Predictive Coding: A Fresh View of Inhibition in the Retina

M. V. Srinivasan; S. B. Laughlin; A. Dubs

Proceedings of the Royal Society of London. Series B, Biological Sciences, Vol. 216, No. 1205. (Nov. 22, 1982), pp. 427-459.

Stable URL:
http://links.jstor.org/sici?sici=0080-4649%2819821122%29216%3A1205%3C427%3APCAFVO%3E2.0.CO%3B2-P

Proceedings of the Royal Society of London. Series B, Biological Sciences is currently published by The Royal Society.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/rsl.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org/
Mon Jul 31 16:35:24 2006

Proc. R. Soc. Lond. B **216**, 427–459 (1982)
Printed in Great Britain

Predictive coding: a fresh view of inhibition in the retina

By M. V. Srinivasan[1, 2]†, S. B. Laughlin[1] and A. Dubs[1]

Departments of Neurobiology[1] and Applied Mathematics[2], Australian National University, P.O. Box 475, *Canberra City, A.C.T.*, 2601, *Australia*

(*Communicated by G. A. Horridge, F.R.S. – Received* 26 *January* 1982 – *Revised* 21 *June* 1982)

Interneurons exhibiting centre–surround antagonism within their receptive fields are commonly found in peripheral visual pathways. We propose that this organization enables the visual system to encode spatial detail in a manner that minimizes the deleterious effects of intrinsic noise, by exploiting the spatial correlation that exists within natural scenes.

The antagonistic surround takes a weighted mean of the signals in neighbouring receptors to generate a statistical prediction of the signal at the centre. The predicted value is subtracted from the actual centre signal, thus minimizing the range of outputs transmitted by the centre. In this way the entire dynamic range of the interneuron can be devoted to encoding a small range of intensities, thus rendering fine detail detectable against intrinsic noise injected at later stages in processing. This predictive encoding scheme also reduces spatial redundancy, thereby enabling the array of interneurons to transmit a larger number of distinguishable images, taking into account the expected structure of the visual world.

The profile of the required inhibitory field is derived from statistical estimation theory. This profile depends strongly upon the signal:noise ratio and weakly upon the extent of lateral spatial correlation. The receptive fields that are quantitatively predicted by the theory resemble those of *X*-type retinal ganglion cells and show that the inhibitory surround should become weaker and more diffuse at low intensities. The latter property is unequivocally demonstrated in the first-order interneurons of the fly's compound eye. The theory is extended to the time domain to account for the phasic responses of fly interneurons.

These comparisons suggest that, in the early stages of processing, the visual system is concerned primarily with coding the visual image to protect against subsequent intrinsic noise, rather than with reconstructing the scene or extracting specific features from it. The treatment emphasizes that a neuron's dynamic range should be matched to both its receptive field and the statistical properties of the visual pattern expected within this field. Finally, the analysis is synthetic because it is an extension of the background suppression hypothesis (Barlow & Levick 1976), satisfies the redundancy reduction hypothesis (Barlow 1961*a*, *b*) and is equivalent to deblurring under certain conditions (Ratliff 1965).

† Present address: Abteilung Neurobiologie, Zoologisches Institut der Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland.

Introduction

Lateral and temporal inhibition are frequently encountered in the preliminary stages of sensory processes, and are commonly observed in visual systems. For example, the bipolar and ganglion cells of vertebrate retinas often generate transient responses and have receptive fields with antagonistic centre and surround organization. Similar patterns of antagonism are found in the first order interneurons of *Limulus* and insect compound eyes (Barlow *et al.* 1957; Werblin 1971; Hartline & Ratliff 1972; Laughlin 1981*a*).

Several functions have been assigned to the inhibition observed in retinal neurons. The most general proposition is that inhibition removes redundancy from the input so as to allow retinal neurons to encode the incoming information efficiently (Barlow 1961*a*, *b*, *c*). By relating the intensity falling on the centre of the receptive field to surrounding levels of illumination, lateral inhibition reduces the range of intensity that neurons need encode without necessarily rejecting information (Barlow 1961*b*). This role of reducing the response range of the centre has also been interpreted as a zero offset, cancelling out the steady (d.c.) bias of the signal so as to keep a neuron's response range centred upon the local mean intensity. Thus lateral inhibition is one of the mechanisms that allow retinal neurons to operate over a wide range of intensities with an appropriate sensitivity (Barlow & Levick 1976; Laughlin & Hardie 1978). Alternatively, lateral inhibition has been interpreted as a filter for enhancing edges (Ratliff 1965), either to deblur the retinal image (see, for example, Marcelja 1979) or to encode a wide range of spatial detail in a form that might be appropriate for analysis in higher centres (Marr & Hildreth 1980). There are difficulties associated with using these hypotheses to establish with certainty the roles that known patterns of inhibition play in retinal processing. As we discuss later, the quantitative predictions of the deblurring hypothesis do not accord with much of the data, the d.c. offset and redundancy reduction hypotheses are qualitative, while the edge enhancement hypothesis depends upon *a priori* assumptions about the relative values of certain signal components to processing at higher levels.

To further our understanding of retinal inhibition's repertoire, and to help define the function of inhibition in known networks, we have developed a new analytical approach, predictive coding. This simple procedure, initially developed for transmission of video data (Oliver 1952; Harrison 1952), reduces the signal amplitude by removing predictable, and hence redundant, components. The reduction in input enables afferent neurons to operate with a higher sensitivity so that the information contained within small signals is less likely to be lost amid the intrinsic noise injected at higher levels. Predictive coding is a unifying concept, giving a clearer perspective of inhibition's full repertoire, because it is a limiting case of both the d.c. offset and deblurring hypotheses, and it also removes redundancy. As one might expect of the simple patterns of inhibition seen in the retina, redundancy is not completely eliminated.

The predictive coding hypothesis has the advantages of being quantative and justifiable without recourse to an evaluation of the importance that subsequent processing gives to various components of the incoming signal. Coding is directed

towards using neurons efficiently to encode *all* the information in the incoming stream of data. Thus the theory can be tested without recourse to *a priori* arguments. Our experiments show that the second order neurons of the fly's compound eye have the patterns of intensity-dependent spatial and temporal inhibition required for predictive coding. As vertebrate retinal neurons show similar properties (Laughlin 1981*a*), we suggest that the simple procedure of predictive coding as used in the fly's eye is of wider relevance. Thus, our findings add weight to earlier suggestions that the 'neat packaging' (Barlow 1961*b*) of information is of fundamental importance to early visual processing.

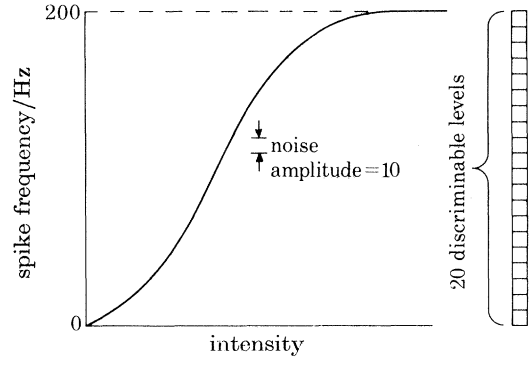

FIGURE 1. Intensity–response function of a typical visual interneuron, schematically shown here to illustrate the effect of intrinsic noise on intensity discrimination. If the response range of this spiking interneuron is restricted to, say 0–200 Hz, and if the amplitude of intrinsic noise corresponds to a fluctuation in spike frequency of, say, 10 Hz, then the neuron can encode $200 \div 10 = 20$ distinguishable intensity levels.

THEORY

The starting point of this coding theory is that intrinsic noise, generated within the nervous system central to the action of inhibition, places a limit upon each neuron's ability to encode incoming signals. As pointed out by Barlow & Levick (1976) a neuron can only produce a limited number of discriminable response levels because biophysical factors such as reversal potentials and maximum spike rates restrict the available response range, and noise limits the number of response levels that can be distinguished within this range. Even if the incoming signal is noise-free, the intrinsic noise divides the response range into a finite number of response levels, as illustrated in figure 1. We propose that lateral inhibition minimizes the range of inputs presented to a neuron, by exploiting the correlation that exists between neighbouring points in a visual scene. It uses the intensity values in the surrounding regions to generate a statistical estimate of the intensity expected at a particular point. By subtracting this best estimate from the signal actually entering at this point, the amplitude of the transmitted signal is minimized. This reduction allows small fluctuations to be monitored with the high sensitivity necessary for them to be resolved against intrinsic noise, yet allows the neurons to operate over a wider range of receptor input levels. This design consideration

is of some importance because visual systems generally have to operate over a wide range of intensities.

The theory is developed within the context of an array of visual interneurons, each of which has an excitatory centre and an inhibitory surround (figure 2). It is first established that spatial correlation exists within the scenes presented to this array. Based upon this correlation, the signal at the centre of the receptive field is estimated from the surround, according to statistical principles. The estimate

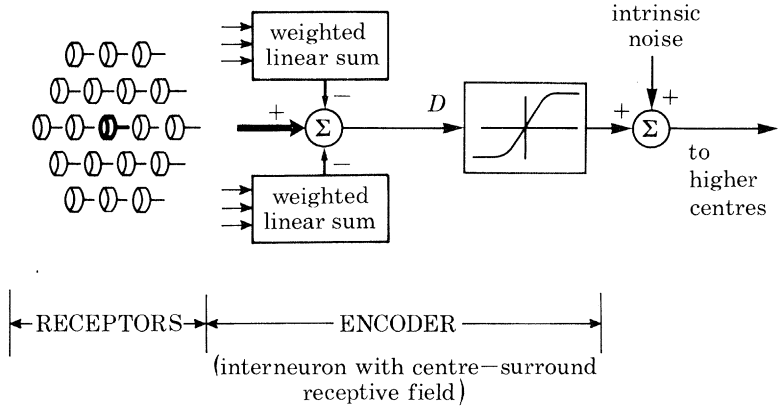

FIGURE 2. Schematic diagram of receptive-field organization, illustrating the hypothesized encoding scheme. The signal produced by the central receptor (shown by heavy lines) is compared with a statistical prediction of this signal based on a weighted linear sum of signals from receptors in the vicinity. The interneuron encodes the difference, D, betwen the actual and the predicted signals, and transmits it to the higher centres.

is derived by summing together weighted contributions from neighbouring points. This weighting function defines the inhibitory surround required to minimize the error between the estimated value of the signal and its actual value, so minimizing the range of signals each interneuron need code. The shape of the weighting function depends strongly upon the signal:noise ratio of the input, and hence intensity, but is only weakly influenced by the degree of spatial correlation in the scene. It is demonstrated that this coding strategy, which minimizes dynamic range, also removes spatial redundancy from the output of the interneuron array, so satisfying the redundancy reduction hypothesis (Attneave 1954; Barlow 1961*a–c*). Finally, the concept of predictive coding is extended to the time domain. Here we exploit the fact that the signal evoked in a receptor by a moving visual scene is temporally correlated because (*a*) the scene is spatially correlated and (*b*) the photoreceptor's impulse–response function has a finite duration. Thus, one can derive the pattern of self-inhibition that predicts the present value of the receptor signal based on its past values. In this case, predictive coding removes temporal redundancy from the interneuron's output.

