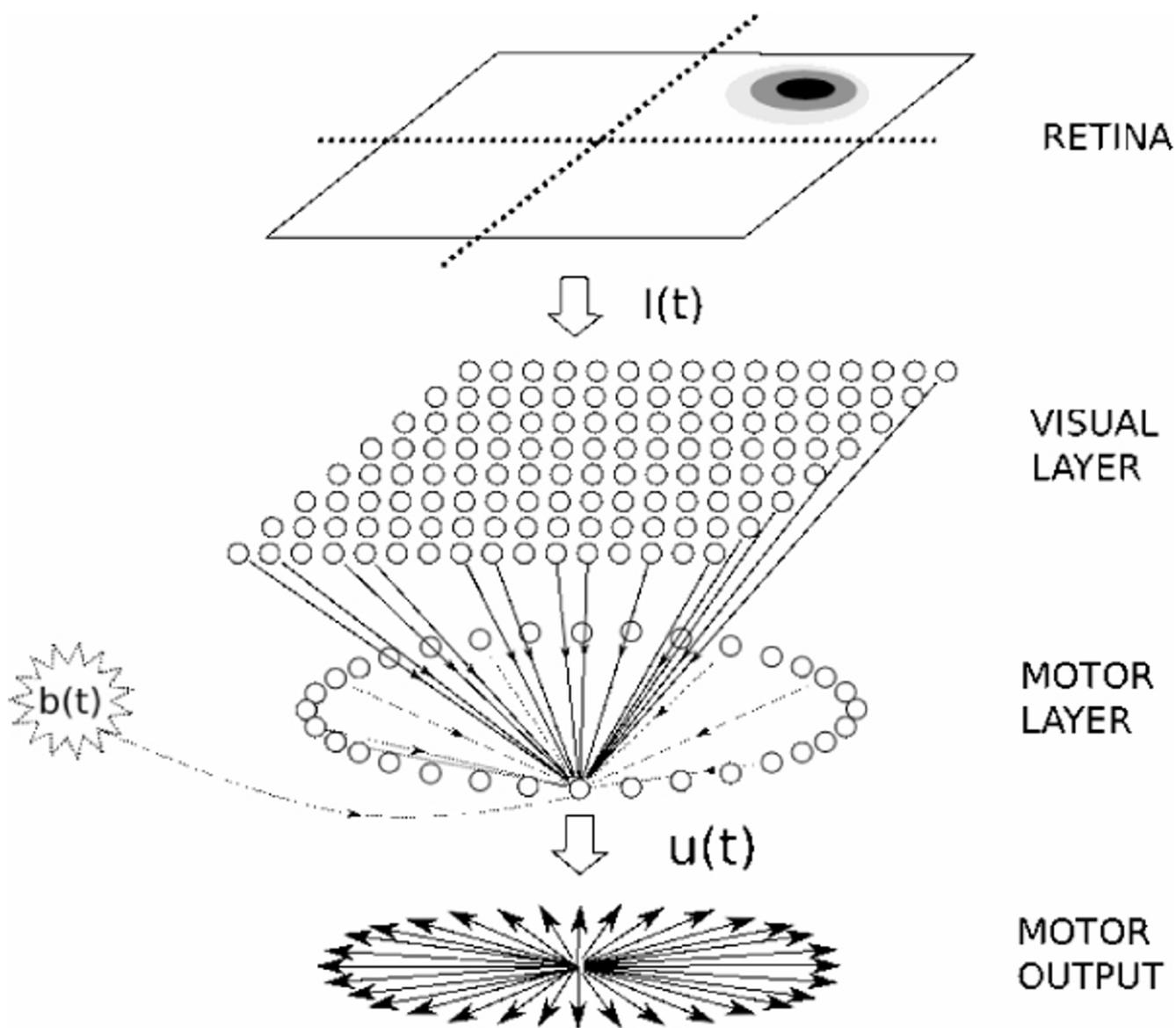


Figure 2
Final average motor output.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."
Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

