

Nonlinearity in the perception of form

ROBERT SHAPLEY

Rockefeller University, New York, New York

and

JAMES GORDON

Hunter College of CUNY, New York, New York

The human visual system seems to determine the apparent brightness of an object mainly by measuring the physical contrast near the border of the object and its background (see, e.g., Cornsweet, 1970; Heinemann, 1955; Land & McCann, 1971; Ratliff, 1965; Whittle & Chalmers, 1969). How one perceives the shape or form of an object is less well understood, but again neural signals derived from contrast near the contour have been hypothesized to be important (Marr, 1981). Contemporary theoretical assumptions about the mechanisms of form perception include: linear spatial filtering into an array of parallel spatial channels, each composed of a chain of visual neurons; the association of channel outputs into groups in an as yet unspecified manner; comparison with internally stored filtered images; and then a stage of decision making for categorization (Ginsburg, 1982; Graham, 1980; Marr, 1981; Robson, 1980).

We have discovered several spatial patterns which reveal an important property of a contour-sensing visual mechanism that is used to perceive form (Shapley & Gordon, 1983). The contour mechanism sums visual signals in an extremely nonlinear manner; it must respond to the *absolute value or some other even nonlinear function* of contrast, rather than to contrast itself. This finding may lead to a new understanding of the way in which the visual system segregates objects from their backgrounds.

Visual patterns were created on the face of a cathode ray tube monitor (Tektronix 608) with an electronic visual display instrument (Milkman et al., 1980) under the control of a microcomputer (PDP-11/23). The instrument produced a raster display. Each frame of the raster consisted of 256 lines with 256 picture elements per line. The frame rate was 270 Hz. The visual stimulator produced four spatial profiles, which could be mapped into each of the $256 \times 256 (=65,536)$ picture elements on the raster display. The mapping was controlled by the values in a $65,536 \times 2$ bit memory, the values of which were read

We thank Norman Milkman, Gary Schick, and Michelangelo Rossetto for the instrument. This work was supported by grants from the U.S. National Eye Institute (EY 1472, EY 188, and EY 2439) and also by a Research Career Development Award from the Eye Institute to R.S. Computer time was provided by the C.U.N.Y. University Computing Center.

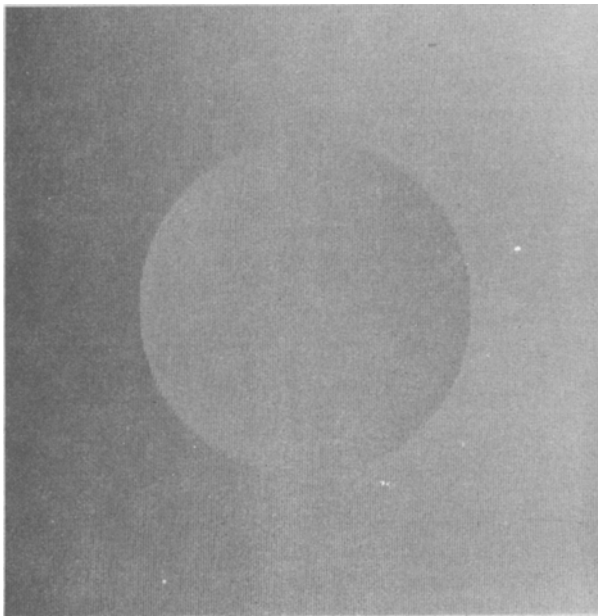
R. Shapley's mailing address is: Rockefeller University, New York, NY 10021. J. Gordon's is: Department of Psychology, Hunter College of C.U.N.Y., New York, NY 10021.

into a high-speed electronic four-position switch at a rate of 20 MHz. The contrast¹ on the monitor screen was controlled by depth-of-modulation values fed to the instrument by the computer. The precision of the depth of modulation was 12 bits. The value of the spatial profile was multiplied in a multiplying D/A converter by the depth-of-modulation, for each profile independently. Thus, for example, spatial profile 1 could be kept at 0 depth of modulation, the mean luminance, while spatial profile 2 could be modulated at a Rayleigh contrast of 0.66. The 608 monitor had a white P4 phosphor. Mean luminance was kept at 100 cd/m² in all experiments. Viewing was binocular. The 10 × 10 cm raster was usually viewed at a distance of 1 m, except for the distance experiment described below.