Scenes contain spatial correlation

In a purely random visual scene, the variation of intensity from one point to the next would be unpredictable. However, real visual scenes are usually composed

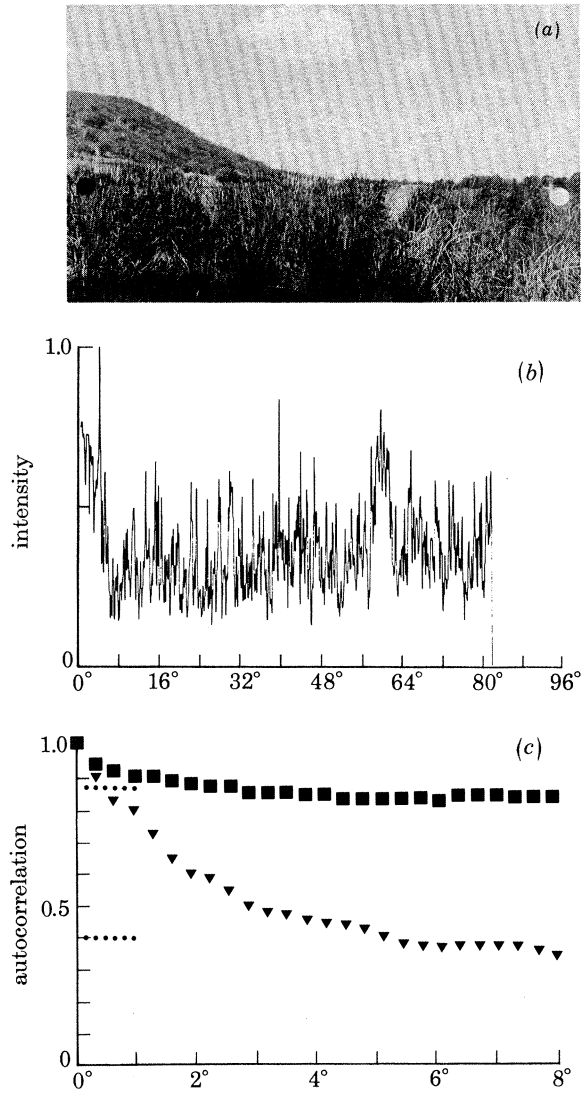

FIGURE 3. The autocorrelation function of visual scenes. (*a*) The scene, a bed of reeds. (*b*) Spatial intensity profile of scene, measured along a line joining the centres of the marker discs. Intensity is plotted on a linear scale with the maximum value normalized to unity. (*c*) Squares: autocorrelation function of intensity profile shown in (*b*). Triangles: autocorrelation function of an intensity profile obtained from a horizontal scan about two disc diameters below that of (*b*). The autocorrelation function $R(x)$ is computed according to

$$R(x) = \frac{1}{L-x}\int_0^{L-x} I(\xi)\, I(\xi+x)\,\mathrm{d}\xi \Big/ \frac{1}{L}\int_0^{L} [I(\xi)]^2\,\mathrm{d}\xi,$$

where $I(\xi)$ is the function describing the spatial intensity profile, and L is the total length of this profile (degrees). Both x and ξ are measured in degrees. The denominator normalizes the autocorrelation function to unity, i.e. $R(0) = 1$. The dotted lines indicate, for each intensity profile, the ratio between the the square of the mean value and the mean square value. As $x \to \infty$ the autocorrelation function should, in theory, asymptotically approach this ratio, provided that the statistical properties of the intensity fluctuations remain constant within the scene. This condition was approximately fulfilled by most of the scenes that were analysed.

of finite areas within each of which the reflectance does not change dramatically. As a result, most visual scenes are not entirely random and intensities are spatially correlated over short distances (Kretzmer 1952). Even a completely random scene (with zero spatial correlation) would give rise to a spatially correlated retinal image, because the optical limitations found in any eye (e.g. diffraction and other

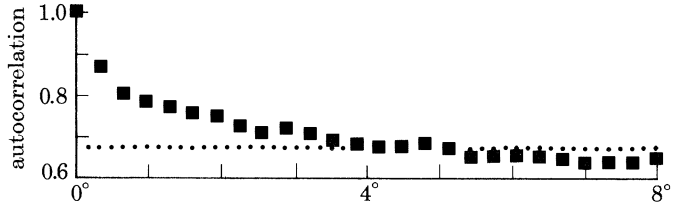

FIGURE 4. Spatial autocorrelation function of intensity profile obtained by scanning another visual scene, a bush at a distance of approximately 5 m (photograph not shown). (Dotted line as in figure 3.)

aberrations) produce on the receptor array a smoothed, blurred version of the original scene. Thus, all retinal images contain spatial correlation. The correlation distance will vary from scene to scene, depending upon the sizes, textures and distances of objects in the scene, and with the quality of the optics. As we will demonstrate that the receptive fields required for predictive coding are rather insensitive to the precise extent of spatial correlation, we present here only two examples to illustrate the range of correlation distances that we have found in natural scenes.

Figures 3 and 4 show spatial autocorrelation functions measured for two different scenes. Figure 3*a* shows a photograph of one of the scenes, a patch of tall reeds. This scene was chosen for illustration because, when scanned horizontally, it is the most random that we have measured so far (i.e. it has the smallest correlation distance). Figure 3*b* shows a spatial intensity profile measured by scanning the scene with a photodetector with a visual field of approximately 0.1° (Laughlin 1981*b*), along a horizontal line joining the centres of the marker discs. Figure 3*c* (squares) depicts the spatial autocorrelation function computed for this profile, as described in the figure legend. It illustrates that intensities are correlated over an angular distance of about 2°. The triangles in this figure depict the autocorrelation function obtained from a horizontal scan two disc diameters below that of figure 3*b*. The spatial correlation now extends to about 4°, presumably because this scan includes more of the grass blades in the foreground, which subtend larger visual angles at the detector. Figure 4 shows the spatial autocorrelation function of a large bush at a distance of 5 m (photograph not shown). It indicates a spatial correlation distance of about 3° for this scene. Scenes depicting human figures are also spatially correlated over substantial angular distances (see, for example, Kretzmer 1952; Rosenfeld & Kak 1976; Gonzalez & Wintz 1977).

Predicting the intensity at a point

Given that intensities are spatially correlated, it is possible to predict the signal expected at a particular receptor, from the signals in the receptors surrounding it. The theory of signal processing offers a well established statistical procedure for making such predictions. It is called 'linear mean-squared estimation' (Papoulis 1965). If a spatial signal is sampled at m different locations 1, 2, ..., m, yielding

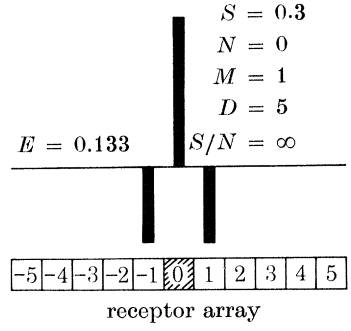

FIGURE 5. Weighting coefficients, calculated from theory, specifying the receptive field of an encoding interneuron that receives input from an array of 11 receptors. The interneuron receives an excitatory input from the central receptor (0), and compares it with a statistical estimate of this signal derived from a suitably weighted linear combination of signals from the ten other receptors. The strengths of the inhibitory coefficients are plotted relative to that of the excitatory coefficient from the central receptor, which is normalized to unity. (The receptive-field profile of the interneuron can be derived from this array of weighting coefficients by taking into account the finite acceptance angle of each receptor.) In this and the following figures M denotes the mean intensity of the visual scene (always normalized to 1), S denotes the standard deviation of the fluctuations of intensity about the mean (i.e. the average contrast of the scene), N denotes the standard deviation of the receptor noise (expressed in terms of equivalent contrast) and D denotes the space constant of the spatial autocorrelation function of the visual scene (expressed in receptor widths). In this example the receptor signal:noise ratio (S/N) is infinity. The figure also shows the standard deviation, E, in intensity units, of the expected error between the predicted and actual values of receptor signal, calculated according to equation (1*c*). Details of the calculations for this and other figures are described in Appendix 1.

samples $x_1, x_2, \ldots, x_m$, respectively, then the statistically best linear prediction, $\hat{x}_0$ of the signal at a location 0 is given by a weighted linear combination of the m samples:

$$\hat{x}_0 = \sum_{i=1}^{m} h_i x_i, \tag{1a}$$

where the h_i are solutions of the system of linear simultaneous equations

$$\left.\begin{aligned} R_{1,1}h_1 + R_{1,2}h_2 + \ldots + R_{1,m}h_m &= R_{0,1} \\ R_{2,1}h_1 + R_{2,2}h_2 + \ldots + R_{2,m}h_m &= R_{0,2} \\ \cdot \quad\quad \cdot \quad\quad \cdot \quad\quad \cdot \quad\quad \cdot \\ R_{m,1}h_1 + R_{m,2}h_2 + \ldots + R_{m,m}h_m &= R_{0,m} \end{aligned}\right\} \tag{1b}$$

and the $R_{i,j}$ are the spatial autocorrelation coefficients between samples taken at locations i and j. That is, $R_{i,j} = \overline{x_i x_j}$. The prediction $\hat{x}_0$ so obtained is 'best' in the sense that it minimizes the mean square error between the predicted value and the true value of the signal at location 0. The expected standard deviation (E) of this error is given by

$$E = \sqrt{\left(R_{0,0} - \sum_{i=1}^{m} h_i R_{0,i}\right)}. \tag{1c}$$

Figure 5 shows the result of using this theory to derive the receptive-field profile of an interneuron that receives input from a one-dimensional array of receptors, spaced one unit of distance apart. In this calculation we have assumed that the retina views a scene in which the distribution of intensities has a mean value of one unit of intensity, a standard deviation about the mean of 0.3 intensity units (i.e. an average contrast of 30 %) and a spatial autocorrelation that is exponential with a space constant of five distance units (i.e. five receptor widths). The result indicates that the best estimate of the receptor signal is obtained primarily from the signals of nearest-neighbour receptors. The interneuron encodes the difference between the actual value of the receptor signal (as indicated by the excitatory centre) and its predicted value (as indicated by the inhibitory surround). This difference, or error, has a standard deviation (E) of 0.133 intensity units in this example. The dynamic range of the interneuron need only be large enough to encode this difference.

Note that the analysis does not require any assumptions regarding how intensity fluctuations in the scene are distributed about the mean level.

Intensity dependence of inhibitory fields

At low light levels, noise caused by the random absorption of photons becomes significant in comparison to the signal. The signal from a given receptor can no longer be predicted reliably from its immediate neighbours, because the receptor signals are contaminated by photon noise. Therefore one expects that the receptive-field surround appropriate for prediction at low light levels (i.e. at poor receptor signal : noise ratios) would differ from that appropriate to high light levels (i.e. at high receptor signal : noise ratios). In quantitative terms, the effect of increased photoreceptor noise is to decrease the non-diagonal ($R_{i,j}$) terms of the R matrix, relative to the diagonal ($R_{i,i}$) terms. This leads to receptive-field profiles in which the inhibitory flanks are weaker and more diffuse.

