Modeling Surround Suppression in V1 Neurons with a Statistically-Derived Normalization Model

Eero P. Simoncelli
Center for Neural Science, and
Courant Institute of Mathematical Sciences
New York University
eero.simoncelli@nyu.edu

Odelia Schwartz
Center for Neural Science
New York University
odelia@cns.nyu.edu

Abstract

We examine the statistics of natural monochromatic images decomposed using a multi-scale wavelet basis. Although the coefficients of this representation are nearly decorrelated, they exhibit important higher-order statistical dependencies that cannot be eliminated with purely linear processing. In particular, rectified coefficients corresponding to basis functions at neighboring spatial positions, orientations and scales are highly correlated. A method of removing these dependencies is to *divide* each coefficient by a weighted combination of its rectified neighbors. Several successful models of the steady-state behavior of neurons in primary visual cortex are based on such "divisive normalization" computations, and thus our analysis provides a theoretical justification for these models. Perhaps more importantly, the statistical measurements explicitly specify the weights that should be used in computing the normalization signal. We demonstrate that this weighting is qualitatively consistent with recent physiological experiments that characterize the suppressive effect of stimuli presented outside of the classical receptive field. Our observations thus provide evidence for the hypothesis that early visual neural processing is well matched to these statistical properties of images.

An appealing hypothesis for neural processing states that sensory systems develop in response to the statistical properties of the signals to which they are exposed [e.g., 1, 2]. This has led many researchers to look for a means of deriving a model of cortical processing purely from a statistical characterization of sensory signals. In particular, many such attempts are based on the notion that neural responses should be statistically independent.

The pixels of digitized natural images are highly redundant, but one can always find a linear decomposition (i.e., principal component analysis) that eliminates second-order cor-

---

Research supported by an Alfred P. Sloan Fellowship to EPS, and by the Sloan Center for Theoretical Neurobiology at NYU.

relation. A number of researchers have used such concepts to derive linear receptive fields similar to those determined from physiological measurements [e.g., 16, 20]. The principal components decomposition is, however, not unique. Because of this, these early attempts required additional constraints, such as spatial locality and/or symmetry, in order to achieve functions approximating cortical receptive fields.

More recently, a number of authors have shown that one may use higher-order statistical measurements to uniquely constrain the choice of linear decomposition [e.g., 7, 9]. This is commonly known as *independent components analysis*. Vision researchers have demonstrated that the resulting basis functions are similar to cortical receptive fields, in that they are localized in spatial position, orientation and scale [e.g., 17, 3]. The associated coefficients of such decompositions are (second-order) decorrelated, highly kurtotic, and generally more independent than principal components.

But the response properties of neurons in primary visual cortex are not adequately described by linear processes. Even if one chooses to describe only the mean firing rate of such neurons, one must at a minimum include a rectifying, saturating nonlinearity. A number of authors have shown that a gain control mechanism, known as *divisive normalization*, can explain a wide variety of the nonlinear behaviors of these neurons [18, 4, 11, 12, 6]. In most instantiations of normalization, the response of each linear basis function is rectified (and typically squared) and then divided by a uniformly weighted sum of the rectified responses of all other neurons. Physiologically, this is hypothesized to occur via feedback shunting inhibitory mechanisms [e.g., 13, 5]. Ruderman and Bialek [19] have discussed divisive normalization as a means of increasing entropy.

In this paper, we examine the joint statistics of coefficients of an orthonormal wavelet image decomposition that approximates the independent components of natural images. We show that the coefficients are second-order decorrelated, but *not* independent. In particular, pairs of rectified responses are highly correlated. These pairwise dependencies may be eliminated by dividing each coefficient by a *weighted* combination of the rectified responses of other neurons, with the weighting determined from image statistics. We show that the resulting model, with all parameters determined from the statistics of a set of images, can account for recent physiological observations regarding suppression of cortical responses by stimuli presented outside the classical receptive field. These concepts have been previously presented in [21, 25].

1 Joint Statistics of Orthonormal Wavelet Coefficients

Multi-scale linear transforms such as wavelets have become popular for image representation. Typically, the basis functions of these representations are localized in spatial position, orientation, and spatial frequency (scale). The coefficients resulting from projection of natural images onto these functions are essentially uncorrelated. In addition, a number of authors have noted that wavelet coefficients have significantly non-Gaussian marginal statistics [e.g., 10, 14]. Because of these properties, we believe that wavelet bases provide a close approximation to the independent components decomposition for natural images. For the purposes of this paper, we utilize a typical separable decomposition, based on symmetric quadrature mirror filters taken from [23]. The decomposition is constructed by splitting an image into four subbands (lowpass, vertical, horizontal, diagonal), and then recursively splitting the lowpass subband.

Despite the decorrelation properties of the wavelet decomposition, it is quite evident that wavelet coefficients are *not* statistically independent [26, 22]. Large-magnitude coefficients (either positive or negative) tend to lie along ridges with orientation matching that of the subband. Large-magnitude coefficients also tend to occur at the same relative spatial locations in subbands at adjacent scales, and orientations. To make these statistical relationships