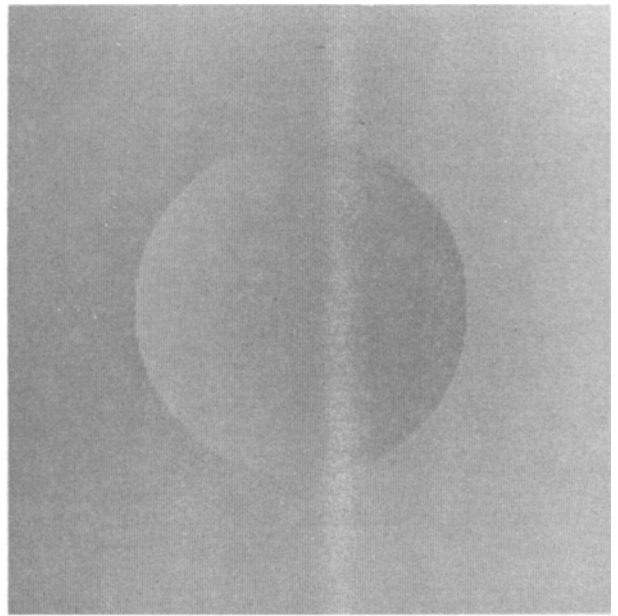
One way to probe the linearity of the perception of form is to create patterns in which the signature of the local contrast (see footnote 1) varies around the boundary of an object on a background. If the contour-sensing mechanisms in our visual system combined visual neural signals in a linear manner, a change in signature of the border contrast should lead to a change in the sign of response of the neural units(s) that sense the contour. The linear combination of positive and negative responses would cause cancellation of signals from neighboring regions of positive and negative contrast. Such cancellation would lead to difficulties in perceiving the border of a sign-reversing object in the region in which the local contrast changes sign.

Figure 1 consists of two patterns in which the local contrast around the border changes sign, together with graphs of the luminance profiles. Each pattern in Figure 1 consists of a circular object on a background. In both panels of the figure, the background is a uniformly increasing gradient of luminance. Thus, the background luminance, as a function of position, is $L_B(x) = L_0(1+mx)$. L_0 is the mean luminance of the background; the slope m is the peak luminance minus the mean luminance divided by the half-width of the screen; and x is the distance along the horizontal from the midpoint of the screen. In Figure 1A, the luminance of the circle is constant at the mean; $L_c = L_0$. In Figure 1B, the circle's luminance is a gradient, but with a slope opposite to that of the background; thus, its luminance is $L_c = L_0(1 - mx)$. In order to achieve equal local contrast around the borders of the circles in Figures 1A and 1B, the slope of the background gradient in 1B is half the value in 1A. Thus, the value of m is not the same in both (see caption for Figure 1). Note also that, for both patterns, the object and background luminances are nearly equal near the middle vertical line of the screen, where they approach L_0 in luminance ($x=0$). The reader may observe this by covering two sides of the figure with opaque cards so that only the central eighth of the figure is visible. No border is seen under these conditions.