Figure 6 shows interneuron receptive-field profiles calculated for receptor signal : noise ratios of 10, 1 and 0.1. As the signal : noise ratio worsens, the inhibitory surround becomes weaker and more diffuse. Evidently, as the receptor signal : noise ratio decreases it becomes increasingly profitable to extract the best prediction from a larger group of surrounding receptors, rather than from just the nearest neighbours. This pooling tends to average out the deleterious effects of photon shot noise from the predicted value because the noise is statistically independent among different receptors. At very low intensities, when the incoming noise in the receptor is large in comparison to the signal, the most reliable prediction corresponds to the mean intensity, which the inhibitory field computes by averaging across the entire receptor array (figure 6*c*). For each value of receptor signal : noise ratio, the

calculated receptive field is that which achieves a minimum error between the predicted and the actual receptor signal. The standard deviation (E) of this error is shown alongside each receptive-field profile. This 'error' is, of course, the signal to be transmitted after predictive coding and, when the signal:noise ratio is high, the transmitted signal is half the amplitude of the incoming signal (measured in

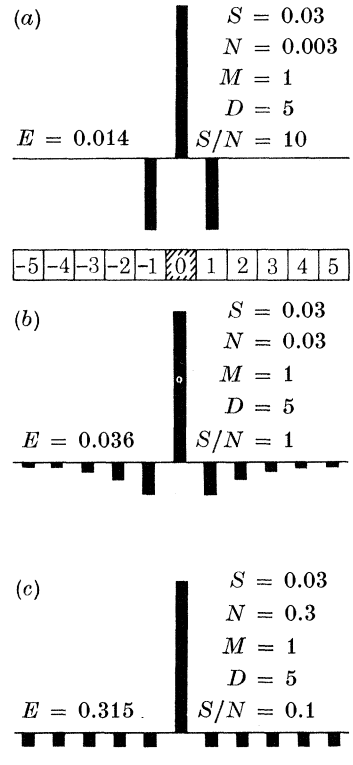

FIGURE 6. As in figure 5, except that the calculations are made for receptor signal:noise ratios of (*a*) 10, (*b*) 1 and (*c*) 0.1.

units of intensity). Thus the interneuron can operate with twice the sensitivity, without additional risk of saturation, as a result of removing some redundant components from the signal by this simple predictive coding procedure.

At lower light levels, the poorer receptor signal:noise ratio causes the error, and hence the required dynamic range, to be larger than the average contrast of the scene (figure 6*b*). Nevertheless, this error is still the smallest that can be achieved at this particular signal:noise ratio; it is smaller than the error that would result from using as a prediction the mean value calculated over the entire receptor array. In the example of figure 6*b*, the standard deviation of the error that results from using an inhibitory field that is tailored for optimum prediction is $E = 0.036$, while the standard deviation of the error that would result from using the global mean as a prediction would be the sum of signal and noise, $S+N$, which is 0.06. In short, the required dynamic range is a function of signal:noise ratio, and hence light intensity.

To summarize the results of figure 6, at high luminances the inhibitory surround forms a tight annulus around the excitatory centre. As the luminance is lowered, the inhibitory surround becomes weaker and more diffuse. Ultimately, at very low luminances, one would expect the surround to be so weak and diffuse as to be experimentally undetectable, except perhaps with a stimulus consisting of a broad annulus of light.

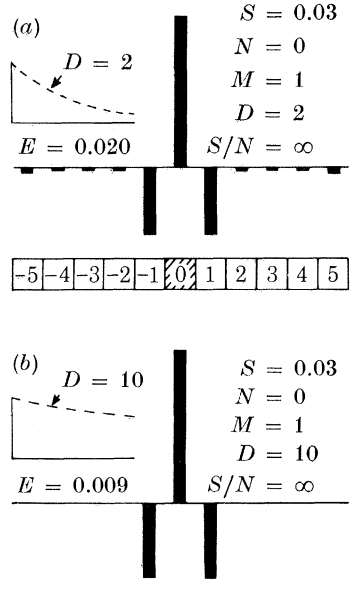

FIGURE 7. As in figure 5, except that the calculations are made for spatial autocorrelation functions with space constants of (*a*) two and (*b*) ten units.

Dependence of inhibitory fields on spatial correlation

Figure 7 *a*, *b* shows receptive fields calculated for spatial autocorrelation functions with space constants of two and ten units, for comparison with the space constant of five units used for figure 5. These figures demonstrate that *the profile required of the inhibitory surround is relatively insensitive to the extent of spatial correlation.* This insensitivity is a desirable feature, because the space constants in real environments vary with the sizes and distances of objects in the environment (figures 3, 4). Photon noise, which can completely randomize the most correlated of inputs, has the more powerful effect upon the inhibitory interactions required for predictive coding. Thus it is most profitable, at least in the first instance, to test for predictive coding by comparing receptive fields at different mean light levels, rather than search for animals that inhabit scenes with radically different spatial characteristics.

Receptive fields in two dimensions

Figure 8 shows receptive fields calculated using the same principles as above, but for a two-dimensional (7×7) receptor array. These receptive fields have the same properties as the one-dimensional fields, with the exception that inhibitory

inputs from remote receptors are weighted comparatively less, partly because in a two-dimensional array of receptors the number of receptors at a given distance from a receptor increases with distance.

Redundancy reduction

An important property of the encoding scheme that we have developed is that it not only increases the accuracy with which retinal images are transmitted to the brain, but also reduces redundancy in the transmitted image by removing all first-order correlations (i.e. correlations described by the autocorrelation function).

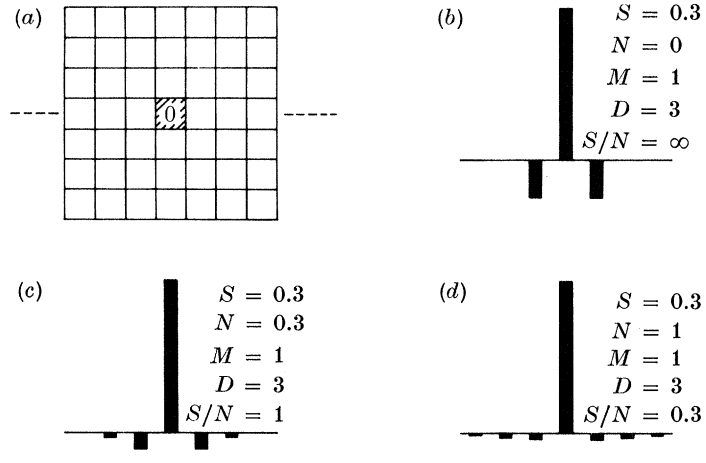

FIGURE 8. As in figure 5, except that the weighting coefficients are now calculated for a two-dimensional (7 × 7) receptor array as shown in (*a*). The encoding interneuron compares the signal from the central receptor (0) with a statistical estimate of this signal derived from the 48 surrounding receptors. The results are shown for receptor signal : noise ratios of (*b*) ∞, (*c*) 1 and (*d*) 0.3, and they depict the weighting coefficients corresponding to the receptors lying along the dashed line of (*a*).

The proof of this statement is by contradiction. If there were any residual first-order correlation in the information transmitted by the array of interneurons, then the signals from at least one pair of interneurons, say i_p and i_q, must be statistically correlated. If this is so, then it should be possible to improve the prediction of the signal of receptor r_p (i.e. reduce the error) by subtracting from the output of i_p a fraction of the output of i_q. However, this would mean that the revised prediction of the signal of r_p is even better than the original prediction, which is itself the best possible linear prediction: a contradiction. Hence, the outputs of the array of interneurons cannot contain any first-order correlations. Thus, in formulating the receptive field that provides for a maximum resistance to contamination by intrinsic noise, we have derived a form of lateral inhibition that satisfies Barlow's (1961*a*, *b*) redundancy reduction hypothesis. To our knowledge this is the first time that a receptive field satisfying this important theory has been developed.

In other words, the lateral inhibition subtracts a statistically predictable

component of the retinal image, and the interneurons transmit only the unpredictable component of the retinal image to the brain. Packing the interneurons' dynamic ranges with the unpredictable component enables the visual system to achieve maximum information capacity, i.e. to transmit the maximum possible number of distinguishable pictures to the brain.

Note that the simple predictive coding scheme described here only attempts to remove the redundancy expected in the *average* scene. It makes no attempt to eliminate the redundancy associated with particular objects, e.g. straight edges and periodic structures. The removal of these specific components probably requires a process that matches its predictive procedures to particular components of the incoming signal, such as oriented edges. There is no evidence at present for such complex procedures being executed at the first or second stages of neuronal processing. Adaptive processing may be more suited to higher levels of the visual system (see, for example, Mackay 1956).

Encoding visual signals in the time domain

Predictive encoding can be extended to the time domain, and a functional role sought for temporal inhibition. Visual interneurons with temporal inhibition (or self-inhibition, as it is also called (Purple & Dodge 1965) have responses that tend to be phasic or transient in nature. They respond primarily to changes of light intensity, and in many cases a constant-intensity light evokes virtually no sustained response. The impulse response of such a unit (i.e. the response to a brief flash) is biphasic, consisting of an initial excitatory phase followed by an inhibitory phase (see, for example, figure 11). The delayed temporal inhibition suppresses the sustained response that would otherwise be evoked by a steady light. Here we formulate the hypothesis that temporal inhibition mediates prediction in the time domain, just as lateral inhibition mediates prediction in the spatial domain.

Consider the problem of encoding the fluctuating signal that is produced by a receptor when the visual scene (or an object within it) moves across the receptor's visual field. The receptor signal will always be correlated over a finite duration; it will never be entirely random. This is so even in the extreme case of a rapidly moving, spatially random visual scene, because the receptor has a finite integration time which smoothes out the intensity fluctuations of the visual scene (Srinivasan & Bernard 1975).

By using a reasoning analogous to that of Snyder (1979), it can be shown that the correlation duration τ of the receptor signal evoked by a spatially random scene moving at an angular velocity W (deg s^{-1}) is given by

$$\tau^2 = (\Delta\rho/W)^2 + (\Delta t)^2, \tag{2}$$

where $\Delta\rho$ is the acceptance angle of the receptor (deg) defined as the half-width of the angular sensitivity function, and Δt is its integration time (s), defined as the temporal half-width of the impulse response function (Howard 1981). Thus, for slow motion ($W \to 0$) the correlation duration would be comparatively large, and would be determined primarily by velocity and the acceptance angle of the receptor; for rapid motion ($W \to \infty$) the correlation duration would be small and

determined primarily by the receptor's integration time (Srinivasan & Bernard 1975).

Given that the receptor signal is temporally correlated, it should be possible to make a statistical prediction of its present value, based upon its past history. The interneuron would encode the difference between the actual present value of the

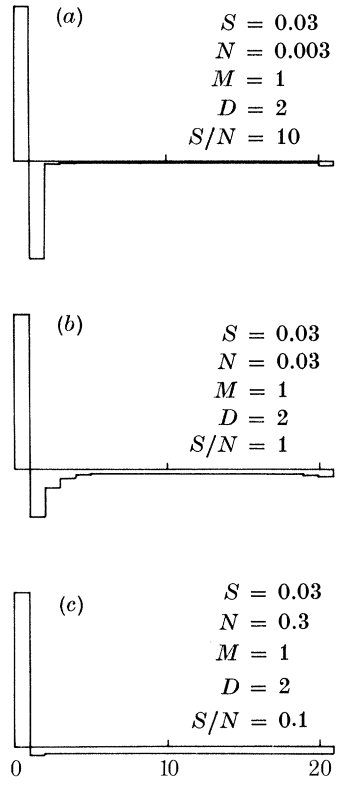

FIGURE 9. Predictive coding in the time domain. The figure shows temporal weighting functions calculated for receptor signal:noise ratios of (*a*) 10, (*b*) 1 and (*c*) 0.1. Time is depicted in units of receptor integration time, and the parameter D now refers to the time constant of the temporal autocorrelation function of the receptor signal, specified in these units. All other parameters have the same meaning as in earlier figures.

receptor signal and a statistical prediction of this value based on a weighted linear sum of past values. In this scheme, then, the role of self-inhibition would be to generate an appropriately weighted sum of the past values, and to subtract the resulting prediction from the excitatory signal corresponding to the present value.

While the predictive temporal encoding that we are about to describe is analogous to predictive spatial encoding, there are two noteworthy differences. First, spatial prediction can be an omnidirectional process (i.e. it can be based on signals from all of the surrounding receptors), while temporal prediction is necessarily unidirectional (it is based only on past values of the receptor signal,

not on future values). Secondly, spatial prediction is based on samples of the signal that are taken at discrete locations (corresponding to individual receptors) while temporal prediction is based on a continuous signal. Neither distinction is problematic in our context, and in our calculations we have, for convenience, treated the receptor signal as being 'sampled' in discrete time bins, each having a duration equal to the receptor's integration time. This treatment renders the temporal prediction problem analogous to the spatial prediction problem, and one can then work out, using the theory outlined above, how the past values of the receptor signal in the various bins should be weighted in order to obtain the best prediction of the present value. (Making the duration of each temporal bin equal to the receptor integration time renders the noises in the various bins statistically independent, as in the spatial case.)

