Dependence of the retinal Ganglion cell's responses on local textures of natural scenes

Xiwu Cao

David K. Merwine

Norberto M. Grzywacz

An important property of natural images contributing to a retinal ganglion cell's (RGC) responses is the temporal modulation of mean intensity (contrast) in the receptive field (RF) center. However, these responses exhibit a significant amount of variability. This variability could arise in part from responses to the spatial intensity variation of the natural images in the RF center, i.e., their local intensity distribution or their local visual texture. We tested five predictions derived from this hypothesis: First, responses tend to increase with the variance of the local visual texture of natural images. Second, the skewed distribution of intensities in natural images leads to asymmetric responses to their onset and offset to and from a gray background of the same mean intensity. Third, repeating this experiment with the negative of natural images inverts the asymmetry. Fourth, performing an intensity histogram equalization of the images eliminates the asymmetry. Fifth, RGCs' responses increase with the spatial contrast of artificial plaids. The hypothesis passed all five tests, which indicate that responses to natural images increase with the variance of their visual texture. To account for this texture sensitivity, we propose a model in which the RFs of most RGCs of the rabbit have multiple nonlinear subunits.

Keywords: retinal ganglion cell, receptive field, texture of natural images, nonlinear, modeling


Introduction

Scenes encountered in nature have many complex statistical characteristics. For example, natural scenes usually contain a high degree of spatial and temporal correlation (Balboa & Grzywacz, 2003; Field, 1987; Ruderman & Bialek, 1994), the power spectra of the scenes tend to fall with the square of spatial frequency (Balboa & Grzywacz, 2003; Burton & Moorhead, 1987; Tolhurst, Tadmor, & Chao, 1992; van der Schaaf & van Hateren, 1996), and their distribution of intensities has skewness, having a median below the mean and a long tail towards high intensities (Brady & Field, 2000; Nual & David, 2000; Olshausen & Field, 2000; Simoncelli & Olshausen, 2001). All these special statistical properties are thought to have had a substantial influence, through evolutionary adaptation, on the design of our visual system, including the retina (Ahmad, Klug, Herr, Sterling, & Schein, 2003; Atick, 1992; Enroth-Cugell & Robson, 1966; Koch, McLean, & Segev, 2006).

However, much of vision research on retinal ganglion cells (RGCs) is performed by using non-natural stimuli, such as flashing spots, drifting gratings, and Gaussian white noise (Carandini et al., 2005; Chichilnisky, 2001; Kaplan & Benardete, 2001; Kim & Rieke, 2001; Meister & Berry, 1999; Reid, Victor, & Shapley, 1997; Rust & Movshon, 2005). A well-studied computational model derived from such artificial stimuli provides a linear–nonlinear account of the RGC’s responses (Carandini et al., 2005; Chichilnisky, 2001; Kaplan & Benardete, 2001; Meister & Berry, 1999; Paninski, Pillow, & Simoncelli,
of these tests has appeared in abstract form (Cao, Merwine, & Grzywacz, 2008). Thus, this model comprises a weighted sum of light-stimulus intensities in the cell’s receptive field (RF), followed by a static nonlinearity. But we still have no good evidence that these models fit the RGC’s responses to natural stimuli (Carandini et al., 2005), and one should not simply extrapolate the knowledge obtained from artificial stimuli to natural stimuli (David, Vinje, & Gallant, 2004; Felsen & Dan, 2005; Simoncelli & Olshausen, 2001; Sterling, 2004).

Therefore, we recently began investigating the responses of RGCs to natural images. We reported that the relative modulation of mean intensity of natural stimuli in the RGC’s RF center (contrast) is an important parameter in determining the cell’s responses. However, as we will show in this paper, although the mean intensity in the RF center explains a large fraction of the variance in the RGC’s responses, one cannot account for much of their variability with just the mean. One possible contributing factor to this variability is the intensity variation of natural stimuli, specifically their local intensity distribution in the RF center. In this paper, we call this local intensity distribution the “local visual texture.” To examine this visual-texture hypothesis for response variability, we derived and tested five different predictions. We also built a nonlinear–linear (NL) model for the texture sensitivity of RGCs observed in the experimental tests. An initial report of these tests has appeared in abstract form (Cao, Merwine, & Grzywacz, 2008).

**Methods**

**Physiological preparation**

The preparation was similar to that used elsewhere (Chatterjee, Merwine, Amthor, & Grzywacz, 2007). Briefly, pigmented rabbits weighing 2–4 kg were dark adapted for 30 minutes before surgery. Surgery was conducted under dim red light. Rabbits were initially anesthetized with I.P. Ketamine (50 mg/kg) and Xylazine (5 mg/kg). Then, sodium pentobarbital (1 ml/kg) was injected into the marginal ear vein to obtain deep anesthesia. Anesthesia was checked by testing corneal reflexes and reactions to a paw pinch. After confirming anesthesia, an eye was enucleated. The eyecup was then flipped and mounted in a recording chamber. A metal ring gently held the paper in place. Once in the chamber, we continuously superfused the retina with oxygenated bicarbonate-buffered Ames solution (Sigma) at 37°C at a flow rate of 3–7 ml/min. The isolated retina remained healthy for at least 6 hr post isolation. In total, 152 RGCs from 6 retinas were obtained for analysis. We classified cells in terms of On, Off, transient, sustained, brisk, and sluggish properties (Merwine, Amthor, & Grzywacz, 1995). The results were generally true for all cell types.

**Electrophysiological recording and visual stimulation**

Methods for both recording with a multi-electrode array (Cyberkinetics) and sorting spikes using the POWERNAP software (Cyberkinetics) were as previously described (Chatterjee et al., 2007). We also described in that same publication general methods of visual stimulation, of determination of the RGCs’ properties, of measurement of spontaneous firing rate, and of mapping of RFs.

