Neuron, Vol. 46, 493–504, May 5, 2005, Copyright ©2005 by Elsevier Inc. DOI 10.1016/j.neuron.2005.03.026

Redundancy in the Population Code of the Retina

Jason L. Puchalla,[1] Elad Schneidman,[1,2] Robert A. Harris,[1] and Michael J. Berry[1,*]
[1]Department of Molecular Biology
[2]Lewis-Sigler Institute for Integrative Genomics
Princeton University
Princeton, New Jersey 08544

Summary

We have explored the manner in which the population of retinal ganglion cells collectively represent the visual world. Ganglion cells in the salamander were recorded simultaneously with a multielectrode array during stimulation with both artificial and natural visual stimuli, and the mutual information that single cells and pairs of cells conveyed about the stimulus was estimated. We found significant redundancy between cells spaced as far as 500 μm apart. When we used standard methods for defining functional types, only ON-type and OFF-type cells emerged as truly independent information channels. Although the average redundancy between nearby cell pairs was moderate, each ganglion cell shared information with many neighbors, so that visual information was represented 10-fold within the ganglion cell population. This high degree of retinal redundancy suggests that design principles beyond coding efficiency may be important at the population level.

Introduction

Retinal ganglion cells in the same vicinity have long been known to have receptive fields that overlap extensively. This means that the retinal code at the level of the ganglion cells intrinsically uses populations of neurons to represent even the sharpest spatial features within a visual image. In addition, natural visual scenes have strong spatial correlations (Field, 1987; van der Schaaf and van Hateren, 1996), which may cause an even wider group of ganglion cells to participate in encoding any single feature within a natural image. A central issue in studying the retina, as well as many other neural circuits, is to understand how information is represented collectively by the activity of a population of neurons.

The retina is an exquisitely organized circuit with many types of ganglion cells whose dendrites sample different sources of visual information within the inner plexiform layer (Dacey et al., 2003; Pang et al., 2002; Pang et al., 2004; Rockhill et al., 2002; Roska and Werblin, 2001; Sterling, 1998; Wu et al., 2000). Furthermore, the dendrites of many types of ganglion cells tile, or just barely cover visual space (Dacey, 1993; Vaney, 1994; Wässle and Boycott, 1991), leading to receptive-field tiling for many kinds of ganglion cells (DeVries and Baylor, 1997; Frechette et al., 2004). Together, these observations have led to the idea that the population of ganglion cells comprises many parallel channels of visual information and that such a code may be highly efficient: cells within each channel have minimally overlapping receptive fields, and cells in different channels encode different visual features. This organization complements classical theoretical ideas of efficient information encoding by sensory systems (Attneave, 1954; Barlow, 1961), where it has been suggested that the center-surround antagonism of ganglion cell-receptive fields acts to decorrelate the spatial power spectrum of natural scenes and reduce the redundancy in the retinal code (Atick, 1992; Atick and Redlich, 1992; van Hateren, 1992). However, the efficiency of the retinal code has not been directly tested.

There are many reasons to think that nearby ganglion cells may not encode independent visual information. First, the receptive/dendritic fields of some types of ganglion cells cover visual space many times (Peichl and Wassle, 1979; Rockhill et al., 2002). Second, functional and anatomical classification is based on determining whether groups of ganglion cells can be clearly distinguished; such methods do not guarantee that these different cell types will encode entirely independent information. Third, natural stimuli cannot be fully characterized by their spatial power spectrum: they also possess strong spatiotemporal regularities due to object motion (Dong and Atick, 1995) as well as higher-order correlations due to extended spatial patterns like edges (Ruderman, 1994; Ruderman, 1997). It is not clear whether retinal processing can detect and eliminate these complex correlations. Furthermore, much of the analysis of population coding in the retina is based on the measurement of receptive fields. Such a characterization is likely to be incomplete, because ganglion cells respond to light with many nonlinearities that are not captured by the spatiotemporal receptive field (Hartline, 1937; Olveczky et al., 2003; Shapley and Victor, 1979; Victor and Shapley, 1979) and because the receptive field can change considerably due to adaptation (Barlow et al., 1957; Smirnakis et al., 1997).

Instead of relying on receptive field maps or simplified models of ganglion cell function, the redundancy of the retinal code can be directly measured by quantifying the amount of information that ganglion cells convey about the stimulus in groups and as individuals (Warland et al., 1997). We present here the first direct measurement of the redundancy in the neural code of the retina in response to a rich set of natural stimuli. To aid in understanding the factors that give rise to redundancy between ganglion cells, we also used simplified artificial stimuli to analyze the sources of retinal redundancy and to relate redundancy to the receptive-field properties of ganglion cells. A simple model is described to explore the consequences of redundancy between cell pairs for the information encoded by the entire ganglion cell population.

Results

We recorded simultaneously from populations of ganglion cells in the salamander retina using a planar multi-

*Correspondence: berry@princeton.edu

electrode array. To explore the manner in which visual scenes are represented within the population of retinal ganglion cells, we calculated the fractional redundancy between all pairs of ganglion cells. This quantity measures the degree to which pairs of ganglion cell spike trains encode the same visual information (Gat and Tishby, 1999; Gawne and Richmond, 1993; Panzeri and Schultz, 2001; Petersen et al., 2001; Schneidman et al., 2003a); its value is sensitive not just to pairwise correlations between spikes, but to all correlations that can be directly sampled (see Experimental Procedures). By directly sampling, the redundancy can be measured without implicitly assuming any model of the light response or of the noise. Redundancy was defined as the difference between the mutual information that the responses of each cell alone conveyed about the stimulus, $I(R_a;S)$ and $I(R_b;S)$, and the information conveyed by their joint responses, $I(R_a,R_b;S)$. As information rates varied widely within the population of ganglion cells, we calculated the redundancy as a fraction of the minimum information of the two individual cells (Reich et al., 2001):

$$\overline{\Lambda}_{ab} = \frac{[I(R_a;S) + I(R_b;S)] - I(R_a,R_b;S)}{\min\{I(R_a;S),I(R_b;S)\}}. \quad (1)$$

This normalization factor, $\min\{I(R_a;S), I(R_b;S)\}$, is the maximum possible redundancy between two cells, so that the fractional redundancy can be no greater than 1. The fractional redundancy is 0 when the two cells encode independent information about the stimulus; its value is 1 when the two cells encode exactly the same information or when one cell's information is a subset of the other's. Negative values of the redundancy mean that the cells are synergistic.

Redundancy between Pairs of Ganglion Cells

In order to assess retinal processing under realistic visual conditions, we stimulated retinas with a set of natural movie clips chosen to represent a variety of environments. An especially important characteristic of natural stimuli is the wide field motion caused by the movement of an animal's eyes or body, as this should strongly stimulate many ganglion cells. We included movies having five different categories of motion: object motion, optic flow, smooth pursuit, saccades, and combinations of these kinds of motion. Movies categorized as having object motion were filmed while the video camera remained stationary and one or more objects within the field of view moved freely (see Experimental Procedures). For the other categories of movies, the camera was moved so as to stimulate eye or body movements. Most movies were taken of woodland scenes, but some were aquatic or man-made. Four examples of movie frames are shown in Figure 1A.

Figure 1B shows spike rasters from ten cells recorded simultaneously during the forest walk movie. As seen, the spike trains were sparse and temporally precise, primarily occurring in well-isolated firing events (Berry et al., 1997). There was a complex pattern of event times across the population, with some cells sharing many narrow events and others sharing none. Figure 1C plots the fractional redundancy for 1838 cell pairs stimulated by 12 different movies versus the distance between ganglion cells, as determined from the receptive field of each cell (see Experimental Procedures). Although the redundancy depended systematically on the distance between cells, there was wide variation among the values for cell pairs of roughly the