Figure 9 shows temporal weighting functions calculated for various values of receptor signal:noise ratio. These functions give the 'impulse response' of the interneuron, which is the temporal equivalent of the receptive field. The impulse responses (figure 9) have been calculated by assuming that the receptor signal is correlated over two receptor 'integration times'. As with spatial coding, the results are relatively insensitive to the extent of correlation, but show a strong dependency on the signal:noise ratio. At high receptor signal:noise ratios the prediction is based upon recent values of response (figure 9*a*) but in the presence of substantial noise the prediction is based upon a mean value determined over a considerable interval (figure 9*c*). This prolongation of inhibition at low signal:noise ratios is analogous to the extension of the inhibitory surround of the receptive field.

As in spatial encoding, it can be proved that predictive temporal encoding suppresses the predictable component of the receptor signal and transmits only the unpredictable component. Thus, there would be no first-order correlation (between bins) in the temporal signal transmitted by the encoding interneuron, and this type of encoding would help the interneuron transmit to the brain the maximum possible number of distinguishable temporal patterns per second.

Comparison of theory with experimental data

If the predictive spatial encoding that we have described is in fact used by visual systems, the most likely application would be in high-acuity form vision. Thus, one might look for such a strategy in the receptive fields of bipolar cells or *X*-type ganglion cells of a vertebrate fovea or area centralis. To test the theory, one has to measure receptive fields at several different luminances, and compare these with receptive fields predicted on the basis of the signal:noise ratios at these luminances. Although there is insufficient information in the literature, the available data fit the predictions of our theory. At high light levels (when the receptor signal:noise ratio is high) the inhibitory surrounds of *X*-type ganglion cells form a tight annulus around the excitatory centre. At low light levels (when the receptor signal:noise ratio is low) the surrounds become weak and diffuse, and eventually undetectable. The experimental evidence for this comes from mapping ganglion-cell receptive fields by using flashing spots of light (Kuffler 1953; Rodieck & Stone 1965) or moving bars of light, by making area-threshold measurements (Barlow *et al.* 1957,

figs 1, 4; Barlow & Levick 1976, fig. 4) or by measuring the spatial contrast-sensitivity functions by using sinusoidal gratings (Enroth-Cugell & Robson 1966, fig. 15). All of these data support the hypothesis that *X*-type retinal ganglion cells have receptive fields that are designed for optimum spatial encoding of slowly moving objects. Similarly, human psychophysics suggests that lateral and temporal antagonism, manifest by the depression of contrast sensitivity at low frequencies, weakens at low intensities (Campbell & Robson 1968; Kelly 1961).

As predictive coding uses the receptive field surround to predict the signal expected at the centre, it formulates neither the centre diameter nor the distance between the centres of adjacent neurons. Our analysis assumes that these critical sampling parameters are determined by factors such as the requirements for regions of high spatial acuity and the need for area summation to enhance sensitivity at low intensities. Thus, if one adopts a more general point of view and considers figures 5–8 as representing arrays of excitatory centres, rather than arrays of photoreceptors, then our analysis predicts that, at a given signal:noise ratio, visual systems with larger excitatory centres (i.e. coarser sampling mosaics) should have proportionately larger inhibitory surrounds. This geometrical scaling would also apply to a system in which interneurons with centres of different sizes coexist in the same retinal region (see, for example, Wilson & Bergen 1979), provided that the respective surrounds are generated from mosaics of correspondingly different coarseness. For many retinal neurons, however, the surround may not be generated from sampling units that have the same diameter as the centre. It is possible to formulate predictive surrounds for given centres without the arbitrary constraint of an unduly coarse sampling mosaic, but the lack of pertinent quantitative data from vertebrate retinal neurons renders the added complications to our analysis superfluous at the present time.

Preliminary calculations with arbitrarily fine sampling mosaics indicate that, at high light levels (i.e. infinite signal:noise ratio), the surround should be created only from sampling units in the immediate vicinity of the excitatory centre, regardless of how large (or small) these units are in comparison with the centre. However, at lower light levels (i.e. moderate signal:noise ratios) the number of units that need to be pooled to create the surround increases with the fineness of the array. The reason is that the finer the array the lower the signal:noise ratio per unit, which in turn requires that more units be pooled to construct the surround (see figure 8). The surround can thus comprise a small number of relatively coarse units or a larger number of fine units. The expectation, therefore, is that at moderate and low light levels the angular extent of the inhibitory surround would not vary dramatically with the fineness of the sampling mosaic.

We can test rigorously for predictive coding in the compound eye of the fly, *Lucilia cuprina*. The receptive fields of a major class of first-order afferent interneurons, the large monopolar cells (l.m.cs), have an antagonistic centre–surround arrangement. The centre corresponds to the finest spatial unit of the photoreceptor mosiac so that the organization of this eye is identical to the disposition of photoreceptors in our model (Kirschfeld 1967; Braitenberg 1967; Zettler & Järvilehto 1972; Mimura 1976). Furthermore, by using established intracellular techniques (Hardie 1979; Laughlin & Hardie 1978), it is possible to

measure the parameters relevant to predictive coding, namely the signal:noise ratios and the shapes of receptive fields at different intensities.

Measurements of signal:noise ratio were made at two mean luminances, 10.0 and 1.26 cd m^{-2}, by using as the signal a sinusoidal grating of contrast 0.4 and spatial

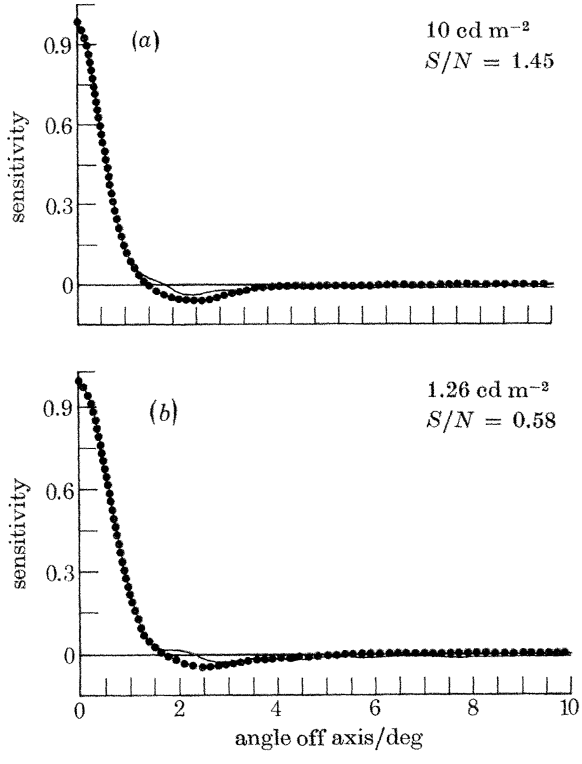

FIGURE 10. Comparison of experimentally measured receptive fields of large monopolar cells (l.m.cs) in the fly's visual system (solid curves) with fields theoretically expected on the basis of predictive spatial encoding (dotted curves). The l.m.c. receptive fields and signal:noise ratios (S/N) were measured at luminances of (*a*) 10.0 and (*b*) 1.26 cd m^{-2}. In each case, the theoretical receptive field was calculated by assuming a 7×7 receptor array, arranged as in figure 8*a*, with an interreceptor spacing of 1.25° in either direction (W. Ribi & D. R. Dvorak, unpublished measurements of *Lucilia* interommatidial angles). The resulting array of weighting coefficients was convolved with a circularly symmetrical, two-dimensional gaussian function of half-width 1.5°, representing the angular sensitivity function of each receptor (Hardie 1979; Dubs 1982) to obtain the overall receptive field of the l.m.c. The theoretical curves depict the mean values of excitation or inhibition expected at various angular distances from the l.m.c. visual axis; these were obtained by averaging the excitation or inhibition from individual receptors in the 7×7 array that lay within concentric rings, each 1° thick, centred on the visual axis.

wavelength 20°, drifting at a constant velocity of 240° s^{-1}. The contrast corresponds to the mean contrast in natural scenes (as determined from 70 scans of the type shown in figure 3) and the combination of pattern wavelength and velocity generates a temporal modulation in each receptor and l.m.c. of 12 Hz, which is close to the peaks of their respective frequency responses (French & Järvilehto 1978). Root mean square noise levels were measured directly from the neuron's graded

responses and, because the relation between response amplitude and contrast is approximately linear for small contrasts, the noise levels were equated with the response to the grating. The receptive fields of the l.m.cs were determined at the same mean luminances by using sinusoidal gratings of various spatial wavelengths to determine the spatial frequency response and then subjecting these data to an inverse Fourier transform so as to obtain the spatial profile of the receptive field (full details of this method are given in Dubs (1982)).

The measured receptive fields of l.m.cs change with mean luminance in the way predicted by the theory. As the signal:noise ratio drops, the surround extends and weakens, but under all the conditions that we have measured the surround is somewhat more diffuse than the theory predicts (figure 10). It is clear that the lateral inhibition exhibited by fly l.m.cs does not minimize the dynamic range required to register fine spatial detail. Can any functional significance be attached to the observation that this class of interneurons does not precisely conform to the predictive encoding hypothesis developed for stationary, or almost stationary, patterns?

We suggest that fly l.m.cs are better suited for sampling rapidly moving scenes partly because the powerful transients essentially abolish the response to a step change in contrast within 250 ms (Laughlin & Hardie 1978). Significant image movement modifies the predictive coding hypothesis in two ways. The first is comparatively trivial. The finite duration of the receptor response blurs moving objects, decreasing their apparent contrast (Srinivasan & Bernard 1975) and so reducing the signal:noise ratio. To explain the broad surrounds of l.m.cs this motion blur would have to demodulate the contrast to 10% of its original value, and for light-adapted fly photoreceptors such demodulation is experienced by temporal frequencies of 120–150 Hz (Zettler 1969; Leutscher-Hazelhoff 1975). To achieve this frequency the finest resolvable grating would be moving at approximately $600° s^{-1}$ and coarser gratings would have to move faster. Flies do turn at these speeds (Collett & Land 1975). However, motion at this, and at lower velocities, requires a second and more substantial modification to the predictive coding hypothesis because the photoreceptors and the inhibitory network cannot adequately resolve the correlated parts of the pattern. The problem of coding moving scenes is discussed more fully below.

The predictive encoding of moving scenes

When a scene moves at constant velocity one can, in principle, exactly predict the present value of the signal in a given receptor by simply recording the past value of the signal from another receptor further 'upstream'. This is not a statistical prediction but a purely deterministic one, relying on *a priori* knowledge of the speed and direction of motion. The simplicity and attractiveness of this scheme is deceptive. One must take into account that motions are generally unpredictable in direction and speed, so that this exact prediction requires neuronal circuitry that continually measures speed and direction of motion and uses this information to locate the inputs from the relevant receptors. This degree of complexity seems unsuitable for early processing and there is little evidence that retinal neurons are involved in such sophisticated computations.

