Concurrent increases in selectivity and tolerance produce constant sparseness across the ventral visual stream

Nicole C Rust and James J DiCarlo

McGovern Institute for Brain Research, MIT

Neural coding schemes that minimize the number of neurons activated at any one time (or equivalently maximize "sparseness") are thought to be both metabolically and computationally efficient [reviewed by 1]. But does sparseness increase as signals propagate through the cortex? To investigate this question, we compared the response properties of neurons at different stages along the pathway supporting object recognition, the ventral visual stream. Specifically, we recorded the responses of neurons in a mid-level visual area (V4) and a high-level visual area (anterior inferotemporal cortex, IT) to a large set of natural images while monkeys performed an object detection task. We found that the distributions of sparseness values in V4 and IT were indistinguishable and that most neurons in both areas were broadly tuned. Similarly, individual images activated the same, large fraction of neurons in V4 and IT. Thus it appears that a coding principle is conserved at each level of processing; however, in opposition to theories of sparse coding, the tuning we observed in mid- and high-level vision is more consistent with a broadly distributed coding scheme [see also 2].

Example natural and scrambled images [3]

If sparseness is not changing, what is happening as signals propagate through the visual system? We began investigating this question by measuring tolerance to position and scale transformations. We found that individual neuron tolerance increased from V4 to IT, and this translated to enhanced performance of the IT over the V4 population on a position- and scale-invariant object recognition task. To determine whether the neurons in each area also differ in terms of the image features that elicit a response, we presented natural and "scrambled" images that have the same local structure but configured randomly [3]. We found that V4 neurons responded similarly to both image sets whereas IT neurons responded much more robustly to the natural images. Likewise, the V4 population discriminated between members of the two image sets with similar fidelity whereas discrimination by the IT population was considerably degraded for the scrambled as compared to the natural images. These results suggest that IT neurons are more selective than V4 neurons in terms of the image features that drive these cells. Moreover, we found that equivalent sparseness values were correlated with higher levels of selectivity and tolerance in IT as compared to V4.

Thus, as signals propagate through the visual system, neurons increase their selectivity for particular image features and, at the same time, neurons increase their tolerance for the position and scale of those features; the rates at which these two factors increase are set such that constant sparseness is maintained at each level of visual processing. Consistent with the observation that the structure of cortex is roughly identical regardless of where it sits in the hierarchy, we speculate that conservation of a broadly distributed coding scheme is an optimal use of resources in equipotential cortex.

References

[1] Sparse coding of sensory inputs. BA Olshausen and DJ Field, *Curr Opin Neurobiol.*, 14:481-7, 2004.

[2] Responses of neurons in primary and inferior temporal visual cortices to natural scenes. R Baddeley et al., *Proc Biol Sci.* 264:1775-83, 1997.

[3] A parametric texture model based on joint statistics of complex wavelet coefficients. J Portilla and EP Simoncelli, *Int J Comp Vis*, 40:49-71, 2000.