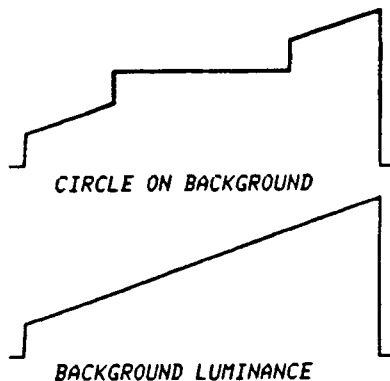
(A)



(B)



(C)



(D)

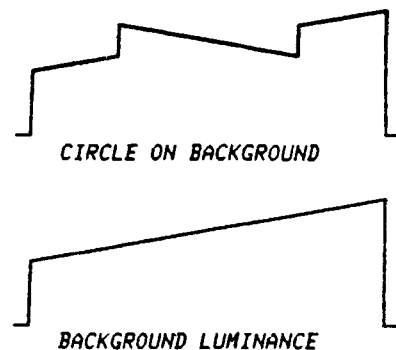


Figure 1. (A and B) Circular figures at whose borders the contrast changes uniformly. The background luminance is a one-dimensional linear gradient. The contrast of the linear gradient was 0.66 in A and 0.33 in B, where contrast is defined as $(L_{\max} - L_{\min}) / (L_{\max} + L_{\min})$. In A, the luminance of the central circular region was constant at $L_{\text{mean}} = (L_{\max} + L_{\min}) / 2$. In B, it was a linear gradient of opposite sign to the background. The extent of the circles was such that the maximum local contrast across the border was 0.5. (C and D) Shown here are the luminance profiles of a horizontal line that cuts through just the background, and also along a line that passes through the center of the circular test patch. The vertical coordinate is relative luminance and the horizontal is distance on the target screen. Panel C displays background and test profiles for pattern A, and panel D shows the profiles for pattern B. Note that the contrast changes from inside positive on the left to inside negative on the right side of the screen. The shaded appearance of the interior of the circle must be due to border contrast, as indicated by the similarity in appearance of the two panels of the figure. The “apparent” contour on the top and bottom of the circles is present when the test target is small enough, even though, at these points, there is no luminance difference across the apparent border.

There are several interesting aspects of these patterns with respect to brightness and form. Subjects often report that the central object has a shaded three-dimensional appearance, whether it actually is nonuniform in luminance (right panel) or of uniform luminance but is surrounded by a border that varies in contrast (left panel). More significant in the present context is the fact that both objects appear to be circular and unitary. Moreover, the unbroken appearance of the contour in the neighborhood of $x=0$, at the top and bottom of the circular object, is direct evidence that the contour-sensing mechanism disregards the

sign of the contrast, because it interpolates a perceptual border right across a region in which there is no physical border, between two regions of opposite border contrast. The distance across which the border is interpolated is of the order of 1° of visual angle (see below). This suggests the surprising and significant idea that neural signals from opposite sides of the object are added together by the nervous system without regard to their signature.

The spatial extent of the border interpolation was estimated by means of the following experiment. Each of four observers viewed the pattern of Figure 1A at different dis-

tances from the stimulus screen. The distances from eye to screen ranged from 20 m down to 5 cm. The slope m of the gradient of illumination was 0.66/half screen. Therefore, the peak magnitude of the local Weber contrast, $(L_{\text{object}} - L_{\text{background}})/L_{\text{background}}$, around the border of the circular object was 0.5, located where the border of the circle intersects its horizontal diameter. The local contrast was -0.25 on the right side of the circle and $+0.5$ on the left. The local Weber contrast went to 0 along the vertical diameter. The appearance of the pattern was an unbroken circle on a background for viewing distances greater than 20 cm on average across the four subjects, that is, when the 5-cm diameter of the circle subtended 14° of visual angle or less. The standard deviation of the distance for perceptual breakdown of the interpolated contour, across subjects, was 4 cm. When the circle subtended more than 14° , the interpolated perceptual contour linking right and left half-circles was not visible, and a blank region filled the middle of the monitor's screen. This procedure amounted to a method of adjustment for distance until perceptual breakdown of the interpolated contour.