Let us return to the original notion of statistical spatial prediction, and see how it can be applied to a moving scene. Assume that the visual scene is spatially correlated over 2° and moves at an angular velocity of 300° s^{-1}. If the photoreceptor integration time is taken to be 20 ms (Howard *et al.* 1982), the scene would move a distance of 6°, or three spatial correlation lengths, during this time. Thus, at any given instant, only $1/\sqrt{3}$ of the amplitude of the photoreceptor's response would, on average, depend upon the spatial intensity pattern that is presently in the vicinity of the receptor (i.e. within one spatial correlation distance, 2°). The remaining fraction of the response amplitude would depend upon patches of the scene that are further away, and would consequently be of no use in extracting any statistical prediction of local intensity other than the mean value. (A quantitative description of this effect is given in Appendix 2). Thus, as the speed of the scene increases, the 'predictive' fraction of the receptor response diminishes. This reduction applies to a receptor response that itself decreases in amplitude as the scene moves faster, because the receptor behaves as a low-pass filter in the time domain (Zettler 1969). Thus, it is clear that, if one assumes that photocreceptor noise is independent of object speed, then the signal:noise ratio of the 'predictive' fraction of the photoreceptor response steadily diminishes as the object speed increases, until one reaches the limiting condition when the only statistically useful component of the receptor signal is its mean value. In this limiting situation, one would expect lateral inhibition to be uniformly spread out over a large number of receptors, because the entire inhibitory field would be geared to predicting only the mean value of the receptor signal. Apparently, this is roughly the condition under which the fly's l.m.c. neurons operate. It is interesting to note that increasing the object speed has the same effect as lowering the luminance: both reduce the signal:noise ratio of the predictive component of the receptor signal, although for different reasons. Thus, diffuse spatial inhibition is to be expected at low luminances or at high object speeds. The former situation applies to *X*-type retinal ganglion cells at low luminances, while the latter situation applies to the fly's l.m.c. neurons at moderate levels of illumination (10 cd m^{-2}). Interestingly, lateral inhibition in the l.m.cs appears more diffuse when measured with rapidly moving gratings than with slowly moving ones (Dubs 1982).

Given that the fly's l.m.c. neurons might be better suited to encode rapidly moving scenes, it is interesting to examine whether the transient signals from the excitatory centres of their receptive fields are involved in *temporal* prediction.

Testing for temporal predictive coding

Using standard intracellular recording techniques, we are able to test whether fly l.m.cs employ predictive coding in the time domain because we can measure the relevant parameters, namely the l.m.c. 'impulse response', the photoreceptor's time constant and the signal:noise ratio. The l.m.cs were adapted to different background levels, spanning a range of three logarithmic units, and were presented with a brief (5 or 10 ms) flash superimposed upon the background to produce a small response (less than 7 mV); 200 or 400 consecutive responses were averaged to enhance the reliability of the data. At each background intensity, the relation between response amplitude and stimulus contrast was determined as described

by Laughlin (1981 *b*), and the voltage noise level was measured. At each background intensity, the known linear relation between contrast and response amplitude was used to convert the cell's voltage noise into the standard deviation of an equivalent contrast signal. Finally, the signal:noise ratio at each background intensity was obtained by dividing the mean contrast in the natural scenes that we surveyed (0.4) by the noise, expressed in equivalent contrast, at that background. The

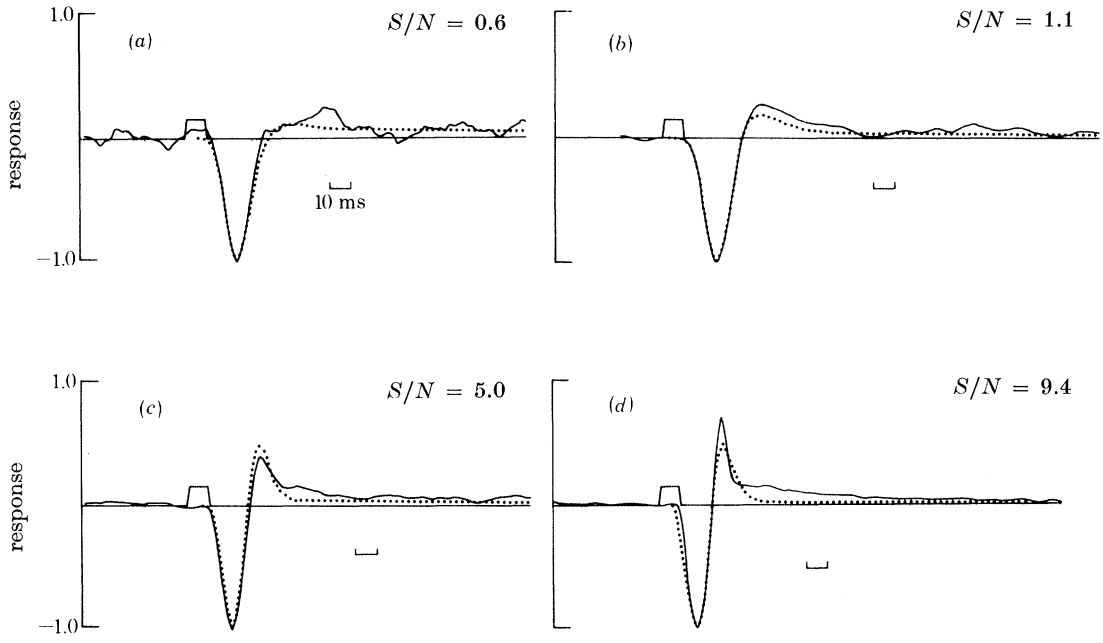

FIGURE 11. Comparison of experimentally measured impulse responses of large monopolar cells (l.m.cs) in the fly's visual system (solid curves) with fields theoretically expected on the basis of predictive temporal encoding (dotted curves). The l.m.c. impulse responses and signal:noise ratios (S/N) were measured at four different luminances, spanning a range of three logarithmic units. In each case, the theoretical weighting function was calculated as shown in figure 9, a receptor integration time of 10 ms being assumed. For each luminance, this theoretical weighting function was then convolved with the receptor's impulse response. Over the range of luminances that was investigated, the receptor's impulse response did not change substantially in time course, and its half-duration lay in the range 8–12 ms. At each luminance, the receptor's impulse response was well approximated by the logarithmic–normal function $v(t) = e^{-[\log(t/t_p)]^2/2\sigma^2}$ (Howard 1981). This approximation was used in computing the convolutions, with $t_p = 14.6$ ms and $\sigma = 0.34$ in (*a*), $t_p = 11.4$ ms and $\sigma = 0.31$ in (*b*) and $t_p = 10.0$ ms and $\sigma = 0.34$ in (*c*) and (*d*). Scale bar in (*a*) applies also to (*b*)–(*d*).

impulse response of the receptor, and its integration time (Δt in equation (2)) was measured at the appropriate background intensities by using the techniques described by Howard (1981). The measured signal:noise ratios and integration times were used to model the impulse response function required for predictive coding as described in the legend of figure 11, and the results were compared with the measured impulse response. There is good agreement between experiment and

theory. In particular, the inhibition appears at the appropriate signal:noise ratio and has the strength and time course required for predictive coding. In the fully light-adapted cell ($S/N = 9.4$ in figure 11), the inhibition is slightly more powerful than is required for predictive coding and this suggests that, in addition to removing correlation and minimizing the dynamic range that is required, the system is further accentuating high frequencies.

It is interesting to note that retinal ganglion cells in the cat exhibit a powerful temporal inhibition of their centre responses which disappears at low intensities (Enroth-Cugell & Shapley 1973). Unfortunately measurements of signal:noise ratios and of receptor integration times are not yet available for this system so that it is impossible to assess the degree to which temporal inhibition effects predictive coding in these cells.

Discussion

The theory of predictive coding provides a quantitative analysis of the function of lateral and temporal inhibition in the retina. As recently and independently suggested by Barlow (1981), the inhibitory interactions exploit the correlations between nearby response levels to generate a statistical estimate of the signal expected at a certain point in space or time. This prediction is then subtracted from the signal that actually arrives at the point, so as to reduce the range of signal amplitudes that are encountered by the removal of a redundant component. Thus incoming signals can be amplified more, so as to render them more resistant to contamination by intrinsic noise generated at higher levels.

Our formulation of the predictive coding hypothesis has a number of advantages. Its execution requires simple neuronal interactions of the type exhibited during the early processing of visual information. The analysis is done in the spatial and temporal domains: so by modelling directly the local interactions it seeks to explain without recourse to the added complications of Fourier analysis. This choice also makes it easier to identify a major form of redundancy from the autocorrelation function. Moreover, the quantitative predictions of the theory are substantive because they are formulated without recourse to assumptions about the value or use of different signal components. Predictive coding regards all information as being equally important, irrespective of whether it is carried by low or high frequencies. It only removes redundancy to help minimize the known limitations of neuronal signal capacity, and it does this in an objective, statistically based fashion. Most important of all, we have verified the quantitative predictions of the theory by analysing receptive fields and patterns of temporal antagonism in the fly's retina.

The analysis of predictive coding, and its successful application to the compound eye of the fly, raises several questions. First, what effects do retinal nonlinearities have on predictive coding, as analysed here? Secondly, what is the relation between predictive coding and the previous theories advanced for the function of retinal antagonism, and to what extent are these theories fulfilled by the available data? Finally, since predictive coding is a simple means of promoting coding

efficiency, how important is the 'neat packaging of information' (Barlow 1961*b*) to retinal function and what factors are likely to influence coding efficiency?

The effects of nonlinear retinal transfer functions

For simplicity we have analysed the case of predictive coding in a linear network, but our conclusions are not substantially altered by two types of nonlinearity found in retinas. These are the compressive intensity–response functions of photoreceptors (Norman & Werblin 1974), which combine with light adaptation to produce an effectively logarithmic transformation of the intensity signal, and the nonlinear synaptic transfer of signals from photoreceptors to interneurons (Werblin 1974; Shaw 1978). These monotonic nonlinearities do not decorrelate retinal signals, they modify space and time constants of correlation and alter signal:noise ratios. Predictive coding is rather insensitive to the space and time constants of correlation so that, in this regard, nonlinearities will have a negligible effect upon the receptive fields it requires. Since the precise forms of retinal transfer functions are often intensity-dependent, such insensitivity could be advantageous. For the more important parameter, signal:noise ratio, the linear analysis shows that predictive coding is most sensitive to changes when signal and noise have similar amplitudes (figures 6, 9) and this is precisely the condition under which nonlinearities have least effect.

In the fly retina, where we rigorously tested and confirmed the validity of our predictive coding model, the relevant nonlinearities have been measured (Laughlin & Hardie 1978). They are not severe enough to necessitate the added complications of a nonlinear model. This is principally because most of the signals generated by natural scenes have contrasts between zero and 0.5 and over this range the amplitude of receptor or interneuron response is proportional to contrast (Dubs *et al.* 1981; Laughlin 1981*b*). Essentially, the receptors' logarithmic transformation ensures that gain is inversely proportional to mean intensity, as required for contrast coding (Laughlin 1981*a*). Thus, in any retinal region operating at one mean luminance, an approximately linear contrast signal is superimposed upon a d.c. bias, and this bias voltage is a logarithmic function of mean intensity. Subtractive inhibition removes the bias from the responses of fly first-order neurons (Laughlin & Hardie 1978). Predictive coding suggests that the area over which this bias is computed should be such that sensitivity to small contrast fluctuations can be maximized. Thus, compressive intensity–response functions operate in conjunction with predictive coding to enable the fly's retina to operate over a wide range of intensities.

However, there are instances of severe retinal nonlinearities that invalidate our linear model (e.g. nonlinear summation within the receptive fields of *Y*-type retinal ganglion cells (Enroth-Cugell & Robson 1966)). None the less, the principle of predictive coding does not require linearity. In its most general form, predictive coding uses not only first-order correlations in the visual scene (as we have done here) but higher-order correlations as well, to generate a prediction based on a nonlinear function of receptor signals in the vicinity. The general theory is outlined in Papoulis (1965). However, its application is warranted only when the existence

of higher-order correlations in visual scenes is established and their statistics are measured. The volume of data and the computational effort increase exponentially with the order of the correlation.