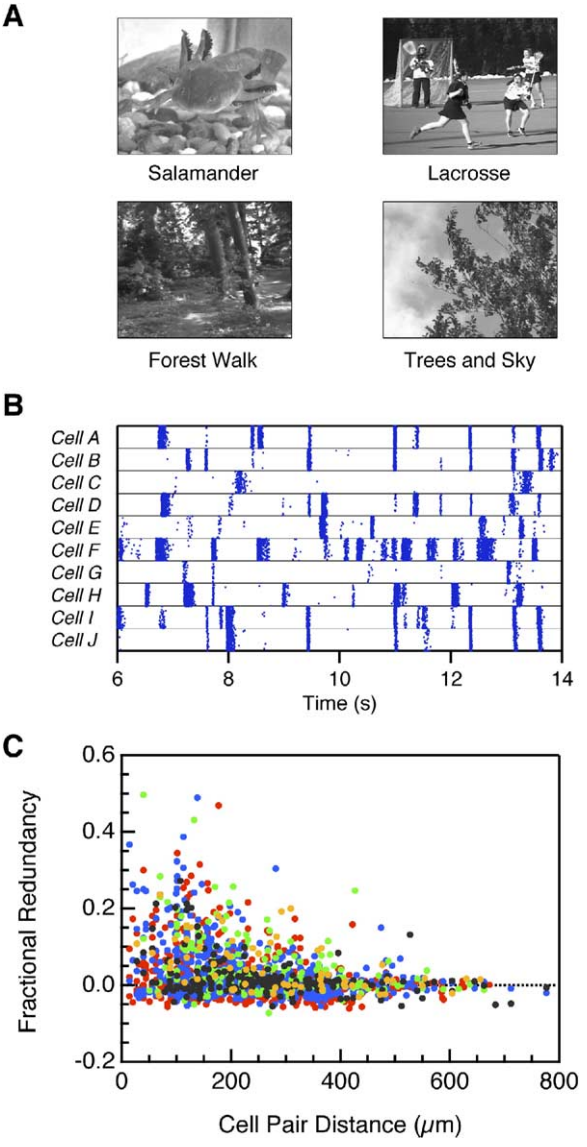

Figure 1. Redundancy under Natural Stimulation

(A) Single frames from four natural movie clips having different categories of motion.

(B) Examples of spike rasters from ten cells recorded simultaneously during the forest walk movie clip. Each dot represents the time of a spike; vertical dimension shows 120 repeated stimulus trials. Cells A and B have a fractional redundancy of 0.17, while cell C does not share significant redundancy with any of the other cells.

(C) Fractional redundancy for 1838 cell pairs in 4 retinas stimulated by natural scenes plotted versus the distance between the cell's receptive-field centers. The type of motion present in each movie clip is shown by the dot color: object motion (red), saccades (blue), optic flow (green), smooth pursuit (black), and combinations of motion (orange).

same distance apart. Redundancies as large as 50% were seen among nearby cells; cells I and J in Figure 1B are such an example. However, most cell pairs—even those that were closely spaced—were nearly independent, indicating that ganglion cells can extract many different visual features from the same spatial location within a natural image. Significant redundancy existed between some cell pairs separated by up to 500 μm. Since the receptive-field center radius of most ganglion cells is ~100 μm, redundant coding was therefore present for some cell pairs with nonoverlapping receptive-field centers. Very few cell pairs exhibited values of synergy that could be not explained by sampling errors (see Experimental Procedures), and the largest values of synergy were less than 5%. No obvious differences in redundancy were found among the five categories of movie motion (Figure 1, colors).

Overrepresentation within the Ganglion Cell Population

Since a large fraction of all the cell pairs that we measured showed at least some redundancy, the pairwise redundancy offers only a partial measure of the total redundancy of the visual messages sent from eye to brain. To quantify the number of neighbors with which a typical ganglion cell shares redundant information, we considered the fraction of cell pairs with redundancies greater than 5%. This threshold is a conservative bound on the random error inherent in our estimation of the redundancy (see Experimental Procedures). Shown in Figure 2A, this fraction was ~40% for nearby ganglion cells. Most redundant neighbors were found within 150 μm of a given ganglion cell. But because the number of neighbors increases with distance, a substantial contribution came from cells spaced up to 500 μm apart. When we estimated the total number of redundant partners within a neighborhood of radius 500 μm (assuming a density of 1400 cells/mm^2), we found 114 ± 11 such cells under natural stimulus conditions. Among those significantly redundant partners, the average redundancy was ~14% at close range and decreased to ~7% at 500 μm away (Figure 2B). Therefore, a typical ganglion cell shares a moderate degree of redundancy with a large number of cells in its local region of the retina.

In order to assess the contribution of pairwise correlations to the total redundancy of the ganglion cell population, we combined all of our redundancy measurements to estimate an overrepresentation factor. The overrepresentation factor is a sum of the redundancies between one ganglion cell and all of its neighbors, divided by the mutual information of the original cell (see Experimental Procedures). This factor measures the number of times, on average, that information conveyed by one neuron is also conveyed by neighboring neurons. Shown in Figure 2C, the overrepresentation factor for a typical ganglion cell was roughly 11.0 ± 1.0 when stimulated by natural movies. In this calculation, we included all pairwise redundancies above a threshold of 5% and treated cell pairs with less redundancy as independent. If we use no threshold and simply average over all cell pairs, the overrepresentation is roughly the same, 11.4 ± 1.2. If we set this threshold at

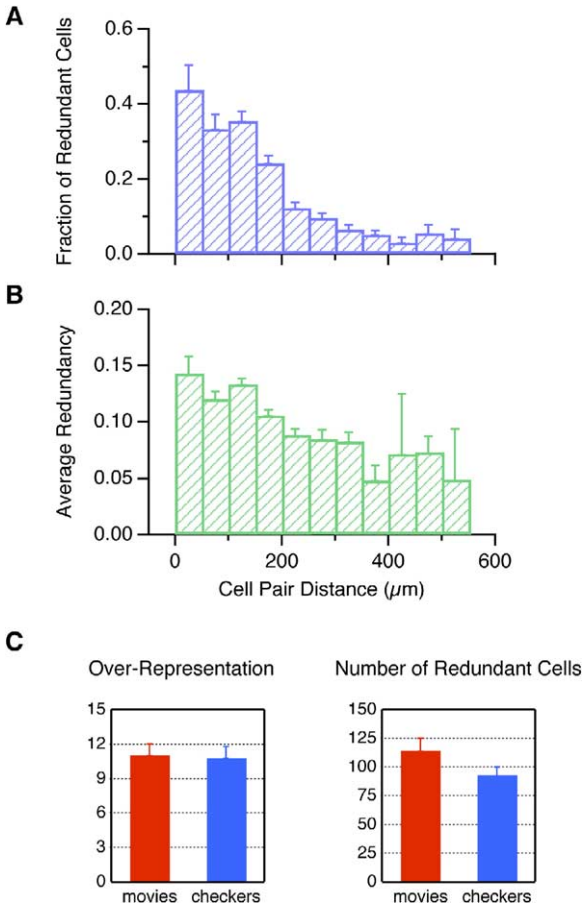

Figure 2. Overrepresentation within the Ganglion Cell Population

(A) The proportion of all measured cell pairs whose redundancy exceeded a significance threshold of 5% plotted as a function of the distance between cells.

(B) The average redundancy of those significantly redundant cell pairs plotted versus cell pair distance.

(C) Left: overrepresentation factor for natural movies (red) and checkerboard flicker (blue, see Experimental Procedures). Right: total number of neighboring cells that share significant redundancy with a single cell, estimated by summing over the data in (A) and assuming a cell density of 1400 cells/mm^2, for natural movies (red) and checkerboard flicker (blue). All error bars were determined by randomly selecting half of the cell pairs and repeating the calculation for many such selections.

10%, the overrepresentation factor only drops to 6.0 ± 0.6, indicating that a significant contribution comes from a small number of cell pairs with much larger redundancies. This analysis suggests that visual messages are, on average, represented many times within the population of ganglion cells.

Sources of Redundancy

Ganglion cells can encode similar visual messages either because of correlations in the stimulus or because of shared retinal circuitry that makes them sensitive to similar visual features. In order to study the contributions from these different mechanisms, we used random flickering checkerboards, where the stimulus had minimal correlations in space and time (55 μm

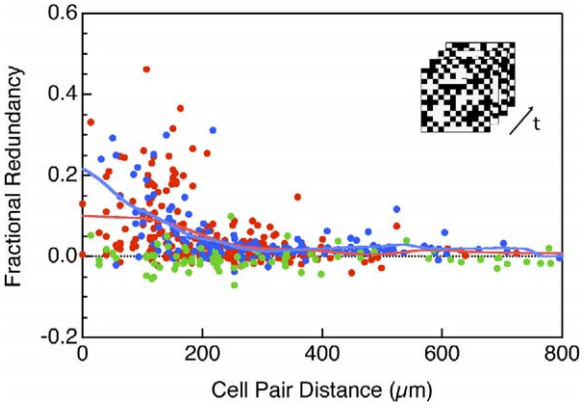

Figure 3. Redundancy under Minimally Correlated Stimulation

Fractional redundancy for 479 cell pairs stimulated by flickering checkerboards plotted versus the distance between the cell's receptive-field centers. Color indicates the functional type of the cells in a pair: pairs where both cells are the same functional type (blue), pairs where one cell is ON-type and the other is OFF-type (green), and other pairs where the cells are of different functional type (red). Lines show the average redundancy for same-type cell pairs (blue) and different-type cell pairs other than ON-OFF pairs (red). ON-OFF cell pairs were independent.

checker size, 33 ms frame time). In such a stimulus ensemble, correlations between ganglion cells are primarily due to shared retinal circuitry. The largest redundancies under checkerboard stimulation (~40%; Figure 3) were similar to those for natural movies, suggesting that receptive-field overlap plays the dominant role in redundant coding for nearby ganglion cells. In fact, the overrepresentation under checkerboard flicker, 10.8 ± 1.0, was roughly the same as found under naturalistic stimulation (Figure 2C), reinforcing this conclusion. For checkerboard flicker, significant redundancy was found primarily for cell pairs closer than ~200 μm (Figure 3), consistent with the typical receptive-field center diameter of salamander ganglion cells. By comparing this result to that for natural movies (Figure 1), one can see that the stimulus correlations present in natural movies give rise to significant redundancy for more widely separated ganglion cells, although such pairs are uncommon.