We need to estimate the contrast threshold for the border of the object in order to infer the spatial extent of interpolation. In separate measurements with two observers, using method of adjustment of contrast, we found that the Weber contrast threshold for a circular object (of the same size as our targets) on a uniform background was 0.02. Therefore, as a first approximation, we assumed that when the local contrast of the border of the object on the gradient background was below 0.02, the border would be invisible in that local region. This is meant to be an approximation only, because the threshold contrast for the border of varying local contrast, as in Figure 1A, is probably not the same as for the uniformly luminous circles on a uniform background. Nevertheless, we believe this approximation is probably useful in estimating the approximate extent of border interpolation.

In the distance experiment reported above, when perceptual breakdown occurred, the length of contour along which the local contrast was below 0.02, the limit for detection, was approximately 1.75° . When the length of the interpolated contour was less than 1.75° , the circle appeared whole. This measurement is a rough estimate of the integration length of the contour-sensing mechanism. In the distance experiment, the direction of gaze was not controlled, but all subjects fixated the region of the interpolated border in order to achieve their best performance. Therefore, the large value of the integration length is remarkable, since it is larger than the diameter of the fovea.

A comparison between objects in which contrast is of the same sign around the boundary and in which it changes sign around the boundary is presented in Figure 2. In this picture, all the circles have a physical luminance of L_0 , and the background, as before, is a ramp $L_B = L_0(1 + mx)$ with $m = (0.66/\text{half-screen})$. One notices the very striking effect of contrast on the apparent brightness of the

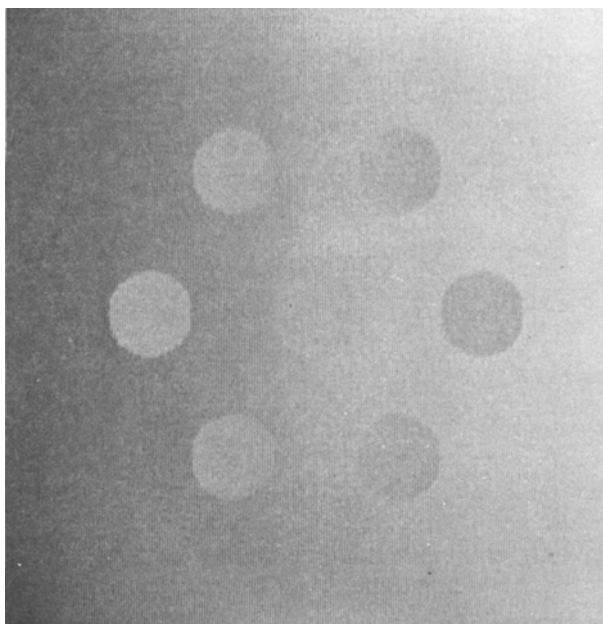


Figure 2. A rosette of circles of constant luminance on a background that is a linear gradient of luminance. The background is again a ramp of luminance of contrast 0.66. The circles are all of constant luminance equal to the mean level of illumination. The apparent differences in the brightness of the circles are results of simultaneous contrast at the borders of the circles with the background. Note that the indistinct central target appears to be roughly the same shape as the surrounding circles of much higher apparent contrast. Note also that for the central circle the contrast changes sign from left to right of a vertical diameter of the circle, whereas the signature of contrast around the border of the other six circles does not change. Yet the first impression is a similarity of shape.

outer circles (see Figure 2 caption). However, it is also striking that the inner circle, although more indistinct because of a lower magnitude of contrast, is perceived as having the same shape as the outer circles. All the outer circles have borders around which the contrast is the same sign. The inner circle, like the circles in Figure 1, has a border of positive contrast to the left of the vertical diameter and negative contrast to the right; interpolated perceptual contours link the physical contours. This indicates a fundamental disregard of the sign of the contrast in sensing the border for pattern recognition.