The functions of lateral and temporal inhibition

Several roles have been advanced for lateral inhibition (see, for example, Hartline & Ratliff 1972) and each of these prospective functions has been, or can be, extended into the time domain to provide a role for temporal inhibition. All these functions are undoubtedly performed by inhibition because each role is simply a reformulation of lateral inhibition's filtering properties. However, to establish the significance of each, one must determine the degree to which the measured inhibitory interactions of neurons execute each of the particular roles. As we have seen in trying to assess the degree to which predictive coding is done, many of the relevant data, and in particular the measurements of signal : noise ratio, are lacking. None the less, a comparison made between the basic philosophies of each proposed role and, wherever possible, the quantitative predictions of the inhibitory patterns that they require is instructive because it suggests that the three roles, predictive coding, redundancy reduction and d.c. bias elimination, are interrelated and are of more importance than either deblurring or edge enhancement.

(i) *Redundancy removal*

This could yet prove to be the most fundamental formulation of the functions of lateral and temporal inhibition. It simply states that sensory systems receive signals that are highly redundant and that neural processing should, in the first instance, be directed towards minimizing this redundancy so that the sensory information can be represented most compactly and processed most economically (Attneave 1954; Barlow 1961*a–c*, 1981). To our knowledge the only previous formulations of inhibition's role in redundancy reduction have been qualitative; so our finding that predictive encoding minimizes a certain form of redundancy is of some interest. The analysis of predictive coding also vindicates the original suggestion that redundancy reduction is beneficial, since it removed redundancy while setting out to protect signals against intrinsic noise. Clearly, it is advantageous to remove the redundant so as to provide more space for the better representation of the essential message.

We emphasize again that predictive coding, as presented here, does not eliminate all redundancy from the visual signal. To do so would be a daunting task since the retinal image is composed of millions of samples, many of which can be related by statistics of a high order (Kretzmer 1952). Predictive coding only removes the redundancy between pairs of points, as manifested by the autocorrelation function, yet this simple coding procedure is obviously effective because it halves the amplitude of the signal, as expressed in intensity units (figure 6).

(ii) *Removing the d.c. bias*

The d.c. bias is the standing component of the signal upon which all fluctuations in intensity are superimposed. There is no question that the d.c. bias is redundant

(Barlow 1961*a*) since it remains constant and unchanging, and that by subtracting the bias away the range of signals is reduced to allow for greater sensitivity (Barlow & Levick 1976). In implementing this scheme the critical question concerns the extent, in space and time, over which the d.c. bias should be determined because, strictly speaking, the d.c. bias refers to the zero frequency component and is computed as a single value spanning either the whole retina or a very long time period. However, this global definition of bias is usually inappropriate for coding the information from a particular point. For example, if a scene is equally divided into an area of highlight and an area of shade, the global bias will fall midway between the biases of each half. This estimation of bias would be very unsuitable because its subtraction will annihilate the shaded part, and leave substantial bias within the area of highlight. The predictive coding theory is an extension of the d.c. bias hypothesis which addresses the question of the appropriate span over which the bias, or local mean signal, should be determined. Because it subtracts an appropriately weighted amount of the signal from each point so as to minimize the dynamic range, it defines the optimum extent of space and time over which the mean should be computed, and hence the optimum extent of lateral and temporal inhibition. Indeed, wherever the space and time constants of retinal inhibition have been measured, the results indicate that the retina computes a 'local' mean of the intensity, rather than a 'global' mean (e.g. in *Limulus* (Brodie *et al.* 1978), fly (Laughlin & Hardie 1978; Dubs 1982) and cat (Enroth-Cugell & Robson 1966; Enroth-Cugell & Shapley 1973)).

(iii) *Deblurring*

The idea that lateral inhibition compensates for optical blurring by squeezing spread edges back to shape is as old as the idea of lateral antagonism itself (Mach 1865), and the requirements for deblurring are readily obtained by applying Fourier analysis (Ratliff 1965). A completely deblurred response will represent all frequencies below the cut-off with equal sensitivity. Because optical sampling and transduction attenuate high frequencies, inhibition must selectively reinforce the high-frequency components, as indeed happens, to produce a mechanism with a flat frequency response (figure 12*b*). By comparison the predictive encoding hypothesis postulates an overall frequency response with a pronounced low-frequency roll-off (figure 12*a*).

Most of the experimental evidence available in the literature favours the predictive encoding hypothesis, and argues against significant deblurring occurring in the early stages of visual processing. First, predictive encoding predicts the widely observed low-spatial-frequency attenuation (in man (DeValois *et al.* 1974), monkey (De Valois *et al.* 1974), cat (Enroth-Cugell & Robson 1966), *Limulus* (Brodie *et al.* 1978) and fly (Dubs 1982)) while deblurring does not. Secondly, signals coming from the surround of a receptive field can be selectively abolished by using stimuli with high temporal frequency, because the surround has a longer time constant than the centre (Robson 1966; Brodie *et al.* 1978; Dubs 1982). Thus, gratings that move rapidly or reverse contrast frequently can be used to evaluate visual performance without the surround, and thereby assess the functional role of the surround. If the surround played an encoding role, one would expect that spatial

contrast-sensitivity functions (c.s.fs) measured with low- and high-temporal-frequency gratings would differ as shown respectively by the solid and dashed curves of figure 12*a*. The low-spatial-frequency attenuation would disappear at high temporal frequencies, while the high-spatial-frequency sensitivity would be essentially unaltered. This is because the primary function of lateral inhibition in this case is to suppress the predictable component of the signal (i.e. d.c. and low

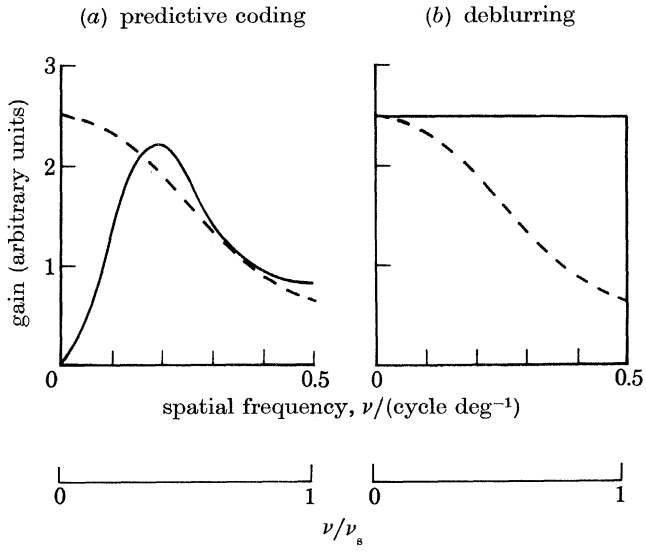

FIGURE 12. Comparison of the spatial contrast-sensitivity functions that are expected from a predictive coding scheme (solid curve, left) and a deblurring scheme (solid curve, right). The contrast-sensitivity function (c.s.f.) shows gain (plotted on an arbitrary linear scale) against spatial frequency (cycle deg^{-1}). We have assumed that the angular sensitivity function of each receptor is gaussian with a half-width of 1.5°, and that the receptors are arranged in a two-dimensional array as in figure 8*a*, spaced 1° apart in either direction. The dashed curve represents the c.s.f. of the optics alone; it is the Fourier transform of the receptor's angular sensitivity function. The c.s.f. for predictive coding (solid curve, left) was obtained by convolving the weighting coefficients of the example of figure 8*d* with the receptor's assumed angular sensitivity function and taking the Fourier transform of the resulting overall receptive field. (The slight levelling-off of this curve at the high-frequency end is an artefact arising from using a finite-sized array (7 × 7) in computing the receptive fields of figure 8.) The solid curve on the right represents the c.s.f. of an ideal deblurring scheme which transmits uniformly all spatial frequencies ranging from zero to the sampling limit of the array (0.5 cycle deg^{-1}). The c.s.f. for predictive coding can be geometrically scaled to apply to visual systems that have finer or coarser mosaics than the example shown here, as discussed in the text. This is illustrated by the lower horizontal scales, wherein spatial frequency ν is normalized to the sampling limit of the array, ν_s.

spatial frequencies). On the other hand, if the surround played a deblurring role, one would expect that the spatial c.s.fs measured with low- and high-temporal-frequency gratings would differ as shown respectively by the solid and dashed curves of figure 12*b*. In this case, the sensitivity to high spatial frequencies would become poorer at high temporal frequencies, while the low-spatial-frequency sensitivity would be essentially unchanged. The available experimental data (for

man (Robson 1966), *Limulus* (Brodie *et al.* 1978) and fly (Dubs 1982)) follow the prediction of figure 12*a*, not figure 12*b*, thus arguing against deblurring and favouring the encoding hypothesis. That is not to say that deblurring is never performed by the visual system. Indeed, the phenomenon of 'contrast constancy' described by Georgeson & Sullivan (1975) suggests that high spatial frequencies are selectively boosted, especially when their contrast (or signal:noise ratio) is high (Snyder & Srinivasan 1979). However, we suggest that this type of image restoration is not a major consideration in early visual processing. In our opinion, the periphery is concerned with encoding the visual image in a simple and efficient way, rather than with accurately reconstructing it.

(iv) *Edge enhancement*

Any mechanism that selectively attenuates low frequencies brings the high frequencies contained by edges into greater prominence. In addition, where objects are uniform across their extent, little information is provided by intensity readings made between edges, so that edge enhancement removes this source of redundancy (Barlow 1961*a*). The distinctive characteristic of the edge-enhancement hypotheses is the concept of feature extraction. For example, the detailed theory of early visual processing set out by Marr & Hildreth (1980) uses centre–surround receptive fields with gaussian profiles to isolate points of inflexion in intensity distributions. These 'zero-crossings' delineate the midpoints of blurred edges, provided that the edges are relatively straight. In addition, given that the spatial frequency components of the relevant signal are limited to one octave, the positions of zero-crossings completely describe the retinal image. The necessary receptive fields are very similar both to the 'difference of gaussian' receptive fields inferred from human psychophysics (Wilson & Bergen 1979) and to the types of receptive field required for predictive coding. This similarity emphasizes that, in common with alternative functions of lateral inhibition, edge detection and predictive coding are in no way exclusive. The more advantages a given filtering or sampling procedure has, the better! The difference is that predictive coding takes into account the qualities of the retinal image in order that it might be encoded within the constraints imposed by neuronal signals. By comparison, edge detection isolates a single characteristic of a scene, that can, through its spatial distribution, provide an adequate and compact description, thought suitable for subsequent processing at higher levels. Until we know more about the dependence of higher-order processing on filtering at lower levels, the assessment of the role of edge detection is difficult.

Coding efficiency in early visual processing

Our comparisons between the different roles of lateral inhibition suggest that deblurring is relatively unimportant, and that the case for edge detection is unproven. Inhibition is far more effective at reducing redundancy and removing d.c. bias, in the manner suggested by predictive coding. The essence of this coding strategy is to remove some components of the incoming signal in the interests of what Barlow called 'the neat packaging' of information into neurons (Barlow 1961*b*). Why is efficient coding necessary and how does it depend upon the properties of the incoming signal and noise, and upon the filter characteristics of

the neurons that process this input? These fundamental questions can be explored by discussing the findings and implications of the predictive coding hypothesis.