The natural images used in this paper were obtained from an online calibrated-image database (van Hateren & Ruderman, 1998). We further calibrated them to be linear with respect to the luminance of our monitor. The source images with 1536 × 1024 pixels were cropped down to either their square central regions or into eight evenly spaced regions with 300 × 300 pixels. Each region was then down-sampled into a square image of 120 × 120 pixels (3 × 3 mm² on the retina). Images were presented in a random sequence, each for 1000 ms. When the natural image was removed, the display was held at a spatially uniform gray for 1000 ms before the next image was presented. The mean intensities of the gray and natural images were the same (9.10 cd/m²) to remove luminance adaptation. The long periods of presentation of gray and natural images allowed us to distinguish clearly between responses to onset and offset of the images (Smyth, Willmore, Baker, Thompson, & Tolhurst, 2003). We presented from 1,000 to 12,000 images in different experiments.

**Data analysis**

**Linear receptive field**

The RGC’s linear RF with natural image stimulation was estimated by using either spike-triggered average (STA) or regularized STA (Willmore & Smyth, 2003). Briefly, given a RGC’s responses to natural images, its linear RF from STA is simply the average of the light stimuli weighed by their elicited responses. Such an STA would yield a filter predictive of linear cells if estimated with white-noise inputs (Marmarelis & Marmarelis, 1978). If the inputs are not white noise, and include

**Methods**

**Physiological preparation**

The preparation was similar to that used elsewhere (Chatterjee, Merwine, Amthor, & Grzywacz, 2007). Briefly, pigmented rabbits weighing 2–4 kg were dark adapted for 30 minutes before surgery. Surgery was conducted under dim red light. Rabbits were initially anesthetized with I.P. Ketamine (50 mg/kg) and Xylazine (5 mg/kg). Then, sodium pentobarbital (1 ml/kg) was injected into the marginal ear vein to obtain deep anesthesia. Anesthesia was checked by testing corneal reflexes and reactions to a paw pinch. After confirming anesthesia, an eye was enucleated. The eyecup was then flipped and mounted in a recording chamber. A metal ring gently held the paper in place. Once in the chamber, we continuously superfused the retina with oxygenated bicarbonate-buffered Ames solution (Sigma) at 37°C at a flow rate of 3–7 ml/min. The isolated retina remained healthy for at least 6 hr post isolation. In total, 152 RGCs from 6 retinas were obtained for analysis. We classified cells in terms of On, Off, transient, sustained, brisk, and sluggish properties (Merwine, Amthor, & Grzywacz, 1995). The results were generally true for all cell types.

**Electrophysiological recording and visual stimulation**

Methods for both recording with a multi-electrode array (Cyberkinetics) and sorting spikes using the POWERNAP software (Cyberkinetics) were as previously described (Chatterjee et al., 2007). We also described in that same publication general methods of visual stimulation, of determination of the RGCs’ properties, of measurement of spontaneous firing rate, and of mapping of RFs.

The natural images used in this paper were obtained from an online calibrated-image database (van Hateren & Ruderman, 1998). We further calibrated them to be linear with respect to the luminance of our monitor. The source images with 1536 × 1024 pixels were cropped down to either their square central regions or into eight evenly spaced regions with 300 × 300 pixels. Each region was then down-sampled into a square image of 120 × 120 pixels (3 × 3 mm² on the retina). Images were presented in a random sequence, each for 1000 ms. When the natural image was removed, the display was held at a spatially uniform gray for 1000 ms before the next image was presented. The mean intensities of the gray and natural images were the same (9.10 cd/m²) to remove luminance adaptation. The long periods of presentation of gray and natural images allowed us to distinguish clearly between responses to onset and offset of the images (Smyth, Willmore, Baker, Thompson, & Tolhurst, 2003). We presented from 1,000 to 12,000 images in different experiments.

**Data analysis**

**Linear receptive field**

The RGC’s linear RF with natural image stimulation was estimated by using either spike-triggered average (STA) or regularized STA (Willmore & Smyth, 2003). Briefly, given a RGC’s responses to natural images, its linear RF from STA is simply the average of the light stimuli weighed by their elicited responses. Such an STA would yield a filter predictive of linear cells if estimated with white-noise inputs (Marmarelis & Marmarelis, 1978). If the inputs are not white noise, and include
considerable spatial or temporal correlations, as for example, in natural images, then, one has to implement a further regularization based on the autocorrelation matrix of the light stimuli (regularized STA—Theunissen, Sen, & Doupe, 2000; Willmore & Smyth, 2003). In brief, given a set of natural stimuli $S$ and their corresponding responses $R$, the cell’s linear RF from STA is calculated by:

$$STA = (S^T S + \alpha I)^{-1}S^T R,$$

where $S^T S$ is the mean autocorrelation matrix for the stimuli, $I$ is the identity matrix, and $\alpha$ (usually $0.5–10$) is the regularization parameter. The good choice of $\alpha$ is dependent on the noise of the data (Calvetti & Reichel, 2004). An optimal $\alpha$ was not necessary in this paper, since we did not apply STA to quantify the true linear RF. Rather, we used STA as a tool to study the asymmetry between onset and offset of natural images. This asymmetry held for a large range of $\alpha$.

Finally, because we calibrate the mean of all images to be equal to the gray background, the linear RFs estimated from STA average out to zero. Thus, a strong excitatory center in the RF will compensate to accompany a large artificial inhibitory surround. In other words, one should not interpret the surround observed in the STA as meaningful in our experiments.