Another pair of patterns that illustrate the unimportance of the sign of the contrast is shown in Figure 3. These are squares with illusory edges based on the same idea as the Kanizsa triangle (Ginsburg, 1975; Kanizsa, 1955, 1979; Marr, 1981). The squares seem to stand out in front of the broken circular figures at the corners. However, in one case the corner figures alternate in contrast, whereas in the other case, as in the Kanizsa triangle, they are all darker than the background. The extrapolated perceptual contours appear equally visible in the two cases. Like the Kanizsa triangle, the central square is of the same physical luminance as the background and the apparent

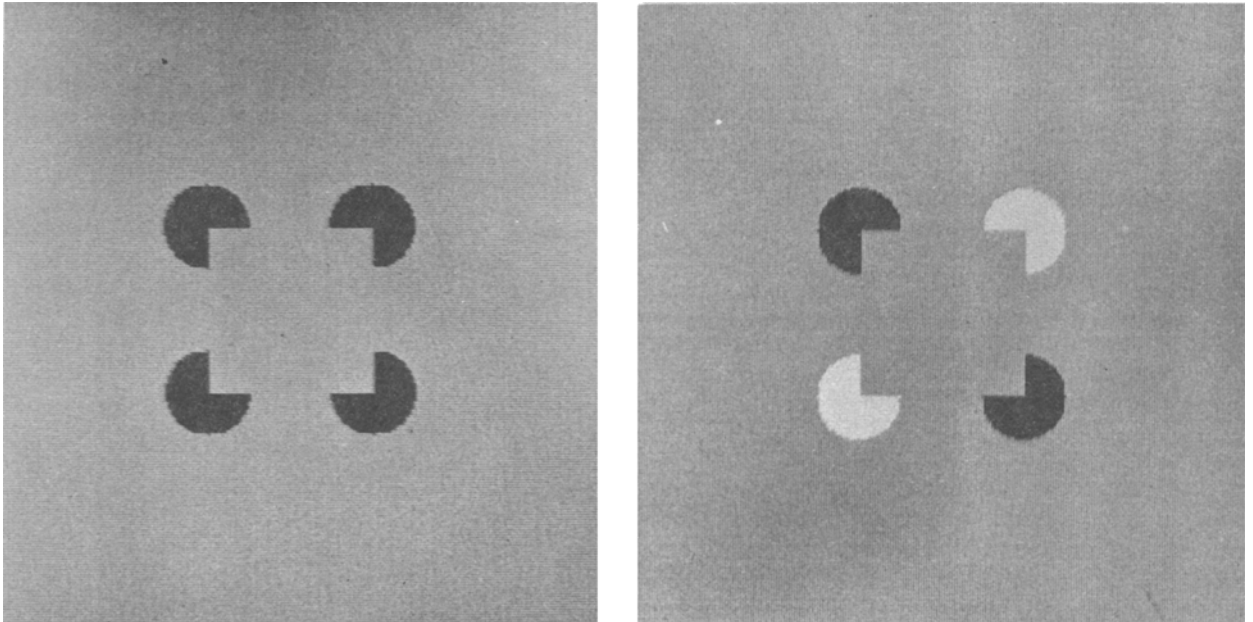


Figure 3. An "illusory" square produced by adjacent contours. Two examples are offered. One (left), like the Kanizsa triangle, has inducing figures at the corners, all of which are darker than the background. On the right, an equally apparent illusory square is produced by inducing figures that alternate in contrast. The central square region and the background have the same luminance, which can be seen by holding the figure close enough to the observer so that the square subtends more than 2° visual angle.

contours between square and background are illusory. One might have thought that a linear spatial filtering mechanism created a neural correlate of these illusory contours by low-pass linear spatial filtering (cf. Frisby, 1980; Ginsburg, 1975), but the evidence of Figure 3 rules this out. Any sort of linear filtering would cause cancellation of the neural signals induced by the opposite contrasts.