As seen for natural movies, ganglion cells under flickering checkerboard stimulation also exhibited a wide variety of redundancy values at a given cell pair distance (Figure 3). One obvious explanation for this diversity is that cells of the same functional type have higher redundancy than cells of different functional type. While a consensus has not yet emerged about how to perform functional classification, there are standard, widely used methods based on the responses of neurons to a broad ensemble of stimuli. Following recent studies (DeVries and Baylor, 1997; Schnitzer and Meister, 2003), we assigned ganglion cells to six functional types using the temporal dynamics of the receptive-field center (Figure 4). Greater redundancy was found, on average, for pairs where both cells were of the same functional type than for pairs where the cells belonged to different functional types (Figure 3, blue lines versus red lines), although considerable variation was again observed (blue dots versus red dots). This variability may reflect the fact that some cells did not fall clearly into a single functional class, but instead had properties that were intermediate between two classes. Cell pairs were found to encode independently when one member of the pair was any kind of ON cell and the other member was any kind of OFF cell (green dots). When one cell in the pair was slow OFF-type, independence was also found, regardless of the functional type of its partner (data not shown). Under naturalistic stimulation, however, cell pairs with one slow OFF-type cell were no longer found to be independent, indicating that only ON and OFF channels of visual information are truly independent.

Interpreting Redundancy Values

Intuitively, high redundancy is expected for cell pairs with overlapping spatial-receptive fields, as such cells receive almost identical stimulation during natural movie clips. But even if ganglion cells have identical spatial-receptive fields, subtle differences in their feature selectivity should limit their fractional redundancy to values less than 1. Noise also serves to limit redundancy, as it causes nominally identical neurons to respond differently on a trial-by-trial basis. In order to better understand how these factors affect coding redundancy, we stimulated the retina with spatially uniform flicker, where every cell received identical stimulation. Although ganglion cell spike trains were locked to the stimulus with great precision under these conditions (Berry et al., 1997), we found an average significant redundancy of 31% and values no greater than 80% (Figure 5). This indicates that noise and functional diversity indeed limit the redundancy to values much less than 1. Comparing Figures 1 and 5, we see that under naturalistic visual conditions, the redundancy between ganglion cells can be a large fraction of what is possible given their noise and feature selectivity. However, a much larger proportion of cell pairs were independent under naturalistic than uniform stimulation (57% versus 21% of cell pairs closer than 50 μm), indicating that ganglion cells with overlapping receptive fields can exploit differences in their spatiotemporal sensitivity to extract distinct information from natural scenes.

The fractional redundancy includes correlations beyond pairs of spikes as well as non-Poisson variability in ganglion cell spike trains. How does this quantity compare to simpler, non-information theoretic measures of shared function? One such measure is the similarity between the spatiotemporal receptive fields of two cells. We investigated the relationship between the full receptive-field structure of ganglion cells and their redundancy by computing the receptive-field overlap between pairs of ganglion cells. This quantity is normalized to have a value of 1 when the two cells have identical spatiotemporal receptive fields (see Experimental Procedures).

As expected, the receptive-field overlap was a strong function of the distance between cells (Figure 6A). Since nonzero values were often found for cells spaced by more than 200 μm, receptive-field surrounds could

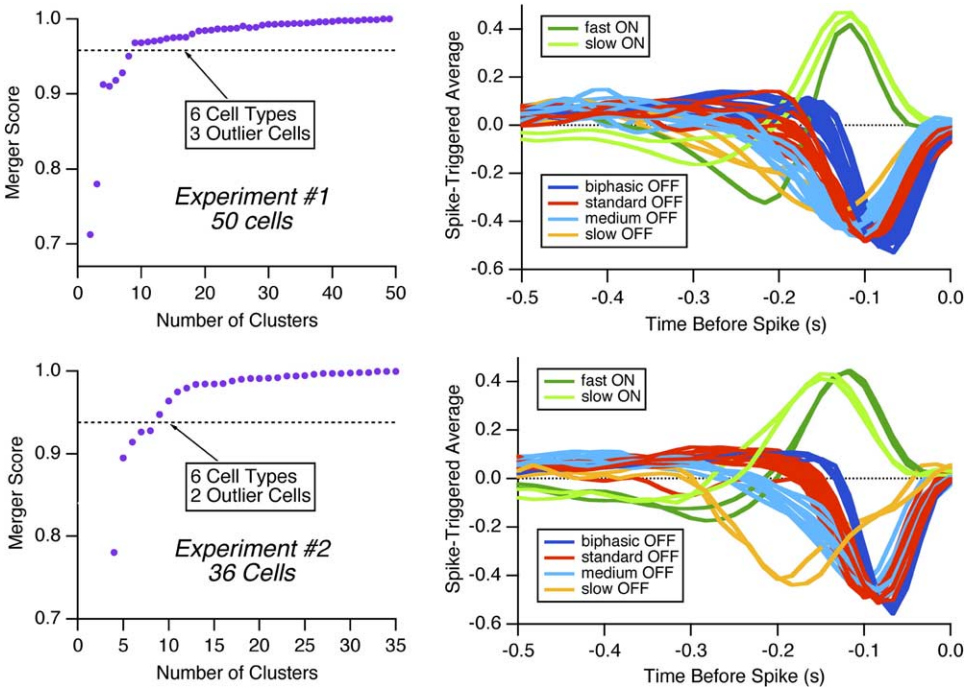

Figure 4. Functional Classification

Left columns: Merger score of our quantitative clustering algorithm (see Experimental Procedures) plotted as a function of the number of clusters. The break in the merger score is indicated by a dashed line; this threshold defines the number of significant clusters. In each experiment, the final merger was between clusters of ON and OFF cells; this merger had a negative score that is off the scale of the graphs. Right columns: Time course of the spike-triggered average for all ganglion cells in the experiment. Color indicates functional type. Outlier cells were combined with their most similar cluster. Rows correspond to recordings from two different retinas.

make a significant contribution to the overlap. Like the redundancy (Figure 3A), the receptive-field overlap took on a wide range of values for nearby cells and was often larger for cells of different functional type than for cells of the same functional type (red dots versus blue dots). Again, this may reflect the fact that not all ganglion cells fell clearly into a single functional class. When directly compared, large values of receptive-field overlap often corresponded to large values of redundancy (Figure 6B). However, extensive variability was found. This surprising result stems from the fact that the spatiotemporal receptive field alone is a poor predictor of ganglion cell responses under naturalistic stimulation (van Hateren et al., 2002). The receptive field overlap is a good measure of the functional similarity between neurons only if it actually does a good job at predicting when each neuron spikes. Since the prediction is poor, receptive-field overlap is an unreliable measure of functional similarity under natural conditions. The redundancy is calculated directly from actual spike trains, so it is does not suffer this inadequacy.

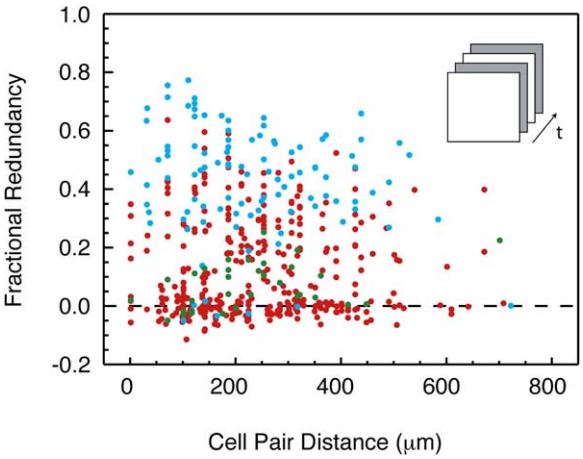

Figure 5. Redundancy under Strongly Correlated Stimulation

Fractional redundancy for 502 cell pairs stimulated by spatially uniform flicker plotted versus the distance between the electrodes on which each cell was recorded. Here all cells received identical stimulation. Color indicates functional type: pairs where both cells are the same functional type (blue), pairs where one cell is ON-type and the other is OFF-type (green), and other pairs where the cells are of different functional type (red).