**Center-mean-contrast-to-response model and quantification of visual texture of natural images**

The center mean contrast (local temporal contrast, in our experiments) of a natural image was defined by

$$C = \frac{M_{\text{Center}} - M_{\text{Gray}}}{M_{\text{Gray}}},$$

where $M_{\text{Center}}$ was the mean intensity of the natural image in the RF center and $M_{\text{Gray}}$ was the mean intensity of the full-field gray background. The center of the RF can be roughly measured by using either a moving edge (Chatterjee et al., 2007) or the linear RF (STA—Willmore & Smyth, 2003). Although a precisely estimated RF center size from the natural images is preferred, it is not required in this study. For example, compared to the true RF center, the center size estimated from an artificial moving edge might be a little smaller, and that estimated from STA might be a little larger. However, the conclusions obtained in this study hold for both estimations of RF.

The center-mean-contrast-to-response relationship ($R_C$) was sigmoidal and modeled well by a cumulative Gaussian function. This function was defined by (Chichilnisky, 2001):

$$R_C = \frac{R_{\text{max}}}{2} \left[ 1 + \text{erf} \left( \frac{C - C_{\text{max}}}{\sigma \sqrt{2}} \right) \right],$$

where $R_{\text{max}}$ was the maximal firing rate, $C_{\text{max}}$ was the contrast value yielding maximal slope, $\sigma$ was the standard deviation of the Gaussian, and $\text{erf}$ is the Gauss error function. A criterion of minimum mean-squared error was applied to obtain the optimal fitting of this function.

As mentioned in the Introduction, we use the term “local visual texture” to mean the intensity variation in the RF center. More specifically, we mean variation away from the mean intensity. Several measures are available to quantify the strength of the visual texture, such as those based on the standard deviation, the entropy, and the histogram of intensities minus the mean. In this study, we used both the standard deviation and the histogram of intensities in the RF center. We describe the histogram method in more detail in the next subsection.

**Response-to-intensity-variation (RIV) function**

We developed a new approach both to test whether local visual textures of a natural image contribute to the response of the RGC and to quantify their strength. We based this approach on what we call the response-to-intensity-variation (RIV) function.

In this method, we assume that one can decompose the response to a natural image ($R_{\text{Total}}$) in two parts. The first part is the response contribution of the change in the local mean intensity of the image, or the local temporal contrast ($R_C$, see Equation 2). The second part is the contribution of the local intensity variation ($R_V$), shown in Figure 1. The local intensity variation is represented by the histogram of relative intensities in the RF center for each image. To define these histograms, let the intensity of a pixel be $I$, the mean in the RF center be $M_{\text{Center}}$, and the intensity of the preceding gray image be $M_{\text{Gray}}$. Then, the relative intensity of this pixel is $I_{\text{Rel}} = (I - M_{\text{Center}})/M_{\text{Gray}}$ and we denote the histogram of relative intensities in an image as $H$. Next, we weigh the relative-intensity histogram ($H(k)$) for each image $k$ with its elicited relative response ($R_V(k)$). This relative response is calculated by $R_V = R_{\text{Total}} - R_C$. Finally, we sum all the response-weighed histograms from all $N$ images and perform a normalization as follows to obtain the RIV:

$$\text{RIV} = \frac{\sum_{k=1}^{N} H(k) * R_V(k)}{\sum_{k=1}^{N} H(k)}.$$

This normalization removes biases due to the specific natural-image data set used.

The rationale behind the RIV function is as follows: Imagine that the local temporal contrast is the only variable controlling the response. In that case, $R_{\text{Total}} = R_C$ and thus, the RIV function, should be close to zero. If instead an intensity variation were more likely to yield a larger response than predicted by the local temporal contrast, then we would get a positive RIV. By the same token, if an intensity variation were more likely to yield a
small response, then the corresponding RIV would be negative. Consequently, the RIV reveals the effects of intensity variations and thus, one can use it to probe visual texture. By comparing the RIV function at small intensity variations with those at large variations, one may be able to tell whether visual texture contributes to the response.

The negative of natural images and their histogram equalization

We generated two new kinds of stimulus from the original natural images to probe the observed response asymmetry of RGCs. The first kind was equivalent to the negative of the image about the mean. To ensure that no intensity was below zero after the inversion of intensities, we compressed their distribution towards the mean by a factor of four. The second new kind of stimulus was a histogram-equalized version of the natural image. This version had a uniform intensity distribution, which we enforced in two steps (Acharya & Ray, 2005):

Step 1: Calculate the intensity distribution of a natural image.

Step 2: Replace the intensity of each point \( I_{\text{Origin}} \) with

\[
I_{\text{New}} = \text{round} \left( \sum_{j=1}^{N_{\text{Origin}}} N_j \times \frac{I_{\text{max}}}{N_{\text{All}}} \right),
\]

where \( I_{\text{max}} \) is the maximum intensity in the image, \( N_{\text{All}} \) is its number of pixels, and \( N_j \) is the number of pixels with the \( j^{\text{th}} \) brightest intensity.

Results

First test of the hypothesis that the RGC’s responses depend on local visual textures of natural scenes: The RIV function

An important property of natural images contributing to an RGC’s responses is the temporal modulation of mean intensity (contrast) in the cell’s RF center. This center-mean-contrast to RGC’s responses can be fitted by a cumulative Gaussian distribution model (Equation 3—Methods). One example of the fit of this model is shown in Figure 2A. The fit is good; for instance, the correlation coefficient between data and fitted model for this cell is 0.86. However, there is still a large amount of additional variability in the cells’ responses that cannot be accounted for by the center-mean-contrast model. Figure 2B shows that the correlation coefficients between data and fitted model are often lower than 0.5. We hypothesize that the response variability
Reducing these correlation coefficients is in part due to the intensity variation in the RF center, i.e., the local visual texture. We will argue in the Discussion and when addressing the results of our fifth test of the hypothesis (the plaid experiments) that the spatial inhomogeneity of the RF cannot explain this response variability.

