The following text on page 242 of *Vision and the Coding of Natural Images* by B.A. Olshausen and D.J. Field, **American Scientist**, vol. 88 (2000) is partially obscured by figure 7 in the pdf version of the reprint:

Redundancy Reduction

The answer comes from a theory that Horace Barlow of the University of Cambridge formulated nearly 40 years ago. He proposed a simple self-organizing principle for sensory neurons—namely that they should arrange the strengths of their connections so as to encode incoming sensory information in a manner that maximizes the statistical independence of their outputs (hence minimizing redundancy). Barlow reasoned that the underlying causes of visual signals are usually independent entities–separate objects moving about in the world–and if certain neurons somewhere in the brain are to represent these objects properly, their responses should also be independent. Thus, by minimizing the redundancy inherent in the sensory input stream, the nervous system might be able to form a representation of the underlying causes of images, something that would no doubt be useful to the organism.

Many years passed before Barlow's theory was put to work in a quantitative fashion to account for the properties of retinal ganglion cells, first by Laughlin and his colleagues in Canberra (in the early 1980s) and then a decade later by Joseph Atick, who was working at the Institute for Advanced Study in Princeton. Atick considered the form of correlations that arise in natural images–namely, the 1/f amplitude spectrum. He showed that the optimal operation for removing these correlations is to attenuate the low spatial frequencies and to boost the highones in inverse proportion to their original amplitudes. The reason is quite simple: A decorrelated image has a spatial-frequency power spectrum that is flat–the spatial-frequency equivalent of white noise–which is just what Atick's transformation yields.

Atick's theory thus explains why retinal neurons have the particular receptive fields they do: The concentric zones of excitation and inhibition essentially act as a "whitening filter," which serves to decorrelate the outputs sent down the optic nerve. The specific form of the receptive fields that Atick's theory predicts nicely matches the properties of retinal ganglion cells in terms of spatial frequency. And recently Yang Dan, now at the University of California, Berkeley, showed that Atick's theory also accounts for the temporal-frequency response of neurons in the lateral geniculate nucleus.

Sparse Coding

The agreement between the theory of redundancy reduction and the workings of nerve cells in the lower levels of the visual system is encouraging. But such mechanisms for decorrelation are just the tip of the iceberg. After all, there is more to natural images than the obvious similarity among pairs of nearby pixels.

One way to get a feel for the statistical structure present is to consider what images would look like if they could be completely characterized by two-point correlations among pixels (Figure 6). One of the most obvious ways that natural scenes differ from such images is that they contain sharp, oriented discontinuities. Indeed, it is not hard to see that most images contain regions of relatively uniform structure interspersed with distinct edges, which give rise to unique three-point and higher correlations. So one must also consider how neurons might reduce the redundancy that comes about from these higher-order forms of structure. A natural place to look is the primary visual cortex, which has been the focus of many studies since the early 1960s, when David Hubel and Torsten Wiesel at Harvard University first charted the receptive fields of these neurons and discovered their spatially localized, oriented and "bandpass" properties. That is, each neuron in this area responds selectively to a discontinuity in luminance at a particular location, with a specific orientation and containing a limited range of spatial frequencies. By the middle of the 1970s, some investigators began modeling these neurons quantitatively and were attempting to represent images with these models.

Stjepan Marcelja, a mathematician at the Australian National University, noticed some of these efforts by neuroscientists and directed their attention to theories of information processing that Dennis Gabor developed during the 1940s. Gabor, a Hungarian-English scientist who is most famous for inventing holography, showed that the function that is optimal for matching features in time-varying signals simultaneously in both time and frequency is a sinusoid with a Gaussian (bell shaped) envelope. Marcelja pointed out that such functions, now commonly known as Gabor functions, describe extremely well the receptive fields of neurons in the visual cortex (Figure 7). From this work, many neuroscientists concluded that the cortex must be attempting to represent the structure of images in both space and spatial frequency. But the Gabor theory still begs the question of why such a joint space-frequency representation is important. Is it somehow particularly well suited to the higher-order statistical structure of natural images?