(i) *Why is coding necessary?*

If the nervous system operated without introducing intrinsic noise then there would be no need to encode for protection against it. However, the inevitable presence of intrinsic noise sets a limit on the information capacity of any interneuron by defining a smallest resolvable signal, even when the input is noise-free. Thus, any component of the signal that is attenuated during processing has its information content reduced because it can no longer be measured so accurately against the background of intrinsic noise. Given that high spatial frequencies are inevitably attenuated by the optics (see, for example, Campbell & Gubisch 1966), a coding strategy is necessary to protect signals from intrinsic noise. Moreover, as intrinsic noise limits the information capacity of single neurons, and the number of neurons available in the retina is finite, there are definite advantages to be obtained from 'neatly packing' (Barlow 1961*b*) the incoming signals into the available channels, by removing redundancy. As we have seen, predictive coding provides for both noise-resistant encoding and redundancy reduction.

(ii) *What information is lost during coding?*

It may not be necessary to throw away information in order to eliminate redundancy but it appears that inhibitory mechanisms make this sacrifice, in the interests of 'neat packaging'. Examination of the frequency response of an inhibitory mechanism shows that it rejects information about low frequencies (figure 12*a*). These components are attenuated so that, when intrinsic noise is added, their resolution is impaired. It is interesting to note that the seemingly drastic procedures of predictive coding, which reduce the input to a small 'difference signal', only completely reject information about one frequency component, the zero frequency or the true d.c. bias (figure 12*a*). All other frequencies are represented to a greater or lesser degree, but low-frequency resolution is sacrified in order that fine detail can be resolved.

(iii) *Some factors influencing coding*

The predictive coding theory illustrates that, to be effective, the coding filters are matched to the characteristics of the incoming data, the incoming noise, the response range of neurons, and the intrinsic noise of the nervous system. The idea of matched sampling and filtering is well established in visual systems. It is argued that the spacing of photoreceptors is matched to the quality of the optical image (Miller 1979; Snyder 1979) and that the mosaics of retinal ganglion cells are matched to the predictable patterns of retinal input generated by factors such as the increasing fineness of texture with distance from the eye, and to the flow patterns that are produced in the visual field by locomotion (Wehner 1981). This matching eliminates the redundancy that might result from oversampling (Hughes 1977, 1981). It has also been shown that the nonlinear contrast coding functions of visual interneurons can be matched to the statistical distributions of contrast

in natural scenes to promote the most efficient use of a neuron's dynamic range, again through the elimination of redundancy (Laughlin 1981*b*). Note that this last finding is related to the role of inhibition advocated here, because the shape of the receptive field and the extent of the temporal inhibition also determine the dynamic range that is required. This important relation is well demonstrated by the predictive coding hypothesis. If there is weak inhibition, or the inhibition has a longer time constant or a wider space constant, then the difference between the predicted and actual signals increases. Thus, in any neuron's functioning, there should be a relation between the statistical distribution of its inputs, its filtering properties and its dynamic range.

(iv) *The influence of movement*

Movement, and in particular the speed and predictability of movement, has a profound effect upon predictive coding for it completely alters the pattern of correlation, at least in the spatial domain. Although we have discussed qualitatively the effects that motion should have on the receptive fields of fly l.m.cs, our rigorous quantitative treatment has been confined to the spatial domain on the one hand, and the time domain on the other. We have not attempted the more general problem of predictive coding in the spatio-temporal domain. Extending the analysis to the spatio-temporal domain is technically challenging, but conceptually straightforward. In principle, one is interested in predicting the present value of the receptor signal, based upon past values of the signal in this receptor and in surrounding receptors. To derive the spatio-temporal characteristics of the desired receptive field, one would first have to measure, for a scene with realistic spatial and temporal (movement) characteristics, a series of spatio-temporal correlation coefficients and use them in a matrix analogous to that of equation (1*b*). A typical element of this matrix would be the correlation coefficient between the present value of the signal at receptor i and the value of the signal at another receptor j, measured t seconds ago; various elements of the matrix would pertain to various i, j and t. The result of inverting this matrix would be a set of weighting coefficients that specify the characteristics of the desired receptive field in space and time. While this type of calculation has to await measurement of the spatio-temporal correlation coefficients, a task that is by no means trivial, one can make tentative guesses as to the kind of spatio-temporal receptive field that is likely to result. Given that the scene is likely to be moving, the future time course of a receptor's signal can best be predicted by looking at the immediate history of the signals of receptors in the immediate vicinity. Remote receptors would not be very helpful in statistical prediction, but would aid in obtaining a reliable estimate of mean intensity. Thus, one would expect fast-acting, short-lived inhibition from nearby receptors and slower longer-lasting inhibition from remote ones. Detailed measurements of the spatio-temporal characteristics of inhibition are not yet available in the literature.

(v) *Is coding efficiency of general importance to retinal function?*

Retinal bipolar and ganglion cells and the large monopolar cells of compound eyes share three important properties: lateral antagonism, temporal inhibition and

the narrow dynamic range that antagonism enables (Laughlin 1981*a*). These three properties enhance coding efficiency by sweeping away redundancy to make room for the better representation of the essential information content of the incoming data. As has been pointed out for the related principle of a nonlinear dynamic range that enhances information capacity (Laughlin 1981*b*), these transformations differ from the usual concept of a neural filter that picks out specific attributes of the incoming data. The type of filtering that enhances coding efficiency is not specific but generalized, for it retains the components that are richest in information at the expense of the poorer, irrespective of what the information actually represents. The observation that several diverse types of retinal neuron probably possess the essential elements required for predictive coding and redundancy reduction, suggests that coding efficiency is one of the fundamental parameters determining the operations of the peripheral visual system. However, the predictive coding hypothesis formulated here has covered but one aspect of coding efficiency, the representation of fine spatial and temporal detail. There could well be other applications. For example, movement has a strong influence on coding procedures and it has been suggested that this accounts for the different sampling procedures of X- and *Y*-cells in the cat retina (Hughes 1981). It has also been argued that suitably organized lateral inhibition can prefilter visual signals in such a way as to promote sensitive detection of movement (Srinivasan & Dvorak 1980). It would be interesting to see if such a division of labour were more effective than employing a single class of cells to cover all situations.

It would be surprising if the principles of efficient coding could not be applied beyond the confines of the visual system. The essence of efficiency as expounded here is to make the best use of available neurons by matching their input–output properties to the expected occurrence of signals. There must be many situations, such as the coordination of muscular activity, where the number of possible patterns of activity are severely constrained. It must be stressed, however, that while coding efficiency may be of general relevance it may not be universal. The nervous system not only gathers and packages information: it weights this information according to its biological significance.

This work stemmed from a series of informal group discussions, covering a variety of topics in vision, held in our laboratory. We are indebted to Jonathon Howard, who joined us in these discussions and drew our attention to the concept of predictive coding. We thank Allan Snyder, Stephen Shaw, Peter McIntyre and Richard Payne for reviewing preliminary versions of the manuscript. This study was enthusiastically encouraged by Adrian Horridge, F.R.S.

Appendix 1

This section describes in detail the computations made in deriving the inhibitory profiles.

We use the example of figure 6*b* to illustrate the procedure. In this example we have

$$\left.\begin{aligned} M \text{ (signal mean)} &= 1.0, \\ S \text{ (signal s.d.)} &= 0.03 \end{aligned}\right\} \quad \text{(A 1)}$$

and

$$N \text{ (noise s.d.)} = 0.03$$

and we have assumed that the spatial autocorrelation function of the scene is exponential, with a space constant D equal to five receptor widths.

The spatial autocorrelation function of the visual scene is then given by

$$R(x) = M^2 + S^2 e^{-|x|/D}, \quad \text{(A 2)}$$

where x is measured in receptor widths. The shape of this function is roughly as in the experimental measurements shown in figures 3 and 4. It has a maximum value of $M^2 + S^2$ at $x = 0$, and it decays exponentially to a value of M^2 as $|x| \to \infty$. (Note that $R(x)$ is not normalized here.)

The receptors are denoted by the numbers $-5, -4, -3, \ldots, 0, 1, 2, \ldots, 5$, as shown in figure 6. We denote by $R_{i,j}$ the correlation between the signals in receptors i and j; that is, $R_{i,j} = x_i x_j$. We then have

$$\left.\begin{aligned} R_{i,i} &= M^2 + S^2 + N^2 \quad (i = -5, -4, \ldots, 5) \\ \text{and} \quad R_{i,j} &= M^2 + S^2 e^{-|i-j|/D} \quad (i \neq j), \end{aligned}\right\} \quad \text{(A 3)}$$

where the N^2 term disappears for $i \neq j$ because the noise in the various receptors is assumed to be uncorrelated.

For our example, equation (1*b*) can be expressed in matrix form as:

$$\begin{bmatrix} R_{-5,-5} & R_{-5,-4} & R_{-5,-3} \ldots R_{-5,5} \\ R_{-4,-5} & R_{-4,-4} & R_{-4,-3} \ldots R_{-4,5} \\ R_{-3,-5} & R_{-3,-4} & R_{-3,-3} \ldots R_{-3,5} \\ \vdots & \vdots & \vdots \\ R_{5,-5} & R_{5,-4} & R_{5,-3} \ldots R_{5,5} \end{bmatrix} \begin{bmatrix} h_{-5} \\ h_{-4} \\ h_{-3} \\ \vdots \\ h_5 \end{bmatrix} = \begin{bmatrix} R_{0,-5} \\ R_{0,-4} \\ R_{0,-3} \\ \vdots \\ R_{0,5} \end{bmatrix}, \quad \text{(A 4)}$$

where $h_{-5}, h_{-4}, \ldots, h_5$ denote respectively the inhibitory weighting coefficients from receptors $-5, -4, \ldots, 5$.

The $R_{i,j}$ values are calculated from equations (A 1) and (A 3), and are inserted into the above matrix equation. Note that the R matrix is symmetrical and all of the diagonal elements are identical. This equation is solved (by matrix inversion on a computer) to obtain the following values for the inhibitory weighting coefficients: $h_{-5} = -0.041$, $h_{-4} = -0.047$, $h_{-3} = -0.070$, $h_{-2} = -0.121$, $h_{-1} = -0.221$, $h_1 = -0.221$, $h_2 = -0.121$, $h_3 = -0.070$, $h_4 = -0.047$ and $h_5 = -0.041$. The profile corresponding to these coefficients is shown in figure 6*b*.

For an exponentially decaying autocorrelation function, as in this example, it

can be shown that the R matrix is non-singular provided that D is finite. In our computations we have experienced no difficulty in inverting the R matrices derived from either exponential or gaussian autocorrelation functions; we have not examined other functions. For a general treatment of the properties of R matrices derived from autocorrelation functions, the reader is referred to Grenander & Szegö (1958).

Receptor noise affects only the diagonal terms of the $R_{i,j}$ matrix; increased noise causes these values to be larger (see equation (A 3)), and the h_i values that result from solving the equation are then smaller and distributed more uniformly over the receptive field.

The procedure for dealing with two-dimensional receptive fields is very similar to that described above. In a 7×7 receptor array, for example, the signal at the central receptor is predicted from the signals of the 48 surrounding receptors. While calculating the $R_{i,j}$ coefficients the geometry of the receptor array has to be taken into account, because this computation requires knowledge of the distance between the ith and the jth receptor.

The calculation for temporal encoding is similar to that for one-dimensional spatial encoding, except that the prediction is one-sided in this case (i.e. it is based only on past samples, not on future ones).