To test whether the local visual texture had an effect on responses that the local temporal contrast could not account for, we had to discount it first. To do so and to quantify the response contributions of local visual textures, we used the RIV function (Methods). This function quantified how much larger or smaller responses tended to be when an intensity different from the mean appeared in the image. Figure 2C shows the RIV function of a typical RGC. The ordinates of the RIV function are the model-discounted mean response, i.e., the responses minus those predicted by the mean local temporal contrasts (Equation 3). In turn, the abscissas are the relative local intensities. Figure 2C shows that the relative response contributions for intensities around the mean tend to be negative, whereas those for intensities away from the mean tend to be positive. Intensities near the mean are most common when images are homogeneous. Therefore, the negative relative responses around the mean imply that images with weak local textures tend to cause less response than predicted by the local temporal contrast. In contrast, the positive relative responses away from the mean imply that images with strong local texture tend to

Figure 2. RGC’s responses to local visual textures of natural images. (A) Center-mean-contrast-to-response model. The dots represent the original responses of an Off-brisk-transient RGC, while the solid cumulative-Gaussian-distribution curve is the model fit (Equation 3). Although the fit yields a correlation coefficient between data and model of 0.86, there is still a large amount of variability. (B) Distribution of cross-correlation coefficients obtained in fits as in Figure 2A for a sample of 67 RGCS. (C) Response-to-intensity-variation (RIV) function. The abscissas represent the relative difference of pixel intensities to the mean. The ordinates represent the relative mean response contribution. The mean responses were negative for intensities around the mean and positive for those away from the mean. (D) Distribution of RIV at low contrast and high relative mean intensities for a sample of 67 RGCS. \( RIV_H \) was defined as the median RIV at high relative intensities (from 0.8 to 1.0). \( RIV_L \) was defined as the median RIV at low relative intensities (from ~0.1 to 0.1). Across the population, i.e., 94% of cells, \( RIV_H > 0 \) and \( RIV_L < 0 \). In addition, these two variables exhibited a negative correlation.
cause more response than predicted by the local temporal contrast.

Out of sixty-seven RGCs, sixty three (94%) yielded an RIV structure similar to that in Figure 2C, that is, with higher relative responses at pixel intensities far from the mean. Only one RGC yielded an inverted RIV structure, while three others did not show clear RIV structures. To quantify the preponderance of RIVs as in Figure 2C, we obtained for each cell the median RIV values at specific low (Relative Intensity = 0.1–0.1) and high (Relative Intensity = 0.8–1.0) relative intensities (RIVL and RIVH, shown in Figure 2B). We then plotted RIVH as a function of RIVL for all cells in Figure 2D. For almost all cells, the RIVH and RIVL were positive and negative respectively. In other words, the RIV structure observed in Figure 2C was generally applicable to almost all cells. In addition, RIVH and RIVL were negatively correlated (Figure 2D). Therefore, although the gain of the responses varied strongly across different types of RGCs, their dependence on texture did not.

In summary, the RGC’s responses depend on the local visual textures of natural images, with a preference for strong visual textures.

Second test of the hypothesis that the RGC’s responses depend on local visual textures of natural scenes: Asymmetry

If RGC’s responses to natural images depend on their local visual textures, then some of the statistical properties of the textures should be evident in the distribution of

--

Figure 3. Symmetry in local center-mean contrast and asymmetry in local intensity distribution. (A) Distribution of center mean contrasts (Equation 3) of 1,000 natural images for an Off cell. The original curve is shown in solid line while the reflection of the curve about 0 is shown in dashed line. The overlap of these two curves indicates that the distribution is symmetric. (B) Asymmetric distribution of local intensities in natural images. For 4,000 natural images, pixel intensities varied from 0 to 5 \times 10^3 \text{ cd/m}^2. More intensities were below the mean (vertical dashed line) than above it, and those above the mean were often very bright. (C) The joint distribution of mean and standard deviation of intensities in local circular areas from 4,000 natural images. The sizes of the local areas were equal to that of an On cell’s RF that we recorded from. As the local mean intensity increased, so did the standard deviation, yielding a positive correlation coefficient of 0.65.
responses. Some of the most relevant statistical properties of natural images appear in Figure 3. Panel A shows the distribution of center mean contrasts (Equation 3) in the RF center of an Off cell for 1,000 natural images. This distribution is approximately symmetric. In contrast, Figure 3B shows that the distribution of intensities in natural images is asymmetric, with more points being darker than the mean (see also Nuala & David, 2000; Olshausen & Field, 2000; Simoncelli & Olshausen, 2001). Hence, the brightest points in natural images have more contrast compared to the mean than the darkest ones. In addition, the local mean and variance in natural images exhibit positive correlation (Frazor & Geisler, 2006; Simoncelli & Olshausen, 2001). Figure 3C illustrates this correlation in the joint distribution of local intensity mean and standard deviations for a sample of 4,000 of our natural images.