Discussion

We have shown that the salamander retina uses a neural code that is extensively redundant to transmit visual information from eye to brain, with visual messages repeated many times within the population of nearby ganglion cells. In large part, this overrepresentation arises from within the circuitry of the retina, as it is observed even for stimuli having minimal spatial and temporal correlation. Redundant coding can therefore be

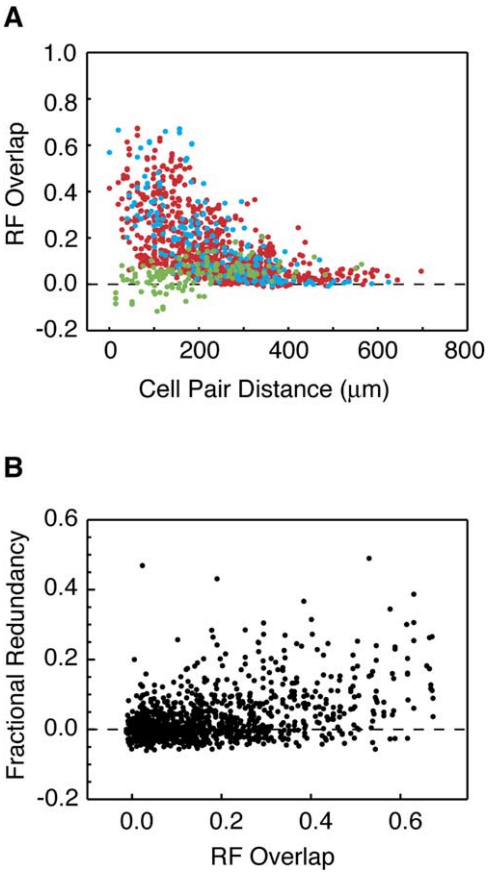

Figure 6. Redundancy and Receptive-Field Overlap

(A) Receptive-field overlap for 1176 cell pairs plotted against their receptive-field center distance. Color indicates functional type, as in Figure 3.

(B) Fractional redundancy plotted against receptive-field overlap for cells recorded under naturalistic stimulation.

thought of as a design choice made by the retina: many ganglion cells sample each point in visual space and convey similar information. In fact, we found extensive mixing of messages among cells assigned to different functional types—even with relatively simple stimuli, such as checkerboard flicker—so that a ganglion cell shared significant redundancy with almost half of its closest neighbors. Furthermore, the correlations present in natural visual scenes induce shared signaling among more widely spaced ganglion cells. In total, the average ganglion cell shared redundancy with over 100 other cells, and as a result, we estimate that the information encoded by an average ganglion cell is represented ~10 times over within the local population under naturalistic stimulus conditions.

Estimating the Population Redundancy

How does the overrepresentation factor relate to the information conveyed by a population of neurons? We can gain insight by considering a simple model of a population of *n* neurons with uniform pairwise correlations. We will calculate the population redundancy under two different, simple assumptions and compare them. The goal of this calculation is not to accurately capture all of the details of the ganglion cell code, but to explore how its population redundancy scales with the number of correlated neurons and the degree of pairwise redundancy.

First, we assume that every neuron conveys an information I about the stimulus and that all pairs have a redundancy of Δ. Furthermore, we assume that there are only pairwise correlations in the population (Martignon et al., 2000; Schneidman et al., 2003b). For n neurons, the population information is I_n, which is less than the sum of each cell's individual information, nI (Figure 7B, i and ii). The redundancy, measured as a fraction of the individual cell's visual information, is defined here as $\tilde{\Lambda}_n = 1 - I_n/nI$. When we add another neuron to the population, we gain information I, but some of this information is redundant. Since this new neuron shares Δ information with all of the previous neurons, one might naively estimate that $n\Delta$ is lost to redundancy. However, much of the $n\Delta$ is, in fact, the same information counted many times.

To estimate how much of the new neuron's information is already encoded by the existing neurons, we make the assumption that the redundancy within a large population of neurons is uniformly distributed. This means that the total possible redundancy, $n\Delta$, has the same overlap as that found within the existing population (Figure 7B, iii and iv). As a result of this uniformity assumption, only $n\Delta(I_n/nI)$ out of the new neuron's I of information is lost to redundancy. The information of the population of $n+1$ neurons is then given by a recursive relationship in terms of the information encoded by the population of n neurons (Figure 7C):

$$I_{n+1} = I_n + I - n\Delta\left(\frac{I_n}{nI}\right) = \left(1 - \frac{\Delta}{I}\right) I_n + I. \qquad (2)$$

As the population size increases, the fraction of the information I that is already represented by existing neurons increases until new neurons add no new information and the population information saturates at a value of $I_\infty = I^2/\Delta$. Equation 2 can be reexpressed as a relationship between the population redundancy for $n+1$ neurons and that present for n neurons, where the superscript u denotes the assumption of uniform correlation in the population:

$$\tilde{\Lambda}^{(u)}_{n+1} = \frac{n}{n+1}\left(1 - \frac{\Delta}{I}\right)\tilde{\Lambda}^{(u)}_n + \frac{n}{n+1}\frac{\Delta}{I}. \qquad (3)$$

An even simpler approximation to the population redundancy can be made using the overrepresentation factor. Since the overrepresentation for n neurons, Θ_n, measures how many times the information of one neuron is also encoded within the rest of the population, the population information is the sum of each cell's information divided by the overrepresentation, $I_n = nI/\Theta_n$. This leads to a population redundancy of:

$$\tilde{\Lambda}^{(O)}_n = 1 - \frac{1}{\Theta_n}, \qquad (4)$$

where the overrepresentation is $\Theta_n = (n - 1)\Delta/I$. This

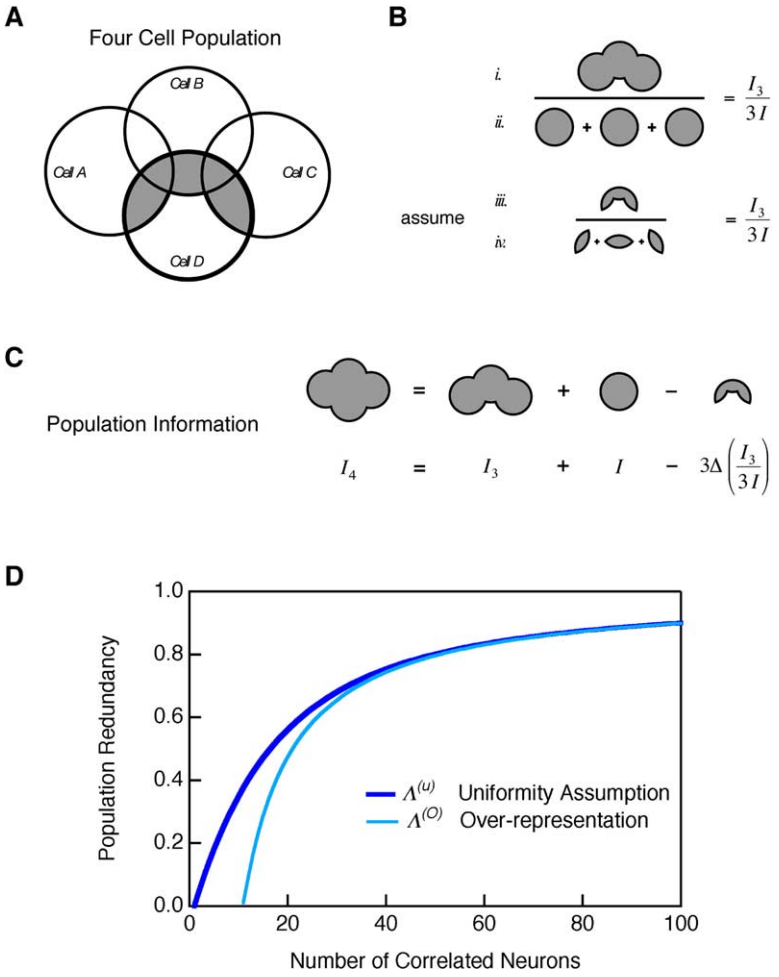

Figure 7. Estimates of the Population Redundancy

(A) Venn diagram showing the overlap in the information encoded by a population of four cells.