Appendix 2

Here we examine the effect of movement on spatial prediction. If one assumes that the receptor response can be modelled as a linear dynamical system that transduces light to receptor potential, the receptor response $v(t)$ evoked by a temporal pattern of optical stimulation $f(t)$ is given by (see, for example: Srinivasan & Bernard 1975; Howard 1981)

$$v(t) = \int_0^\infty f(t-\xi)\, h(\xi)\, \mathrm{d}\xi,$$

where $h(\xi)$ is the impulse response function of the photoreceptor, i.e. the response to a brief flash of light, and $f(t)$ and $v(t)$ are defined in terms of the deviations from their respective mean values. If $h(\xi)$ is approximated by a rectangular pulse with duration equal to the receptor's integration time Δt, the above equation can be rewritten as

$$v(t) \approx \int_0^{\Delta t} f(t-\xi)\, \mathrm{d}\xi.$$

Since, in our example on page 444, the distance moved in one integration time is three correlation lengths, the above integral can be viewed as the sum of three statistically independent samples of $f(t)$ taken in the interval $0 \leqslant \xi \leqslant \Delta t$:

$$v(t) \approx \tfrac{1}{3}\Delta t\, f(t-\tfrac{1}{6}\Delta t) + \tfrac{1}{3}\Delta t\, f(t-\tfrac{1}{2}\Delta t) + \tfrac{1}{3}\Delta t\, f(t-\tfrac{5}{6}\Delta t).$$

Let f denote the root mean-square (r.m.s.) amplitude of $f(t)$ and v denote the r.m.s. amplitude of $v(t)$. Since the three samples of $f(t)$ are statistically independent, we have $v \approx \frac{1}{3}\Delta t\sqrt{3} f = \frac{1}{\sqrt{3}}\Delta t\, f$. The r.m.s. contribution of each sample to the receptor response is $\frac{1}{3}\Delta t\, f$, which is $\frac{1}{\sqrt{3}}$ of the r.m.s. value of the response.

References

Attneave, F. 1954 Informational aspects of visual processing. *Psychol. Rev.* **61**, 183–193.

Barlow, H. B. 1961*a* Possible principles underlying the transformation of sensory messages. In *Sensory communication* (ed. W. A. Rosenblith), pp. 217–234. Cambridge, Massachusetts: M.I.T. Press.

Barlow, H. B. 1961*b* Three points about lateral inhibition. In *Sensory communication* (ed. W. A. Rosenblith), pp. 782–786. Cambridge, Massachusetts: M.I.T. Press.

Barlow, H. B. 1961*c* The coding of sensory messages. In *Current problems in animal behaviour* (ed. W. H. Thorpe & O. L. Zangwill), pp. 331–360. Cambridge University Press.

Barlow, H. B. 1981 The Ferrier Lecture 1980. Critical limiting factors in the design of the eye and the visual cortex. *Proc. R. Soc. Lond.* B **212**, 1–34.

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

Barlow, H. B. & Levick, W. R. 1976 Threshold setting by the surround of cat retinal ganglion cells. *J. Physiol., Lond.* **259**, 737–757.

Braitenberg, V. 1967 Patterns of projection in visual system of the fly. I. Retina–lamina projections. *Expl Brain Res.* **3**, 271–298.

Brodie, S. E., Knight, B. W. & Ratliff, F. 1978 The spatiotemporal transfer function of the *Limulus* lateral eye. *J. gen. Physiol.* **72**, 167–202.

Campbell, F. W. & Gubisch, R. W. 1966 Optical quality of the human eye. *J. Physiol., Lond.* **186**, 558–578.

Campbell, F. W. & Robson, J. G. 1968 Application of Fourier analysis to the visibility of gratings. *J. Physiol., Lond.* **197**, 551–566.

Collett, T. & Land, M. F. 1975 Visual control of flight behaviour in the hoverfly, *Syritta pipiens* L. *J. comp. Physiol.* **99**, 1–66.

DeValois, R. L., Morgan, H. & Snodderly, D. M. 1974 Psycophysical studies of monkey vision. III. Spatial luminance contrast sensitivity tests of macaque and human observers. *Vision Res.* **14**, 75–81.

Dubs, A. 1982 The spatial integration of signals in the retina and lamina of the fly compound eye under different conditions of luminance. *J. comp. Physiol.* **146**, 321–343.

Dubs, A., Laughlin, S. B. & Srinivasan, M. V. 1981 Single photon signals in fly photoreceptors and first order interneurons at behavioural threshold. *J. Physiol., Lond.* **317**, 317–334.

Enroth-Cugell, C. & Robson, J. G. 1966 The contrast sensitivity of retinal ganglion cells of the cat. *J. Physiol., Lond.* **187**, 517–552.

Enroth-Cugell, C. & Shapley, R. 1973 Adaptation and dynamics of cat retinal ganglion cells. *J. Physiol., Lond.* **233**, 271–309.

French, A. S. & Järvilehto, M. 1978 The transmission of information by first and second order neurons in the fly visual system. *J. comp. Physiol.* **126**, 87–96.

Georgeson, M. A. & Sullivan, G. D. 1975 Contrast constancy: deblurring in human vision by spatial frequency channels. *J. Physiol., Lond.* **252**, 627–656.

Gonzalez, R. C. & Wintz, P. 1977 *Digital image processing.* London: Addison-Wesley.

Grenander, U. & Szegö, G. 1958 *Toeplitz forms and their application.* Berkeley and Los Angeles: University of California Press.

Hardie, R. C. 1979 Electrophysiological analysis of fly retina. I. Comparative properties of R1–6 and R7 and 8. *J. comp. Physiol.* **129**, 19–33.

Harrison, C. W. 1952 Experiments with linear prediction in television. *Bell Syst. tech. J.* **31**, 764–783.

Hartline, H. K. & Ratliff, F. 1972 Inhibitory interactions in the retina of *Limulus*. In *Handbook of sensory physiology*, vol. 7 (2) (ed. M. G. F. Fuortes), pp. 381–447. Berlin: Springer.

Howard, J. 1981 The temporal resolving power of the photoreceptors of *Locusta migratoria*. *J. comp. Physiol.* **144**, 61–66.

Howard, J., Dubs, A. & Payne, R. 1982 The dynamics of phototransduction in insects – a comparative study. *J. comp. Physiol.* (Submitted.)

Hughes, A. 1977 The topography of vision in mammals of contrasting life style: comparative optics and retinal organisation. In *Handbook of sensory physiology*, vol. 7 (5) (ed. F. Crescitelli), pp. 613–756. Berlin: Springer.

Hughes, A. 1981 Cat retina and the sampling theorem: the relation of transient and sustained brisk-unit cut-off frequency to α and β mode cell density. *Expl Brain Res.* **42**, 196–202.

Kelly, D. H. 1961 Visual responses to time-dependent stimuli. 1. amplitude sensitivity measurements. *J. opt. Soc. Am.* **51**, 422–429.

Kirschfeld, K. 1967 Die Projektion der optischen Umwelt auf das Raster des Rhabdomere im Komplexauge von *Musca*. *Expl Brain Res.* **3**, 248–270.

Kretzmer, E. R. 1952 Statistics of television signals. *Bell Syst. tech. J.* **31**, 751–763.

Kuffler, S. W. 1953 Discharge patterns and functional organisation of mammalian retina. *J. Neurophysiol.* **16**, 37–68.

Laughlin, S. B. 1981*a* Neural principles in the peripheral visual systems of invertebrates. In *Handbook of sensory physiology*, vol. 7 (6B) (ed. H. Autrum), pp. 133–280. Berlin: Springer.

Laughlin, S. B. 1981*b* A simple coding procedure enhances a neuron's information capacity. *Z. Naturf.* **36**c, 910–912.

Laughlin, S. B. & Hardie, R. C. 1978 Common strategies for light adaptation in the peripheral visual systems of fly and dragonfly. *J. comp. Physiol.* **128**, 319–340.

Leutscher-Hazelhoff, J. T. 1975 Linear and non-linear performance of transducer and pupil in *Calliphora* retinula cells. *J. Physiol., Lond.* **246**, 333–350.

Mach, E. 1865 Über die Wirkung der raumlichen Vertheileng des Lichtreizes auf die Netzhaut. *Sber. Akad. Wiss. Wien* (II) **52**, 303–322.

Mackay, D. M 1956 Towards an information-flow model of human behaviour. *Br. J. Psychol.* **47**, 30–43.

Marcelja, S. 1979 Initial processing of visual information within the retina and the LGN. *Biol. Cybernetics* **32**, 217–226.

Marr, D. & Hildreth, E. 1980 Theory of edge detection. *Proc. R. Soc. Lond.* B **207**, 187–218.

Miller, W. H. 1979 Ocular optical filtering. In *Handbook of sensory physiology*, vol. 7 (6A) (ed. H. Autrum), pp. 69–143. Berlin: Springer.

Mimura, K. 1976 Some spatial properties in the first optic ganglion of the fly. *J. comp. Physiol.* **105**, 65–82.

Norman, R. A. & Werblin, F. S. 1974 Control of retinal sensitivity. 1. Light and dark-adaptation of vertebrate rods and cones. *J. gen. Physiol.* **63**, 37–61.

Oliver, B. M. 1952 Efficient coding. *Bell Syst. tech. J.* **31**, 724–750.

Papoulis, A. 1965 *Probability, random variables, and stochastic processes.* Tokyo: McGraw-Hill Kogakusha.

Purple, R. L. & Dodge, F. A. 1965 Interaction of excitation and inhibition in the eccentric cell in the eye of *Limulus*. *Cold Spring Harb. Symp. quant. Biol.* **30**, 529–537.

Ratliff, F. 1965 *Mach bands: quantitative studies on neural networks in the retina.* San Francisco: Holden-Day.

Robson, J. G. 1966 Spatial and temporal contrast-sensitivity functions of the visual system. *J. opt. Soc. Am.* **56**, 1141–1142.

Rodieck, R. W. & Stone, J. 1965 Response of cat retinal ganglion cells to moving visual patterns. *J. Neurophysiol.* **28**, 819–832.

Rosenfeld, A. & Kak, A. C. 1976 *Digital picture processing.* New York: Academic Press.

Shaw, S. 1978 Signal transmission by graded slow potentials in the arthropod peripheral visual system. In *The neurosciences: fourth study program* (ed. F. O. Schmitt & F. G. Worden), pp. 275–295. Cambridge, Massachusetts: M.I.T. Press.

Snyder, A. W. 1979 Physics of vision in compound eyes. In *Handbook of sensory physiology*, vol. 7 (6A) (ed. H. Autrum), pp. 225–313. Berlin: Springer.

Snyder, A. W. & Srinivasan, M. V. 1979 Human psychophysics: functional interpretation for contrast sensitivity versus spatial frequency curve. *Biol. Cybernetics* **32**, 9–17.

Srinivasan, M. V. & Bernard, G. D. 1975 The effect of motion on visual acuity of the compound eye. *Vision Res.* **15**, 515–525.

Srinivasan, M. V. & Dvorak, D. R. 1980 Spatial processing of visual information in the movement-detecting pathway of the fly. *J. comp. Physiol.* **140**, 1–23.

Wehner, R. 1981 Spatial vision in arthropods. In *Handbook of sensory physiology*, vol. 7 (6C) (ed. H. Autrum), pp. 287–616. Berlin and New York: Springer.

Werblin, F. S. 1971 Adaptation in a vertebrate retina: intracellular recording in *Necturus*. *J. Neurophysiol.* **34**, 228–241.

Werblin, F. S. 1974 Control of retinal sensitivity. II. Lateral interactions at the outer plexiform layer. *J. gen. Physiol.* **63**, 62–87.

Wilson, H. R. & Bergen, J. R. 1979 A four mechanism model for spatial vision. *Vision Res.* **19**, 19–32.

Zettler, F. 1969 Die Abhängigkeit des Übertragungsverhaltens von Frequenz und Adaptationszustand, gemessen am einzelnen Lichtrezeptor von *Calliphora erythrocephala*. *Z. vergl. Physiol.* **64**, 432–449.

Zettler, F. & Järvilehto, M. 1972 Lateral inhibition in an insect eye. *Z. vergl. Physiol.* **76**, 233–244.