(depending on the exact experiment) centering on the classical receptive field and covering some of the near surround. The spike times were recorded with 1 ms resolution. The stimulus sets included simple bright and dark bars on a gray background, patches of oriented sine wave gratings from 1 to 6 cycles per degree, and Walsh patterns. In the LGN, the stimuli included bright and dark bars, and center-surround stimuli of both positive and negative contrasts, with centers ranging from smaller than to larger than the neuron's receptive field center. The gratings, Walsh patterns, and center-surround stimuli were contrast-modulated around the background luminance.

Data analysis included statistical descriptions and information theoretic methods. Information theory quantifies how well an output signal, here a neuron's response, can be used to identify the input, here a visual stimulus (Shannon & Weaver, 1949; Cover & Thomas, 1991). Before knowing a neuron's response, we are uncertain about which member of a stimulus set was presented. Information is the reduction in uncertainty about which stimulus was presented, provided by knowing the neuron's response. Intuitively, information quantifies how distinct the responses to different stimuli are. The less the responses to different stimuli overlap, the smaller the chance that more than one stimulus elicited a particular response, and the higher the information. The more the responses to different stimuli overlap, the more difficult it is to determine which stimulus elicited a particular response, and the lower the information. Thus, our minimal description of neuronal responses must include both the average responses to the stimuli and their variabilities, namely, we must know the distribution of responses to each stimulus. Estimating information from data can be difficult and should be undertaken carefully. The problems that are associated with estimating information from small data sets are beyond the scope of this discussion, but are discussed extensively elsewhere (Kjaer et al., 1994; Golomb et al., 1997; Panzeri, Treves, 1996).

For the work below, our use of information theory is largely comparative. We would like to know which representations of neural responses allow the classification of stimuli with the greatest certainty, namely, which representations carry more information. Often we want to know whether adding an element to the representation of the response (that is, increasing the dimension of the code) adds information.

RESULTS

Response Strength

Since the earliest single-neuron recordings, it has been clear that the number of impulses or response strength is easily modulated by changing experimental parameters. For example, in V1 of monkey cortex, the responses to a stationary sine wave grating or bar centered on the receptive field change as the orientation is changed, giving rise to the classic tuning curve.

Despite this obvious relation between orientation and response strength, the interpretation of the responses is more complicated than first appears because the number of impulses varies widely across repeated trials. This variability is usually dealt with by averaging across (hopefully many) repetitions of the stimulus condition. Because the brain does not have the luxury of the experimentalist (observing many responses) and must decode the response from a single stimulus presentation, the brain cannot average across trials. Instead, the response must be decoded across neurons. Asking how many independent neurons are needed to decode the response to determine, for example, the orientation of a bar or grating, is natural. This question can be answered by examining how much information is needed to determine the orientation, and how much information is provided by a single neuron.

Past methods for calculating information have estimated response distributions directly from the data, although, as was pointed out above, doing so can be difficult. In the past few years, reliable methods have been developed (Kjaer et al., 1994; Golomb et al., 1997; Panzeri, Treves, 1996; Victor, Purpura, 1996). Recently we have taken another approach. Rather than estimating such