(B) Graphical description of the assumption of uniform correlations. (i and ii) The redundancy among the first three cells is the information they jointly encode (i) divided by the sum of their individual informations (ii). (iii and iv) The ratio between the part of the new cell's information that is redundant (iii) and the total possible redundancy (iv) is assumed to be equal to the redundancy among the first three cells.

(C) Recursive formula for the information encoded by all four cells, showing the terms in Equation 2.

(D) Redundancy in a simple model neural population plotted versus the number of neurons for an estimate based on the assumption of uniform redundancy within the population (dark blue line) and for an estimate based on the overrepresentation factor (light blue line). The parameters of the model are roughly matched to experimental values: n = 100 and $\Delta I = 0.1$, giving an overrepresentation of $O_{100} = 9.9$ (see Figure 2). The population redundancy is $R_{100} = 0.9$, indicating that the population has 10 times less information than the sum of each cell's information.

formula only makes sense in the limit of many neurons, where $\Theta_n > 1$.

These two measures of population redundancy differ when there are few neurons in the population, but they both reach the same value for many neurons (Figure 7D). This indicates that in the limit of large overrepresentation, $\Theta_n >> 1$, the population information is simply related to the overrepresentation factor, $I_n \approx nI/\Theta_n$, for a population of uniformly correlated neurons. What this calculation illustrates is that correlations of moderate strength shared among many cells can significantly reduce the population information. This results from the fact that a group of n neurons has $n(n-1)/2$ cell pairs, so that the number of pairs grows rapidly as the population size increases. Therefore, to understand the impact of pairwise correlations on a population code, one should consider the overrepresentation factor, which is a sum over pairwise redundancies, not just the average pairwise redundancy. Such an analysis predicts that the retinal code is ~10-fold redundant.

The calculation described above is only an approximation to the effect that pairwise redundancy has on the information encoded by a population of neurons. The real answer will depend on the detailed relationship between all pairwise correlations. In addition, the population information may be altered by correlations among three or more neurons that cannot be explained by existing pairwise correlations. Schneidman et al. (2003b) have described a formalism that gives a precise definition to the impact of pairwise correlations on the information conveyed by a population of neurons by reference to the maximum entropy distribution for n neurons consistent with all pairwise correlations. This formalism also defines the contribution to the population information from unique correlations at third order or higher. It will be interesting to determine how the connected information of second order described in this work relates to the overrepresentation factor.

Amphibian versus Mammalian Retinas

We carried out our experiments in the salamander retina. How might these results differ in the mammalian retina? Previous studies of ganglion cells in the mammalian retina reveal that the population has rather high overlap: 40–55 in the cat, depending on eccentricity (Peichl and Wassle, 1979), and ~40 in the rabbit in the far periphery (Rockhill et al., 2002). This degree of overlap makes it likely that the ganglion cell code in mammals also possesses extensive redundancy. Physiological measurements in the salamander give a total coverage of ~60 for receptive fields (Segev et al., 2004), similar to these mammalian retinas. Furthermore,

salamander ganglion cells have a center-surround spatial organization with roughly equal weight in center and surround (Smirnakis et al., 1997), indicating that they are expected to remove spatial correlations in natural scenes, as in mammalian retinas.

On the other hand, notable differences exist between salamander and mammalian retinas. In the salamander, we have identified six types of ganglion cells based on functional classification, but as many as 15–20 types of mammalian ganglion cells have been described, primarily using anatomy (Dacey, 1993; Masland, 2001; Wässle and Boycott, 1991). Thus, if there is greater specialization of function within the mammalian retina, its ganglion cell code may be less redundant than the salamander's. It is important to note, however, that not all of these many types of ganglion cells are expected to be functionally independent. For instance, brisk transient and brisk sustained cells are very likely to respond to many of the same stimulus patterns. Another distinction is that the salamander retina has a large fraction of ganglion cells that are OFF-type with very transient responses, and these ganglion cells tend to share high redundancy. Because the mammalian retina has many ganglion cells with more sustained responses and a more equal balance between ON and OFF cells, somewhat lower redundancy might be expected.

Redundancy in Other Neural Codes

Using the same definition of fractional redundancy as we used (Equation 1), Reich and Victor (Reich et al., 2001) found extensive redundancy among neurons in primary visual cortex of the macaque monkey during stimulation with flickering checkerboards. Notably, their values ranged up to ~50% between pairs of cells and had great variability, as seen for the retina under naturalistic stimulation (Figure 1). Petersen and Diamond explored how pairs of spikes both within the same cell and between pairs of cells contributed to the joint mutual information in primary somatosensory cortex of the rat in response to whisker plucks (Petersen et al., 2001; Pola et al., 2003). They found that pairs of cells in the same barrel were highly redundant (20% of the joint mutual information, on average), and those in different barrels also had significant redundancy. Perhaps it is quite common for neural circuits to use codes with extensive and overlapping redundancies between neurons.

Multineuronal Firing Patterns

Many previous studies have reported an enhanced tendency for nearby ganglion cells to fire spikes synchronously—both under constant illumination (Mastronarde, 1989) and visual stimulation (Meister et al., 1995). Such correlations suggest that the brain might interpret instances when two or more cells fire spikes synchronously as a different visual message than when the same cells spike individually. The receptive field for such synchronously spiking cell groups are smaller than the union of individual cell's receptive fields (Schnitzer and Meister, 2003), leading to the suggestion that recognizing synchrony among cells may allow the brain to achieve higher spatial resolution than otherwise possible and that the retinal code may thus be significantly synergistic (Meister, 1996). It is perhaps then surprising that pairs of ganglion cells almost never exhibited net synergy. Synchronous spike pairs were, however, relatively rare, so that they might serve as a synergistic coding symbol without necessarily leading to average synergy between cells. In addition, it is possible that other multineuronal firing patterns, such as combinations of spiking from some cells and silence from other cells, may constitute synergistic coding symbols.

Warland et al. (1997) have used linear decoding to study how a population of ganglion cells encode spatially uniform flicker. Identifying four functional types of neurons, they found that cells of the same type contributed almost no extra information to the population code (high redundancy), while the first cell of a different type contributed almost its entire individual information (low redundancy). A similar study using flickering checkerboards again found that cell pairs with large receptive-field overlap were strongly redundant (D. Warland, personal communication), in agreement with our results (Figure 3). Furthermore, recognizing synchronous spikes as distinct coding symbols did not increase the information captured by a linear decoder. However, a linear decoder is only one possible algorithm with which to extract information from a spike train, and the information extracted by such a decoder is only a *lower bound* on the encoded information. Thus, this technique is not as well suited to evaluating the synergy or redundancy of the neural code as the direct method (Schneidman et al., 2003a). Dan et al. (1998) studied the role of synchronous spikes for the transmission of information from LGN to visual cortex, finding that a linear decoder could extract up to 40% more information about flickering checkerboards if synchronous spikes were treated as distinct coding symbols. This result suggests that noise correlations contribute net synergy to this neural code, in contrast to what Warland et al. (1997) found in the retina.

Recently, Nirenberg et al. (2001) used information theoretic techniques to study the importance of noise correlations for the retinal code, concluding that ganglion cells were largely independent encoders. Our study differs from theirs in two important ways. First, we have reported redundancy, which is the direct measure of whether neurons are independent encoders. The redundancy, which includes both signal and noise correlations, was found to be as high as 50% of the information conveyed by a single cell (Figure 1). Second, we have studied the strength of correlations not just at the level of cell pairs, but also at the population level. Because ganglion cells share redundancy with many nearby cells, we find that the retinal code has extensive redundancy (Figure 2).

Efficiency of the Retinal Code

Theories of efficient coding predict that processing within sensory systems acts to remove correlations present in the natural world and reduce the redundancy of the neural code. The fact that redundancy was present for cell pairs with nonoverlapping receptive-field centers (Figures 1 and 3) shows that the retina achieves only partial redundancy reduction. However, theories

that have formalized the notion of efficient coding predict perfect redundancy reduction only when the noise in the system is negligible (Atick, 1992; Laughlin, 1989; van Hateren, 1992). In fact, most cell pairs with nonoverlapping receptive-field centers were independent, so that the average redundancy of such cell pairs was not very large (Figure 2). Furthermore, the theory of Atick and Redlich relies on several key assumptions: that the statistics of natural scenes are captured by their spatial power spectrum, that ganglion cells convey information using a firing rate code, that retinal processing is described by the action of a linear filter, and that ganglion cell-receptive fields tile visual space. We and others have found that none of these conditions actually hold (Berry et al., 1997; Dong and Atick, 1995; Peichl and Wassle, 1979; Ruderman, 1994), so that many important factors influence the efficiency of the retinal population code in ways that are not yet understood.