Interpolation of contours in Figure 3 also breaks down if the suprathreshold inducing borders are separated too widely in space. If the vertices of the square are separated by more than 2° , the interpolated contours are much less visible. Such an estimate of the distance over which perceptual interpolation fails is quite consistent with our measurements on the circular figures of Figure 1. Perceptual breakdown of the Kanizsa "illusion" can be seen quite clearly in standard demonstrations of the Kanizsa triangle (Kanizsa, 1955, 1979) by varying viewing distance.

One hypothesis that may provide a unified explanation of the results presented here, and other phenomena of form perception, is as follows. Suppose that the borders of objects are sensed by nonlinear spatial channels that respond to the absolute value or some other even nonlinear function of the contrast at the border. An even nonlinear function of contrast is, by definition, insensitive to the sign of the contrast. Extrapolation of real contour into illusory contour, as in Figures 1-3, would be a consequence of the extent of the spatial receptive fields of these nonlinear border sensors and their overlaps. A diagram of what such a receptive field might be like is offered as Figure 4.

Another explanation for interpolated contours is quite different from the one we have proposed here. It is the idea that interpolated contours are due to "cognitive" processes (Gregory, 1972). Our results cast doubt on the "cognitive" hypothesis in two ways. First, the unitary appearance of the circles in Figures 1 and 2 is perceived within a fraction of a second. This is a much faster time frame than most "cognitive" processes. The second problem for the "cognitive" hypothesis is the rather short distance over which the contours may be interpolated. Although we remarked on how long this distance seemed in terms of the size of the fovea, it is a very short distance when compared with the spatial extent of the mind's image of the world. If the interpolated contour were really the mind's best solution to the appearance of the visual objects in Figures 1-3, there is no reason why greater separation would cause the interpolation to fail, as it most definitely does. The nonlinear contour-sensor that we have proposed in Figure 4 has a finite interpolation length built into it.

Recently, Prazdny (1983) has presented a diamond-shaped figure of the same general type as our Figure 3, a Kanizsa-like figure with objects of opposite contrast at the vertices. He uses the existence of such figures as evidence that mechanisms like those that produce simultaneous contrast cannot be the explanation of the Kanizsa triangle. Thus, he refutes Frisby (1980) and others who have interpreted the borders of the Kanizsa triangle as having been induced by such mechanisms. Note that Ginsburg's (1975) explanation of the appearance of Kanizsa-like figures in terms of spatial filtering is qualitatively equiva-

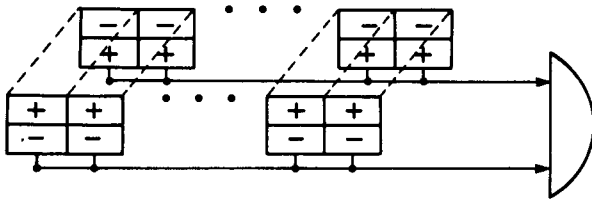


Figure 4. Diagram of a nonlinear border-detecting mechanism. It is supposed that signals from local contrast-sensitive neurons of one polarity, indicated by the small + and - receptive fields, are summed in a nonlinear manner with signals from local contrast neurons of the opposite sign, indicated by the - and + signs, respectively. The fields of the border detectors are elongated compared with the shape of the local contrast-sensitive neurons because of (non-linear) summation along the length of the receptive field. The non-linearity of summation—indicated by a circular sector symbol in the diagram—may be an absolute value or any other nonlinearity of even order.

lent to the simultaneous contrast mechanism explanation of Frisby (1980); they both predict the sensation of a border where none exists in the image because of (linear) neural blur. We agree with Prazdny's (1983) argument that the existence of interpolated contours in Figure 3 is evidence against the linear spatial filtering and the equivalent simultaneous contrast hypotheses. However, Prazdny goes further and asserts that his analogue to our Figure 3 can only be explained in terms of "cognitive" mechanisms. We have offered here an alternative to the "cognitive" explanation, namely the nonlinear spatial filter sketched in Figure 4. Furthermore, we find the *failure* of the interpolated contours to complete the figures, when the distance between inducing contours is too great, to be compelling evidence against a strictly "cognitive" explanation.