The asymmetric statistical distribution of intensities in natural images leads to our second test of the hypothesis that the RGC’s responses depend on local visual textures. In our main experiment, natural images were alternated with full-field, gray backgrounds. Thus, RGC’s responses could be recorded in two types of transitions. One transition was from gray to natural images and we called it “onset.” The other was from natural images to gray and we called it “offset.” Now, consider, for example, two natural images A and B, such that the onset to A yields the same positive temporal center mean contrast as the offset from B. For a situation like this to happen, Image A must be brighter than the mean in the RF center. In turn, Image B must be darker than the mean. Hence, from the results in Figures 3B and 3C, Image A must have more individual pixels with high contrast than Image B. Therefore, the RIV result (Figure 2) leads to the prediction that an On RGC would respond more strongly to the onset of Image A than to the offset of Image B. Similarly, an Off RGC would respond more strongly to the offset of Image A than the onset of Image B. The predicted asymmetry in the responses to onset and offset are due only to visual texture. This is because the local mean contrasts are similar. Figure 4 shows that the predicted asymmetry holds.

Post-stimulus response histograms of example On and Off cells to the alternation of 1,000 natural images and gray backgrounds appear in Figures 4A and 4D. We found asymmetric responses to the onset and offset transitions for both On and Off RGCs. On cells responded more strongly to the onset of natural images than to the offset (Figure 4A). In turn, Off cells showed similar asymmetries, but with the opposite polarity (Figure 4D). We also calculated the linear RFs of these On and Off cells separately for onsets and offsets of natural images. The linear RFs showed similar response asymmetries as the post-stimulus response histograms. The magnitudes of the linear RFs of On cells were high in response to the onsets of natural images (Figure 4B) but low for offsets (Figure 4C). Instead, Off cells showed a large magnitude for offsets (Figure 4F) but small for onsets (Figure 4E).

The linear RFs for both the onset and offset of natural images were calculated for 96 cells, including 69 Off and 27 On RGCs. To visualize the response asymmetries for all cells, we plotted in Figure 4G a point per RGC, with the peak magnitude of its onset linear RF as abscissa and that for the offset as ordinate. We included in the plot a diagonal line to indicate equal responses to onsets and offsets. If RGCs responded symmetrically to both onsets and offsets, the points should have had a symmetric distribution about the diagonal line. However, 87 cells (91%) were below the equality line. Sixty three of 69 Off cells had larger peak magnitude in their offset linear RFs than in their onset counterparts. In turn, 24 of 27 On cells had stronger magnitude in their onset linear RFs. This Onset–Offset asymmetry held generally for all RGCs recorded, regardless of cell type.

Hence, the hypothesis that the RGC’s responses depended on visual texture passed the second test. RGCs showed asymmetric responses to the onset and offset of natural images across the population, with the Off cells’ polarities being opposite to those of On cells.

Third test of the hypothesis that the RGC’s responses depend on local visual textures of natural scenes: Image negatives

The prediction of this Onset–Offset-asymmetry was based on the special statistical properties of the visual texture in natural images (Figures 3B and 3C). The direction of the asymmetry depended on the distribution of intensities having positive skewness (Figure 3B). Such skewness led to the prediction that On cells prefer the onset of natural images, while Off cells prefer their offset. However, if the skewness were negative, then the prediction would be that On cells would prefer offsets, while Off cells would prefer onsets. To test this prediction, we used the negative of natural images to stimulate RGCs. An original natural image (positive) and its negative appear in the top panels of Figures 5A and 5B respectively. Because the negatives of natural images had an inverted asymmetric intensity distribution about the mean intensity (compare the middle panels of Figures 5A and 5B), our hypothesis predicted the inversion of the asymmetry of the RGC’s responses. That was exactly what we observed. An example from an Off cell is shown in the third row of Figures 5A and 5B. As expected, this cell showed stronger responses to the offsets of original natural images than to onsets. However, when stimulated with the negatives of natural images, this asymmetry was inverted. This inversion occurred across our sample of 27 cells (Figure 5E).
Fourth test of the hypothesis that the RGC’s responses depend on local visual textures of natural scenes: Histogram equalization

The test of the negative images inverted the direction of the Onset–Offset asymmetry. Can one come up with an image transformation that according to our hypothesis would not only invert the asymmetry, but eliminate it altogether? The answer is positive: one can use histogram-equalized natural images to stimulate RGCs (Methods). Two of these images, one normal and one histogram equalized, are shown in the top panels of Figures 5C and 5D. Because the histogram equalization eliminated the skewness of the intensity distribution (the middle panels in Figures 5C and 5D), our hypothesis predicted the elimination of the asymmetry of the RGC’s responses. Again, that was exactly what happened. For example, consider the Off cell, shown in the bottom panels of Figures 5C and 5D. This cell did not show any significant asymmetries in their linear RF for either positive (Figure 5C) or negative (Figure 5D) histogram-equalized natural images. The same holds for a sample of 29 cells (Figure 5F).

Fifth test of the hypothesis that the RGC’s responses depend on local visual textures: Artificial plaids

We found the dependence of RGC responses on visual texture in the RF center using natural images. The reason why they helped us reach this conclusion is that natural images are rich in texture. However, one can also build artificial textures that can test the hypothesis. For instance,
we built plaid images comprising small square elements of only two intensities with the opposite contrast. As in the previous experiments, we alternated these plaids with a gray full-field background of the same mean intensity. Hence, the temporal local mean contrast of image transitions was zero. We then measured the response of the RGC as a function of the strength (i.e., spatial contrast) of the local visual texture in the plaid. Two example plaids with weak and strong visual textures are respectively shown in Figures 6A and 6B. Based on the RIV result, our hypothesis led to the prediction that the response should increase as a function of the strength of the local visual texture. Figure 6D shows that this prediction holds for the responses of an example RGC for plaid elements of 150 μm².