The Uses of Redundancy

Although the principle of redundancy reduction has been highly influential in neuroscience in the 50 years since it was first proposed by Attneave and Barlow (Attneave, 1954; Barlow, 1961), Barlow himself recently revisited this notion critically (Barlow, 2001). There are, in fact, several advantages to a redundant neural code. Most obviously, such a code can be highly tolerant to the misfiring of a single ganglion cell. The surprising degree of overrepresentation that we have observed indicates that there is great scope for the brain to use error-correction mechanisms to ensure that behaviorally vital information is not corrupted by noise. Redundancy may also constitute a form of passive attention or salience map, as it allows highly overrepresented visual information to have a stronger impact on downstream neurons than other visual information that is not overrepresented. Perhaps the ability to rank the importance of different stimuli with such a saliency map more than compensates for the coding inefficiency it requires. Finally, redundancy among neurons also confers the potential to form combinatorial codes, where unique stimulus features can be represented by the difference in the responses of two or more neurons. Such combinatorial codes may allow high-order features to be extracted more simply by subsequent neural circuits than would be possible with a highly efficient code. We should also keep in mind, though, that retinal redundancy may serve no positive purpose.

Whatever the possible advantages of redundancy, it clearly can be achieved only at a cost in efficiency: more neurons and more spikes are deployed to represent the same information than are truly necessary. While there has been a great focus on efficiency as a fundamental design principle for neural codes, robustness is less well understood (but see LeMasson et al., 1993). Quantifying our intuitive notion of robustness, as has recently been demonstrated for the chemical network underlying bacterial chemotaxis (Barkai and Leibler, 1997), promises to enrich our understanding of design principles in neural networks. Especially interesting will be to explore how redundancy and efficiency trade off as the signal-to-noise ratio of visual stimuli changes, as it does, for instance, with changes in the mean light level (Atick, 1992; Laughlin, 1989; van Hateren, 1992).

Experimental Procedures

Physiological Preparation

Larval tiger salamanders were light adapted for 1 to 3 hr prior to removal of the eye. A section of the eye one-quarter to one-third of the original size was isolated. The preparation left the retina, pigment epithelium, and sclera as an intact unit. The isolated section was placed ganglion side down on a microelectrode array and held in place with a dialysis membrane. The sample chamber was continuously perfused with Ringer's solution at room temperature (Balasubramanian and Berry, 2002). Recordings of 12 to 24 hr were consistently achieved. This study is based on data from five different retinal preparations.

Visual Stimulation

Natural movie clips were acquired using a Canon Optura Pi video camera at 30 frames per second. The spatial power spectrum of movie clips had an approximately power-law dependence on spatial frequency, $S(k) \sim k^{-\alpha}$. The exponent, α, ranged from 2.05 to 2.30 in our set of movies. The temporal power spectrum also followed a power law, $S(\omega) \sim \omega^{-\beta}$, with an exponent ranging from 0.84 to 1.82. The average luminosity at the retina of all movies was within 15% percent of the total mean value of 8 Lux (12 mW/m^2). The original color movies were converted to gray scale, incorporating a gamma correction for the computer monitor. All visual stimuli were displayed on an NEC FP1370 monitor and projected onto the retina from the ganglion cell side with standard optics. Movies were displayed at 60 Hz with 2 repeated frames. A display pixel was 11 μm by 11 μm when projected onto the retina. Movie clips covered the entire piece of dissected retina.

During visual optic flow, the retina sees a range of radial object motions (defined as the radial speed of passing edges) that depend monotonically on the distance of the object from the center of view. This can be thought of as a range of angular speeds starting at zero in the center of the field of view to some maximum in the periphery. In the optic flow movies shown here, this maximum angular speed is 5°/s for the region of the movie viewed by the ganglion cells. In saccadic movies, rapid camera movements occurred at a rate of 0.75 ± 0.2 saccades/s. The typical duration of each simulated saccade was 0.2 ± 0.1 s with angular speeds averaging 150°/s ± 50°/s. In between simulated saccades, the camera was still and there was no intrinsic motion in the scene. In smooth pursuit, the camera was panned slowly and steadily across a static scene with a speed of about 2°/s.

In the checkerboard random flicker, square regions (55 μm on the retina) were randomly chosen to be either black or white on a fast time scale (every 33 ms), giving the impression of flicker everywhere in space. In spatially uniform flicker, intensity values for the entire screen were randomly chosen every 33 ms. Natural movie clips of 15 to 30 s duration as well as checkerboard and uniform flicker sequences of 30 s duration were presented to the retina between 100 and 300 times for redundancy calculations. With the exception of one experiment, each retinal preparation was presented with many stimuli: a collection of movies, repeated checkerboard flicker, repeated uniform flicker, and nonrepeated checkerboard to map receptive fields.

Electrophysiological Recording

Extracellular voltages were recorded by a MultiChannel Systems MEA 60 microelectrode array and amplified with band pass filtering of ~1 Hz to 3 kHz. Experiments used either a square microelectrode array with 10 μm diameter electrodes spaced 100 μm apart, or a hexagonal array with a spacing of 30 μm in the center and 90 μm for the outermost ring. Custom recording hardware and software based on a National Instrument MIO 64 A/D PCI card and LabView software streamed data to disk with a sampling rate of 10 kSamples/s. Voltage data from each of the 60 electrodes were sorted into spike trains of single ganglion cells based on the overall spike size and shape in a 2.5 ms window. Only spike waveforms

that were clearly distinguished from the noise were used. We required that isolated spike trains had fewer than 0.5% of their interspike intervals shorter than 2 ms. In order to determine whether the same spikes were being counted on multiple channels, the crosscorrelation function between isolated spike trains was computed. Whenever there was a narrow (±0.2 ms) peak in the crosscorrelation that contained more than 1% of the spikes, one of the spike trains was disqualified.

Receptive-Field Analysis

To map the receptive fields of ganglion cells, we stimulated them with nonrepeated checkerboard flicker for 30–45 min and calculated the spike-triggered stimulus average (Schnitzer and Meister, 2003). The spike-triggered average, *RF(x,y,t)*, contains the spatial and temporal features of the stimulus that on average caused a cell to spike. The spatial profile showed center-surround antagonism, and the temporal dynamics displayed a mixture of temporal integration and differentiation. The spatial profile of the receptive-field center was fit with a two-dimensional Gaussian function, and the center coordinates, $\bar{x} = (x_0, y_0)$, were used to determine the distance between pairs of ganglion cells, $D_{ab} = |\bar{x}_a - \bar{x}_b|$. The receptive-field center was identified as all of the checkers having the same polarity as the checker with the largest value, and its temporal profile was calculated by summing over the temporal profile of all the checkers that belonged to the receptive-field center.

Receptive-Field Overlap

The receptive field can be viewed as a vector with separate components for each spatial location and time bin before the spike, giving a total of 80 × 80 × 32 components. To calculate the overlap between cells *a* and *b*, O_{ab}, we computed the inner product between their receptive fields and divided by the noise-corrected magnitude of each receptive-field vector:

$$O_{ab} = \frac{\vec{RF}_a \bullet \vec{RF}_b}{\sqrt{\|\vec{RF}_a\|^2 - N_a\sigma_a^2}\sqrt{\|\vec{RF}_b\|^2 - N_b\sigma_b^2}}. \quad (5)$$

The overlap has a maximum value of +1 when the two receptive fields are identical, a minimum value of –1 when they are of opposite polarity but otherwise the same, and a value close to 0 when they are far apart in space or very different in their temporal dynamics. The noise correction term σ was defined as the standard deviation per pixel far in space and/or time from the maximum of the spike-triggered average, and *N* was the number of pixels. This correction compensated for the overestimation in magnitude of each receptive-field vector due to averaging over only a finite number of spikes. Noise correction was unnecessary for the numerator, because noise causes no bias in this inner product. In most cases, the noise correction resulted in roughly a ~10% change in the overlap.