The data on the finite range of the interpolation mechanism may be related to the spatial extent of the "Gestalt-like grouping operations" postulated by Marr (1981). These were required in his scheme to unify the elements of the "primal sketch," the earliest stage in his proposed computational theory of vision. The nonlinear border mask we propose in Figure 4 may produce many of the desirable effects of "Gestalt-like grouping" in a mechanistic fashion.

REFERENCES

- CORNISWEET, T. (1970). *Visual perception*. New York: Academic Press.
- FRISBY, J. P. (1980). *Seeing*. Oxford: Oxford University Press.
- GINSBURG, A. P. (1975). Is the illusory triangle physical or imaginary? *Nature*, **257**, 219-220.
- GINSBURG, A. P. (1982). On a filter approach to understanding the perception of visual form. In D. G. Albrecht (Ed.), *Recognition of pattern and form* (pp. 175-192). New York: Springer.
- GRAHAM, N. (1980). Spatial frequency channels in human vision: Detecting edges without edge detectors. In C. Harris (Ed.), *Visual coding and adaptability*. Hillsdale, NJ: Erlbaum.
- GREGORY, R. L. (1972). Cognitive contours. *Nature*, **238**, 51-52.
- HEINEMANN, E. G. (1955). Simultaneous brightness induction as a function of inducing- and test-field luminances. *Journal of Experimental Psychology*, **50**, 89-96.
- KANIZSA, G. (1955). Margini quasi-percettivi in campi con stimolazione omogenea. *Rivista di Psicologia*, **49**, 7-30.
- KANIZSA, G. (1979). *Organization in vision*. New York: Praeger.
- LAND, E., & McCANN, J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America*, **61**, 1-11.
- MARR, D. (1981). *Vision*. San Francisco: Freeman.
- MILKMAN, N., SCHICK, G., ROSSETTO, M., RATLIFF, F., SHAPLEY, R., & VICTOR, J. (1980). A two-dimensional computer-controlled visual stimulator. *Behavior Research Methods & Instrumentation*, **12**, 283-292.
- PRAZDNY, K. (1983). Illusory contours are not caused by simultaneous brightness contrast. *Perception & Psychophysics*, **34**, 403-404.
- RATLIFF, F. (1965). *Mach bands*. San Francisco: Holden-Day.
- ROBSON, J. G. (1980). Neural images: The physiological basis of spatial vision. In C. Harris (Ed.), *Visual coding and adaptability*. Hillsdale, NJ: Erlbaum.
- SHAPLEY, R., & ENROTH-CUGELL, C. (1984). Visual adaptation and retinal gain controls. In N. N. Osborne & G. J. Chader (Eds.), *Progress in retinal research* (Vol. 3, pp. 263-346). Oxford: Pergamon.
- SHAPLEY, R., & GORDON, J. (1983). A nonlinear mechanism for the perception of form. *Investigative Ophthalmology and Visual Science*, **24** (Suppl.), 238.
- WHITTLE, P., & CHALLANDS, P. D. C. (1969). The effect of background luminance on the brightness of flashes. *Vision Research*, **9**, 1095-1110.

NOTE

1. There are two generally accepted definitions of physical *contrast*. For aperiodic objects, the *Weber contrast* is $(L_{\text{object}} - L_{\text{background}}) / L_{\text{background}}$. For periodic spatial patterns, the *Rayleigh contrast* is $(L_{\text{max}} - L_{\text{min}}) / (L_{\text{max}} + L_{\text{min}})$. These names for the different kinds of contrast were introduced by one of us in a review article (Shapley & Enroth-Cugell, 1984). When we use the term "local contrast," we mean the local value of the Weber contrast.

(Manuscript received August 24, 1984;
accepted for publication December 24, 1984.)