As predicted by our hypothesis and the RIV curve, the RGCs’ responses to the artificial plaids increased with the strength of the local visual texture although the temporal mean contrast was zero. In addition, when we used the negative image of the plaid, the response was roughly equal to the response to the positive image. This negative-image control ruled out the possibility that the increase in responses in Figure 6D were due to inhomogeneities of the RGC’s RF shape. For instance, consider an On cell. Light squares in the middle of its Gaussian-looking RF would count more than dark squares in its periphery. However, with the negative image, we would now have dark squares in the middle of the RF. Hence, if the responses in Figure 6D were due to RF spatial inhomogeneities, the positive and negative images would rarely give the same responses as we observed.

One of the advantages of using an artificial image is the possibility of manipulating a few of its parameters. For the plaid, the other main parameter besides texture strength is the square-element size. When we varied this parameter, the response dependence on texture strength did not
change for square elements of 100 x 100, 150 x 150, 200 x 200, and 300 x 300 μm^2. However, when we reduced the size of the squares elements of the plaid to 50 x 50 μm^2 (Figure 6C), the RGC showed no response to this plaid.

Discussion

Dependence of RGC’s responses on texture of natural scenes

An important property of natural images contributing to an RGC’s responses is the temporal mean contrast in the cell’s RF center. The center-mean-contrast-to-response curve can be fitted by a cumulative-Gaussian-distribution model (Figures 2A and 2B). Although the model largely accounts for the RGC’s responses, it cannot explain their large variability. There are many candidates to explain this variability, such as the contribution from visual texture of natural stimuli, the response of the RGC’s RF surround, or simply neural noise. In this study, we obtained evidence in favor of the hypothesis that the visual texture of natural scenes in the RF center contributes to the response variability (Figures 2, 4–6). This contribution is significant when compared to the noise of the neurons. For instance, with the plaid experiment, we could estimate how much of the variance one could explain with the visual-texture responses. To make this estimation, we first computed the total variance (VT) in data like those in Figure 6D. Then, we performed linear regression on the data and measured the residual variance (VR) from the best linear fit. The percentage of variance explained by the texture was then (VT - VR)/VT. For the data in Figure 6, about 75% of the variance was explained by the dependence of the responses on visual texture. For a sample of 13 RGCs, the percentage of the variance explained by visual-texture dependence was 50% ± 20%.

To quantify the visual texture of natural images, we used the histogram of intensity distribution (Figure 1), rather than variability estimators, such as standard deviation or entropy. Standard deviation is not a good parameter for natural stimuli, because it cannot differentiate the points with intensities above the mean and those with intensities below it. In other words, the standard deviation cannot capture the asymmetry of the intensity distribution in natural stimuli. Similarly, entropy...
does not capture this asymmetry. In contrast, the full histogram can quantify the asymmetry with a richer representation of the variation of intensities.

With the histogram representation of the intensity variation, we developed the RIV method and used it to show that RGCs prefer strong visual textures to relatively homogenous natural stimuli (Figure 2C). The RIV method has four important limitations: First, it neglects the response contribution from the RF surround. Second, the method assumes that an RGC has the same dependence on visual texture for different temporal center mean contrasts. Third, the method is not sensitive to the effect of different positions in the RGC’s RF. Fourth, positive correlation between neighboring-pixel intensities in natural stimuli (Olshausen & Field, 2000) might influence the RIV results. Nevertheless, none of these limitations detract from our conclusion that the texture in the RF center modulates the responses of RGCs.

Asymmetry of RGCs’ responses due to dependence on visual texture

In our experiments, the onset and offset of natural stimuli did not elicit equal responses from RGCs. On cells responded more to the onset of natural stimuli than to the offset, whereas Off cells responded in the opposite way (Figure 4). Although we predicted these responses from the statistical properties of visual textures (Figure 3), we now consider two alternate explanations.

Several known mechanisms of RGCs may possibly explain these asymmetric responses. One is that RGCs can adapt to temporal and spatial contrasts of light stimuli (Baccus & Meister, 2002; Chander & Chichilnisky, 2001; Kim & Rieke, 2001; Smirnakis, Berry, Warland, Bialek, & Meister, 1997). In our experiments, there was a difference in the initial states of the onset and offset of natural images. The onset started with full-field gray background that included no contrasts, whereas the offset started with natural images that included luminance contrasts. If the initial contrast states played an adaptive role in the RGC’s responses to the alternation of natural images with the gray background, then these states could possibly generate the observed asymmetry. Our results using the negative natural images, however, ruled out this possibility (Figures 5A, 5B, and 5E). With the negative natural images, we did not alter the initial contrast states of the onset and offset. Yet, we observed the opposite asymmetries. Thus, the response asymmetries did not result from the difference of initial states of contrast adaptation in the onset and offset conditions. This conclusion strongly suggests that the RGC’s response dependence on texture of natural images is not due to sensory adaptation to their statistical properties.

A second possible explanation that we considered for the Onset–Offset asymmetry was that the distribution of temporal center mean contrasts in natural stimuli could be asymmetric. Considering this explanation was important, because an important stimulus variable controlling the RGC’s responses is the center mean contrast (Figure 2A). Moreover, an asymmetric center-mean-contrast distribution seemed plausible at first, because the distribution of intensities has positive skewness (Figure 3B). However, the distribution of center mean contrasts turned out to be approximately symmetric (Figure 3A). The symmetry of this distribution of center mean contrasts may be explained by the Central Limit Theorem (Rice, 1995). Each of the center mean contrasts is determined by a large number of points in the RF center. Hence, although the distribution of intensities has positive skewness, the distribution of intensity means tends to be symmetric (approximately Gaussian).

We finally demonstrated that the asymmetric responses of RGCs to the onset and offset of natural images were caused by the asymmetric distribution of intensities of natural images (Figure 3B). This demonstration was based on two experiments. In the first experiment, stimulating RGCs with the negatives of natural images caused the response asymmetries to invert (Figures 5A and 5B). In the second experiment, performing a histogram equalization on natural images eliminated the response asymmetries (Figures 5C and 5D).