Functional Classification

Following previous studies (DeVries and Baylor, 1997; Schnitzer and Meister, 2003), we assigned cells to functional classes based on the temporal dynamics of their receptive-field center. We used an agglomerative clustering algorithm to help inform our choice of functional classes. Receptive fields were compared by computing their functional overlap, using only the temporal dynamics of the receptive-field center (Equation 5). At the outset of our iterative algorithm, each cell formed its own cluster. First, we found the pair of clusters that had the largest overlap. Then, these two clusters were merged into a single cluster by averaging their receptive fields, weighted by the number of cells in each cluster. This procedure was repeated until all cells were merged into a single cluster. The significance of each merger was evaluated using the overlap between the two clusters. If the two clusters were very similar, then they will have an overlap close to +1, but if they were quite different, then the overlap will be significantly lower. Six cell classes could be identified, as shown by a break in the merger score as a function of number of clusters (Figure 4). Including the spatial profile of the receptive field had little effect on the assignment of functional classes. Similar results were found across multiple retinas. Our functional types corresponded closely with previous classification schemes. For instance, when tested with diffuse flashes of light, fast OFF, standard OFF, and medium OFF cells were ON/OFF-type; slow OFF cells were OFF-type; and fast ON and slow ON were ON-type.

Redundancy

Spike trains were discretized into the spike count in successive time bins of 10 ms; bins were allowed to have more than one spike. Next, spike words were formed by concatenating *K* such time bins. For the joint response of two cells, spike counts from each cell were concatenated together into a word of *2K* digits. The probability of finding the i^{th} spike word W_i at time *t* during the movie, $p(W_i|t)$, was compiled over repeated stimulus presentations, and the probability of finding the i^{th} spike word over the entire movie was found by averaging over time, $p(W_i) = \langle p(W_i|t)\rangle_t$. Entropies were estimated directly from the probability distribution of spike words:

$$H_{signal} = -\sum_i p(W_i)\log_2 p(W_i) \quad (6a)$$

$$H_{noise} = -\langle \sum_i p(W_i|t)\log_2 p(W_i|t)\rangle_t, \quad (6b)$$

and the mutual information between stimulus and response was the difference between signal and noise entropies:

$$I(R;S) = H_{signal} - H_{noise}.$$

Entropies were corrected for the finite size of the data set by dividing the data into S = 2, 4, and 8 blocks and using standard techniques (Strong et al., 1998). An extrapolation to infinite word length was made using word lengths ranging from 2 to 4 digits (Strong et al., 1998). Using these methods, we found signal entropies in the range of 5.2 to 13 bits/spike, with an average of 8.6 bits/spike; noise entropies in the range of 2.4 to 7.9 bits/spike, with an average of 4.6 bits/spike. A very wide range of firing rates was found under naturalistic stimulus conditions, from 0.06 up to 10.2 spikes/s with a median of 1.2 spikes/s; the information rates for individual cells correspondingly varied widely, from 0.35 to 21.8 bits/s with a median of 4.4 bits/s.

Random and Systematic Error in Redundancy Values

Signal entropies were much better sampled than noise entropies, and single-cell entropies were better sampled than joint entropies. As a result, any bias that remained after correction for finite data size tended to be largest for the joint noise entropy. Such a bias would decrease the measured redundancy. To estimate the size of this bias, we assumed a Poisson model of spike generation with a 5 ms absolute refractory period; the free firing rate in this model can be uniquely determined from the measured spike trains (Berry and Meister, 1998). With this model, we generated random spike trains with enough repeats to sample the noise entropy thoroughly. By comparing the information of simulated spike train having many repeats (1800) to the information for spike trains having the same number of repeats as used in our data (120–180), we found that the redundancy was underestimated by ~0.5%–1% under our experimental conditions.

The redundancy depends on the choice of time bin used to form spike words. We chose 10 ms because the timing precision of neurons was rarely any better, so that most of the details of the spike train on a finer temporal scale are noise. However, fluctuations in firing on a fine temporal scale may serve to further distinguish the responses of different neurons and thus decrease their redundancy. On the other hand, excess synchrony can also exist on a fine temporal scale, and these correlations may increase the redundancy. We tested the effect of temporal scale by subdividing spike words of 10 or 20 ms duration into words of the same duration but having smaller time bins and recomputing the redundancy. When words with 2 bins of 10 ms were divided into 4 bins of 5 ms, the redundancy increased by a factor of 1.10 ± 0.06; when words with 1 bin of 10 ms were divided into 4 bins of 2.5 ms, the redundancy increased by a factor of 1.11 ± 0.11. These tests indicate that noise correlations on a fine temporal scale add about 10% to the redundancy.

The redundancy also depends on the total temporal duration of

the chosen spike words. Longer words are desirable, but words with many digits cannot be adequately sampled. Typically, the probability of joint responses could only be sampled with spike words up to 4 digits long. As a result, there is a trade-off with the choice of time bin: smaller time bins resolve synchrony on a fine timescale, but cannot sample correlations on a long timescale. Such correlations significantly affect the redundancy: when we used 5 ms bins instead of 10 ms bins, extrapolating up to 4 digit spike words in each case, the redundancy was smaller by a factor of 1.35 ± 0.10. This underscores the fact that forming spike words with a duration of 40 ms allows one to sample correlations not observed within 20 ms spike words. Pairs of ganglion cells often exhibited weak correlation on timescales greater than 40 ms, which are only partially captured by extrapolation to infinite word length. Our choice of time bin and word length reflects a good balance between resolving fine temporal structure and sampling long timescale correlations. Were we able to sample spike words with smaller time bins and many more digits, it is likely that the redundancy would be greater than our estimate.

The significance of our calculated redundancy values was determined by two different methods. First, we examined the distribution of redundancies found under checkerboard stimulation for cell pairs separated by more 250 μm (see Figure 3). Since such pairs of cells have little or no overlap of their receptive-field centers, we expect that they should mostly be independent. We fit a Gaussian to the distribution over these putatively independent cell pairs; the width of the Gaussian was ~2.5%. Second, we shifted the spikes of one cell relative to another by a constant, random time. Again, such a shift should render the cell pair nearly independent. We calculated the redundancy between cell pairs for many values of this relative time shift; the standard deviation across different time shifts was 3%. We chose a conservative threshold of 5% redundancy for significance, as the probability that independent cells would have a larger apparent redundancy by chance is $p < 0.01$.

Overrepresentation Factor

The overrepresentation factor measures how many times, on average, the information encoded by one ganglion cell is also encoded by neighboring cells. For a single cell a, this quantity is $\theta_a = \Sigma_b \Lambda_{ab}/I_a$, where Λ_{ab} is the (nonnormalized) redundancy and I_a is the information, both measured in bits. Because we record only a small fraction of the ganglion cells in a patch of retina, this quantity is dominated by the details of which cells we happened to sample. Instead, we pooled all of our measurements to estimate the total overrepresentation for the population. For this, we averaged all values of redundancy for cells spaced by a distance r, $\theta(r) = \langle \Lambda_{ab}/I_a \rangle_{|\vec{x}_a - \vec{x}_b| = r}$, where the sum is over a and b, so that the individual information of each cell appears once in the denominator in this average. The total overrepresentation was given by the sum of this contribution over all distances up to the maximum for which we had sufficient data, $\Theta = \int_0^{r_{max}} \theta(r)\, 2\pi r\, n_{cell}\, dr$, where the cell density n_{cell} was taken to be 1400 cells/mm^2 (Segev et al., 2004) and the integral was evaluated with a bin size of 25 μm. At large separations, the few cell pairs that were redundant tended to have values close to the significance threshold, so that a calculation of the overrepresentation at this range had great uncertainty. As a result, we included only cell pairs separated by up to r_{max} = 500 μm in our calculation. Error bars were determined by randomly selecting half of the cell pairs and repeating the calculation for many such selections. Since the overrepresentation factor is an average quantity, it is still possible for an individual ganglion cell to encode some unique information even when the overrepresentation factor is larger than 1.

Acknowledgments

This work was supported by NIH Grant R01 EY14196, a grant from the E. Matilda Ziegler Foundation, a Pew Scholars award to M.J.B., a Bristol-Meyer Squibb/Lewis Thomas Fellowship to J.L.P., and a Wellcome Trust Fellowship to R.A.H. We thank Bill Bialek, Adrienne Fairhall, and Blaise Agüera y Arcas for many useful conversations and Markus Meister for comments on the manuscript.

Received: July 2, 2004
Revised: January 14, 2005
Accepted: March 17, 2005
Published: May 4, 2005

References

Atick, J.J. (1992). Could information theory provide an ecological theory of sensory processing? Network *3*, 213–251.

Atick, J.J., and Redlich, A.N. (1992). What does the retina know about natural scenes? Neural Comput. *4*, 196–210.

Attneave, F. (1954). Some informational aspects of visual perception. Psychol. Rev. *61*, 183–193.

