Temporal Coherence – A Generic Coding Principle for Multiple Cortical Areas

Wolfgang Einhäuser[1,2], Jörg Hipp[2], Reto Wyss[1,2], Paul F.M.J. Verschure[2], Peter König[2,3]

[1]California Institute of Technology, Division of Biology, Pasadena, CA, USA

[2]Swiss Federal Institute of Technology, Institute of Neuroinformatics, Zurich, Switzerland

[3]University of Osnabrück, Institute of Cognitive Science, Osnabrück, Germany

The cerebral cortex reveals a remarkably homogeneous architecture, which is in sharp contrast to the functional specialization of different cortical areas. This raises the question as to what extent the specific properties of cortical areas can be derived from the statistics of their typical input by using a small set of generic computational principles. Here we test the hypothesis that temporal coherence constitutes such a principle.

We train neurons to optimize the stability of their responses and to minimize redundancy in their representations within multi-layer artificial neural networks. In order to compare the resulting response properties to neurons in different areas of mammalian cerebral cortex, we use 3 different sets of real-world stimuli.

a) We train neurons to have optimally stable responses on images drawn from a standard image library. The trained representations are better suited to readily classify objects than the corresponding input representation. In particular, classification performance is largely invariant to the position and viewing angle of an object, even for previously unseen views. These properties correspond well to response properties of neurons in inferotemporal cortex.

b) We use a physical implementation of an artificial vibrissae (whisker) system and collect whisker deflection signals for various textures. On the output of a simple model of early sensory neurons, we train neurons to optimize the stability of their responses. We find that the trained neuronal representation is better suited to classify the textures than that of the early sensory neurons and compatible with the rodent's performance. Hence the representation is likely to reflect neuronal properties of higher somatosensory cortices.

c) We train a multi-layer network on natural stimuli as provided by a robot randomly exploring a real world office environment. Early processing stages show receptive fields similar to those observed in the primary visual cortex. Subsequent stages are selective for increasingly complex configurations of local features, as observed in higher visual areas. The last stage of the model displays place fields as observed in entorhinal cortex and hippocampus.

These results demonstrate that functionally heterogeneous cortical areas can be generated by only a few coding principles and highlight the importance of the variability of the input signals in forming functional specialization.

Supported by the Swiss National Science Foundation (WE, RW, PK) and the EU-BBW "AMOUSE" project (JH, PK).

Correspondence to WE (wet@caltech.edu).