That RGCs respond to asymmetries in visual textures may be significant for perception. Psychophysically, humans are sensitive to the asymmetries of intensity distribution in natural stimuli (Chubb, Landy, & Econopouly, 2004; Motoyoshi, Nishida, Sharan, & Adelson, 2007; Ratliff, Borghuis, Kao, Sterling, & Balasubramanian, 2010; Tkacik, Prentice, Victor, & Balasubramanian, 2010). For example, humans may use the skewness of the reflected light to estimate surface qualities of objects (Motoyoshi et al., 2007) and may use the “blackshot” mechanism to sense the prevalence of the darkest intensities (Chubb et al., 2004). Our studies added to these perceptual experiments by showing that in early visual processing, RGCs respond to the asymmetric intensity distribution in natural stimuli.

Model for RGC texture sensitivity

How can one explain the sensitivity of RGCs to visual texture in the RF center? One may seek explanations in either biophysical or computational models of RGC function. We begin with the latter kind of model and then comment on the former.

Arguably, the most successful computational model for many RGCs is the linear–nonlinear (LN) model (Carandini et al., 2005; Chichilnisky, 2001; Shapley & Victor, 1978; Willmore & Smyth, 2003). This model consists of a spatiotemporal linear filter followed by a static, pointwise nonlinearity. The LN model has been successful in accounting for responses of visual cells to artificial stimuli, including Gaussian white noise (Chichilnisky, 2001; Willmore & Smyth, 2003) and other orthogonal
stimuli (Reid & Alonso, 1995; Ringach, Sapiro, & Shapley, 1997). However, natural images are neither white nor orthogonal. They have considerable spatial (and for movies, temporal) correlations (Balboa & Grzywacz, 2003; Field, 1987; Ruderman & Bialek, 1994). Therefore, using natural images to test the LN model is harder than with white or orthogonal stimuli. Perhaps this difficulty is the reason why only one study has so far shown evidence that this model may account for responses to natural stimuli (Mante, Bonin, & Carandini, 2008).

We now argue that a LN model is inconsistent with the data in our paper. In this model, the linear filter calculates a weighted mean of the image. If two images had the same weighted mean intensity in the RF center, then they would elicit the same response regardless of the nonlinearity. Consequently, two natural images of similar temporal local mean contrast should yield similar responses. However, as shown by the RIV result (Figure 2C), differences in visual texture can break the response similarity. Therefore, visual-texture dependence argues against a LN model for the responses to natural images.

A counterargument is that one cannot be sure that two images with similar temporal mean contrasts in the RF have similar weighted mean intensities. This is because, as explained in Results, RF profiles are inhomogeneous. However, we already argued when discussing the data in Figure 6 that the methods of the plaid experiment are immune against this criticism. We now extend those arguments. Let us say that a Plaid A generates a response larger than a Plaid B. Then, the output of the linear filter of the LN model must be different for the two plaids. Hence, if we stimulated the RGC with the negative images of the plaids, the order of the linear-filter outputs would be inverted, leading to an inversion of the order of the responses. Consequently, if responses increase with the contrast of a plaid (Figure 6D), then according to the LN model, they cannot increase if we use the negative of the plaid. This prediction violates the result of Figure 6D, which shows no effect of using the negative of plaids.

We thus sought an alternate computational model that could account for the visual-texture dependence of RGCs. The model that we propose has multiple spatially distributed nonlinear subunits integrated linearly. Thus, this model has a NL structure. A linear mechanism could precede the first nonlinearity, allowing us to account for the subunit responses with a LN model. Similarly, a nonlinearity (e.g., from a spiking mechanism) could follow the linear integration in our model. Therefore, in its more general form, the model would have a LNLN structure, of which the central NL mechanism is the part crucial for texture dependence.

To explain how this model works, let us consider a simple example of both an RGC that is receiving inputs from only two nonlinear subunits (Figures 7A and 7B) and an image with only two pixels. (Although we explain this model by using only two subunits and two-pixel images, it generalizes for more subunits and more pixels.) Each subunit responds to only one pixel of the image with a “threshold-like” nonlinear function (Figure 7C). The output of this ganglion cell is equal to the sum of the outputs of these two subunits. The nonlinearity of the threshold-like function, the response of the subunit to the blue pixel is slightly smaller than that to the green pixel, and the response to the red pixel is much larger than that to the yellow pixel. As a result, the response sum for the blue and red pixels (A) is higher than the sum for the green and yellow—B. Because of the nonlinearity of the threshold-like function, the response of the subunit to the blue pixel is slightly smaller than that to the green pixel, and the response to the red pixel is much larger than that to the yellow pixel. As a result, the response sum for the blue and red pixels (A) is higher than the sum for the green and yellow pixels (B). Thus, the RGC shows larger responses to strong visual textures than those to weak visual textures. (D) If the first stage of the model were linear, then we would not observe the texture dependence of the RGC’s responses.

![Figure 7](Image)
situation is that of an image with the same mean intensity as the first but with weak visual texture (Figure 7B). This image is represented by two pixels intensities close to the mean, shown as the green and yellow dots. The mean intensity of the two pixels is equal in both situations (Figure 7C), but the responses are not equal. Because of the concave “threshold-like” nonlinearity of the subunits (Figure 7C), the response to the blue pixel is slightly smaller than that to the green pixel. In turn, the response to the red pixel is much larger than that to the yellow pixel. As a result, the sum of the responses to the blue and red pixels (Figure 7A) is higher than the sum for the green and yellow pixels (Figure 7B). Thus, the RGC shows a stronger response to the image with strong visual texture than to that in the image with weak visual texture. The nonlinearity preceding the linear summation is crucial for this texture dependence. Figure 7D shows that if the first stage were linear as in the LN model rather than nonlinear, an image with strong visual texture would tend to yield a similar response as an image with weak visual texture. Consequently, we would not get texture dependence as observed in RGCs.