Balasubramanian, V., and Berry, M.J., II (2002). A test of metabolically efficient coding in the retina. Network *13*, 531–552.

Barkai, N., and Leibler, S. (1997). Robustness in simple biochemical networks. Nature *387*, 913–917.

Barlow, H.B. (1961). Possible principles underlying the transformation of sensory messages. In Sensory Communication, W.A. Rosenblith, ed. (Cambridge, MA: MIT Press), pp. 217–234.

Barlow, H. (2001). Redundancy reduction revisited. Network *12*, 241–253.

Barlow, H.B., Fitzhugh, R., and Kuffler, S.W. (1957). Change of organization in the receptive fields of the cat's retina during dark adaptation. J. Physiol. *137*, 338–354.

Berry, M.J., II, and Meister, M. (1998). Refractoriness and neural precision. J. Neurosci. *18*, 2200–2211.

Berry, M.J., II, Warland, D.K., and Meister, M. (1997). The structure and precision of retinal spike trains. Proc. Natl. Acad. Sci. USA *94*, 5411–5416.

Dacey, D.M. (1993). The mosaic of midget ganglion cells in the human retina. J. Neurosci. *13*, 5334–5355.

Dacey, D.M., Peterson, B.B., Robinson, F.R., and Gamlin, P.D. (2003). Fireworks in the primate retina: in vitro photodynamics reveals diverse LGN-projecting ganglion cell types. Neuron *37*, 15–27.

Dan, Y., Alonso, J.M., Usrey, W.M., and Reid, R.C. (1998). Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. Nat. Neurosci. *1*, 501–507.

DeVries, S.H., and Baylor, D.A. (1997). Mosaic arrangement of ganglion cell receptive fields in rabbit retina. J. Neurophysiol. *78*, 2048–2060.

Dong, D.W., and Atick, J.J. (1995). Statistics of natural time-varying images. Network *6*, 345–358.

Field, D.J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. J. Opt. Soc. Am. A *4*, 2379–2394.

Frechette, E.S., Sher, A., Grivich, M.I., Petrusca, D., Litke, A.M., and Chichilnisky, E.J. (2004). Fidelity of the ensemble code for visual motion in primate retina. J. Neurophysiol. Published online December 29, 2004. 10.1152/jn.01175.2004.

Gat, I., and Tishby, N. (1999). Synergy and redundancy among brain cells of behaving monkeys. In Advances in Neural Information Processing Systems, M. Kearns, S. Solla, and D. Cohn, eds. (Cambridge, MA: MIT Press), pp. 465–471.

Gawne, T.J., and Richmond, B.J. (1993). How independent are the messages carried by adjacent inferior temporal cortical neurons? J. Neurosci. *13*, 2758–2771.

Hartline, H.K. (1937). The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. Am. J. Physiol. *121*, 400–415.

Laughlin, S.B. (1989). The role of sensory adaptation in the retina. J. Exp. Biol. *146*, 39–62.

LeMasson, G., Marder, E., and Abbott, L.F. (1993). Activity-dependent regulation of conductances in model neurons. Science *259*, 1915–1917.

Martignon, L., Deco, G., Laskey, K., Diamond, M., Freiwald, W., and Vaadia, E. (2000). Neural coding: higher-order temporal patterns in

the neurostatistics of cell assemblies. Neural Comput. *12*, 2621–2653.

Masland, R.H. (2001). The fundamental plan of the retina. Nat. Neurosci. *4*, 877–886.

Mastronarde, D.N. (1989). Correlated firing of retinal ganglion cells. Trends Neurosci. *12*, 75–80.

Meister, M. (1996). Multineuronal codes in retinal signaling. Proc. Natl. Acad. Sci. USA *93*, 609–614.

Meister, M., Lagnado, L., and Baylor, D.A. (1995). Concerted signaling by retinal ganglion cells. Science *270*, 1207–1210.

Nirenberg, S., Carcieri, S.M., Jacobs, A.L., and Latham, P.E. (2001). Retinal ganglion cells act largely as independent encoders. Nature *411*, 698–701.

Olveczky, B.P., Baccus, S.A., and Meister, M. (2003). Segregation of object and background motion in the retina. Nature *423*, 401–408.

Pang, J.J., Gao, F., and Wu, S.M. (2002). Segregation and integration of visual channels: layer-by-layer computation of ON-OFF signals by amacrine cell dendrites. J. Neurosci. *22*, 4693–4701.

Pang, J.J., Gao, F., and Wu, S.M. (2004). Stratum-by-stratum projection of light response attributes by retinal bipolar cells of *Ambystoma*. J. Physiol. *558*, 249–262.

Panzeri, S., and Schultz, S.R. (2001). A unified approach to the study of temporal, correlational, and rate coding. Neural Comput. *13*, 1311–1349.

Peichl, L., and Wassle, H. (1979). Size, scatter and coverage of ganglion cell receptive field centres in the cat retina. J. Physiol. *291*, 117–141.

Petersen, R.S., Panzeri, S., and Diamond, M.E. (2001). Population coding of stimulus location in rat somatosensory cortex. Neuron *32*, 503–514.

Pola, G., Thiele, A., Hoffmann, K.P., and Panzeri, S. (2003). An exact method to quantify the information transmitted by different mechanisms of correlational coding. Network *14*, 35–60.

Reich, D.S., Mechler, F., and Victor, J.D. (2001). Independent and redundant information in nearby cortical neurons. Science *294*, 2566–2568.

Rockhill, R.L., Daly, F.J., MacNeil, M.A., Brown, S.P., and Masland, R.H. (2002). The diversity of ganglion cells in a mammalian retina. J. Neurosci. *22*, 3831–3843.

Roska, B., and Werblin, F. (2001). Vertical interactions across ten parallel, stacked representations in the mammalian retina. Nature *410*, 583–587.

Ruderman, D.L. (1994). The statistics of natural images. Network *5*, 517–548.

Ruderman, D.L. (1997). Origins of scaling in natural images. Vision Res. *37*, 3385–3398.

Schneidman, E., Bialek, W., and Berry, M.J., II. (2003a). Synergy, redundancy, and independence in population codes. J. Neurosci. *23*, 11539–11553.

Schneidman, E., Still, S., Berry, M.J., II, and Bialek, W. (2003b). Network information and connected correlations. Phys. Rev. Lett. *91*, 238701.

Schnitzer, M.J., and Meister, M. (2003). Multineuronal firing patterns in the signal from eye to brain. Neuron *37*, 499–511.

Segev, R., Goodhouse, J., Puchalla, J.P., and Berry, M.J., II. (2004). Recording spikes from a large fraction of the ganglion cells in a retinal patch. Nat. Neurosci. *7*, 1155–1162.

Shapley, R., and Victor, J.D. (1979). The contrast gain control of the cat retina. Vision Res. *19*, 431–434.

Smirnakis, S.M., Berry, M.J., Warland, D.K., Bialek, W., and Meister, M. (1997). Adaptation of retinal processing to image contrast and spatial scale. Nature *386*, 69–73.

Sterling, P. (1998). Retina. In Synaptic Organization of the Brain, G. Shepard, ed. (New York: Oxford University Press).

Strong, S.P., Koberle, R., de Ruyter van Steveninck, R.R., and Bialek, W. (1998). Entropy and information in neural spike trains. Phys. Rev. Lett. *80*, 197–200.

van der Schaaf, A., and van Hateren, J.H. (1996). Modelling the power spectra of natural images—statistics and information. Vision Res. *36*, 2759–2770.

van Hateren, J.H. (1992). A theory of maximizing sensory information. Biol. Cybern. *68*, 23–29.

van Hateren, J.H., Ruttiger, L., Sun, H., and Lee, B.B. (2002). Processing of natural temporal stimuli by macaque retinal ganglion cells. J. Neurosci. *22*, 9945–9960.

Vaney, D.I. (1994). Territorial organization of direction-selective ganglion cells in rabbit retina. J. Neurosci. *14*, 6301–6316.

Victor, J.D., and Shapley, R.M. (1979). The nonlinear pathway of Y ganglion cells in the cat retina. J. Gen. Physiol. *74*, 671–689.

Warland, D.K., Reinagel, P., and Meister, M. (1997). Decoding visual information from a population of retinal ganglion cells. J. Neurophysiol. *78*, 2336–2350.

Wässle, H., and Boycott, B.B. (1991). Functional architecture of the mammalian retina. Physiol. Rev. *71*, 447–480.

Wu, S.M., Gao, F., and Maple, B.R. (2000). Functional architecture of synapses in the inner retina: segregation of visual signals by stratification of bipolar cell axon terminals. J. Neurosci. *20*, 4462–4470.