

What the statistics of natural images tell us about visual coding.

David J. Field

Physiological Laboratory, University of Cambridge
Cambridge England, CB2 3EG

ABSTRACT

Why does the mammalian visual system represent information as it does? If we assume that visual systems have evolved to cope with the natural environment then we might expect the coding properties of the visual system to be related to the statistical structure of our environment. Indeed, images of the natural environment do not have random statistics. The first-order statistics (e.g., distribution of pixel values) and second-order statistics (e.g., power spectra) of natural images have been discussed previously and they bear important relations to visual coding. Statistics higher than second-order are difficult to measure but provide crucial information about the image. For example, it can be shown that the lines and edges found in natural images are a function of these higher-order statistics. In this paper, these higher-order statistics will be discussed in relation to the coding properties of the mammalian visual system. It is suggested that the spatial parameters of the cortical 'filters' (e.g., bandwidths of simple and complex cells) are closely related to these higher-order statistics. In particular, it will be shown that the spatial non-linearities shown by cortical complex cells provide the early visual system with the information required to learn about these statistics.

1. INTRODUCTION

The precise function of cells in the mammalian visual cortex have puzzled researchers since they were first mapped out by Hubel and Wiesel¹ in 1962. Theories as to why these cells behave as they do have varied from edge and bar detection to Fourier analysis. More recently, it has been suggested that insight into the behaviour of cells in the mammalian visual pathway may be gained by a better understanding of the statistics of the visual environment^{2,3}. This line of approach will be continued in this paper. We will consider a particular property of natural images, and how this property relates to the behaviour of cortical cells. In particular, we wish to emphasize a type of cell described as 'complex'^{1,4}. Although such cells are suggested to represent as much as half of primary visual cortex⁵, no widely accepted theory of their function is available. Complex cells show a particular type of spatial nonlinearity which distinguishes them from the more linear and more widely studied 'simple' cells.

There are two main points to this paper. First, natural images show a particular form of redundancy which does not show up in measures of second-order statistics. This redundancy results in local features like edges and lines. In the frequency domain, an edge or bar can be described as a local correlation in phase: at the location of the edge, the Fourier components in the same phase. However, when the image contains a large number of edges and lines at different positions, the correlation between global Fourier coefficients is lost. There may still be redundancy, but it is now in the form of a higher-order statistic. It might loosely be described as a redundancy across different scales of the image; the information present in different frequency bands is not independent.

In the second part of this paper, a visual model will be described which represents images with 'filters' that share many properties of cortical receptive fields. The most important property that will be investigated will be the spatial non-linearity shown by cortical complex cells. By manipulating the bandwidths in this model, we will attempt to get a better idea of why cortical cells have their particular parameters and how these parameters relate to the statistics described above.

1.1 The statistics of natural images

Randomly selected images from the natural world do not have random statistics. Such images do not fit a description of randomly generated pixels. Many of the regularities of natural images are best described in terms of "nth-order" statistics. For example, first-order statistics refer to the probability distributions of the values of each of the pixels (i.e., the luminances at different points). If we assume that all pixels in an image represent samples from the same distribution, then it is possible to pool the pixels from a single image to get an estimate of first order statistics.