The model not only accounts for the dependence on visual texture. The threshold-like nonlinearity of the subunits ensures that the model also accounts for the sigmoidal response dependence on the mean contrast. This dependence was observed here (Figure 2A) and in previous studies (Carandini et al., 2005; Chichilnisky, 2001; Shapley & Victor, 1978; Willmore & Smyth, 2003).

What are possible biophysical implementations of the nonlinear subunits in the new model? Without detailed intracellular studies, we cannot be certain of what the cellular underpinnings of the NL model possibly mediating visual-texture dependence may be. Their subunit nonlinearities could come, for example, from the synaptic outputs of bipolar cells. In this case, the segregation between subunits could be morphological (e.g., different bipolar cells). Alternatively, the subunit nonlinearities could be due to voltage-dependent processes in the dendritic trees of RGCs (Burkhardt & Fahey, 1998; Dacey et al., 2000; Demb, Zaghoul, Haarsma, & Sterling, 2001; Euler & Masland, 2000; Koch, Poggio, & Torre, 1983; Wu, Gao, & Maple, 2000). Hence, the segregation between subunits would be functional (e.g., electrotonic isolation of dendrites). Although we do not have enough data to settle the nonlinear mechanism, we can comment on the spatial properties of the subunits. These properties are revealed by the dependence of the responses on the size of the square elements of the plaids. The responses vanish when this size falls to around 50 × 50 μm² but not at around 100 × 100 μm² (Figure 6E). Therefore, the spatial integration of individual subunits should be between 50 and 100 μm. Such integration may be consistent with bipolar-cell RFs (Dacheux & Miller, 1981; Ghosh, Bujan, Haverkamp, Feigenspan, & Wassle, 2004; Jeon & Masland, 1995; MacNeil, Heussy, Dacheux, Raviola, & Masland, 2004; Mills & Massey, 1992).

Nonlinear subunits in the RF center of RGCs have been proposed by other investigators (Enroth-Cugell & Robson, 1966; Hochstein & Shapley, 1976; Soodak, Shapley, & Kaplan, 1991; Victor, 1987). The proposals covered a variety of species, including rabbit, the animal in our study (Barlow & Levick, 1965). However, the experiments leading to the proposed subunits used artificial stimuli, rather than natural ones. Our studies show that even to explain RGCs’ responses to natural images, one should not ignore the possible role of nonlinear subunits.

**The importance of studying visual neurons with natural images**

We decided to study the dependence on visual texture after observing the variability of the responses of RGCs to natural images (Figure 2A). That such stimuli motivated us to work on an important property of RGCs seemed already to add to the strong arguments in favor of using natural images in visual research. However, the use of these kinds of images incurs some complications. The most important complication is that because one does not design natural images, one may not be able to control all aspects of the stimulation. This “lack” of control does not mean that we do not know the intensity and color of every pixel. These variables can be well calibrated. However, natural images can lead to larger response variability than artificial ones, since the exact geometry of the inputs within and outside the RF varies across stimuli. Furthermore, some studies have argued that one can predict the responses to natural images from models based on artificial-stimulation data (Mante et al., 2008).

Nevertheless, we argue that the lessons learned by using natural images outweigh these difficulties. Responses of cells in the visual system may depend on the statistics of the environment in ways that are hard to glean from results of experiments using artificial stimuli. For instance, in our data, the onset–offset asymmetry depends on an interaction between peculiar statistics of natural images and nonlinear mechanisms of RGCs. One can get textures with images like gratings, but these are symmetric and have special properties due to their periodicity. Random-dot patterns also have textures, but cells often respond poorly to them.

In addition, although many studies have described some special statistics of natural images (Simoncelli, 2003; Simoncelli & Olshausen, 2001), we cannot yet be certain that we can capture all the relevant statistics artificially. Early studies, for example, emphasized the self-similarity apparent in the power spectrum of natural images (Field, 1987). In contrast, later studies both challenged self-similarity (Balboa, Tyler, & Grzywacz, 2001) and pointed out that power spectrum lacks phase information (Thomson, 2001) and thus, is insensitive to contours. Other studies yet argued that power spectra contained only second-order statistics and thus, failed to capture the skewness (Ratliff
et al., 2010) and kurtosis (Field, 1994) present in natural-image statistics. Finally, natural images tend to be dominated by low contrasts (Balboa & Grzywacz, 2000, 2003; Ruderman & Bialek, 1994; Zhu & Mumford, 1997). Consequently, what one obtains with these images may be different from the majority of the data with artificial stimuli, which tend to use high contrasts to increase responses.

Conclusion

In conclusion, although the main variable of natural images that RGCs encode may be the mean temporal contrast in the RF center, they may also respond to the statistical properties of natural visual textures.

Acknowledgments

We thank Eun-Jin Lee, Junkwan Lee, Joaquín Rapela, Mónica Padilla, Susmita Chatterjee, and Jeff Wurfel for their helpful discussions of the material, and Denise Steiner for administrative assistance. This work was supported by National Eye Institute grants EY08921, EY11170, and EY016093, and a National Science Foundation grant EEC-0310723 to N.M.G.

Commercial relationships: none.
Corresponding author: Xiwu Cao.
Email: caoxiwu@gmail.com
Address: 1042 Downey Way, DRB 140, Los Angeles, CA 90089, USA.

References


