

Informatics 1 Cognitive Science – Tutorial 8

Matthias Hennig

Week 9

In this tutorial we will look at cortical representations of sensory information and their possible interpretation in a semantic context. In the lectures we have seen that the cortex has a hierarchical organisation, with simple feature-like receptive fields in early areas, and more complex ones at later stages. In a higher area such as the visual IT cortex, neurons are no longer selective to simple features, but show signs of object selectivity instead.

While this suggests the visual system may indeed perform some form of object recognition, how precisely this happens is an open question. We also have seen that nearby cortical receptive fields tend to be sensitive to similar features (e.g. pinwheels in orientation maps), suggesting similarity in the stimulus can be preserved in the cortical topology (note that in an artificial neural network we would not expect this).

The reading for this tutorial is this following paper:

Huth, A. G., Nishimoto, S., Vu, A. T., & Gallant, J. L. (2012). *A continuous semantic space describes the representation of thousands of object and action categories across the human brain.* *Neuron*, 76(6), 1210-1224. <https://www.sciencedirect.com/science/article/pii/S0896627312009348>

As so often, the paper is a long read and quite technical. Therefore, start by reading the following and watch the video the authors made:

<https://xcorr.net/2013/01/24/categories/>

<https://www.youtube.com/watch?v=0FDtsbLZBuM>

The paper uses Principal Component Analysis (PCA), a method for dimensionality reduction you may not be familiar with yet. In simple terms, PCA looks for the directions of largest variance in multi-dimensional data. In many cases, this allows removing irrelevant dimensions, to describe and visualise complex data in a low-dimensional space. [More precisely, PCA is the eigen-decomposition of the data covariance matrix. High eigenvectors with high eigenvalues correspond to directions in data space with high variance, these capture essential data features.]

Questions:

1. State, in simple terms, the two main hypotheses this study attempts to address.
2. Explain the experimental design in this study.
3. What are the two main results of this study?
4. Can you think of any limitations of this study that could affect this interpretation?
5. What is your take on these results? What would be interesting questions to ask next?