Reprinted from

JOURNAL
of the
OPTICAL SOCIETY
of
AMERICA

Smallest channel in early human vision

D. Marr
MIT Psychology Department and A. I. Laboratory, 79 Amherst Street, Cambridge, Massachusetts 02139

T. Poggio
Max-Planck-Institut für biologische Kybernetik, Spemannstrasse 38, D 7400 Tubingen, Federal Republic of Germany

E. Hildreth
MIT A. I. Laboratory, 545 Technology Square, Cambridge, Massachusetts 02139

pp. 868–870

JULY
1980

VOLUME 70 NUMBER 7

Smallest channel in early human vision

D. Marr
MIT Psychology Department and A. I. Laboratory, 79 Amherst Street, Cambridge, Massachusetts 02139
T. Poggio
Max-Planck-Institut für biologische Kybernetik, Spemannstrasse 38, D 7400 Tubingen, Federal Republic of Germany
E. Hildreth
MIT A. I. Laboratory, 545 Technology Square, Cambridge, Massachusetts 02139

(Received 23 October 1979)

We suggest from the psychophysical data on two-point and line acuity that the smallest foveal channel in human vision must have an excitatory center with a diameter of around 1′ 20″. Taking into account the optics of the eye and the finite size of the receptors, we show that this would correspond well to the properties of a retinal ganglion cell receiving excitatory input from a single cone.

The idea that there may be a range of different size or spatial-frequency-tuned mechanisms in human vision was originally introduced on the basis of psychophysical evidence by Campbell and Robson.[1] This led to a series of papers dealing with spatial frequency analysis in the visual system. Recently, Wilson and Giese[2] and Wilson and Bergen[3], (see also Ref. 4) integrated these and other anatomical and physiological data into a coherent logical framework. The key to their framework is (a) the partitioning of the range of sizes associated with the channels into two components, one due to spatial inhomogeneity of the retina, and one due to local scatter of respective field sizes, and (b) the correlation of these two components with anatomical and physiological data about the scatter of receptive field sizes and their dependence on eccentricity.

On the basis of detection studies, Wilson and Bergen proposed a specific four-channel model with the following characteristics: (1) At each position in the visual field, there exist four size-tuned filters or masks, the smaller two (called the N and S channels) showing relatively sustained temporal responses, and the larger two (called T and U) being relatively transient. (2) The half-power bandwidth of the N and S channels is about 1.3 octaves, but may be slightly larger for the T and U channels. (3) Wilson and Bergen described the receptive field shapes of these channels as the difference of two Gaussian distributions. (4) In the fovea, and using line stimuli, the widths w of the central excitatory regions of these receptive fields have the following values: N channel, 3.1′; S channel, 6.2′; T channel, 11.7′; U channel, 21′. The S channel is the most sensitive under both sustained and transient stimulation, and the U channel is the least, having only 1/4 to 1/11 the sensitivity of the S channel. (5) The receptive field sizes increase linearly with eccentricity, being in this case about double at 4° eccentricity.

Essentially all of the psychophysical data on the detection of spatial patterns below 16 cpd at contrast threshold can be explained by this model, together with the hypothesis that the detection process is based on a form of probability summation.

Our current theory of these channels is that they are the first step in the detection of intensity changes in the image.[5–7] The critical idea is that the first operation performed by the sustained channels can be regarded as (nonoriented) spatial second-derivative filters acting on the image at two scales. Sharp intensity changes correspond to zero-crossings (fast transitions from positive to negative values or vice versa) in the filtered output from these channels. These zero-crossings may be represented (by cortical cells) as oriented zero-crossing segments. The physiological detection of zero-crossings can be performed by detecting nearby on-and-off cells activity and need not depend on the detection of cells with zero response.[6,7] It has been shown: (a) that the optimal differential filter has a cross section very similar to that found by Wilson and Bergen;[7] and (b) that the zero-crossings together provide rich information about the image.[6]

Although Wilson and Bergen's experiments were carried out with oriented line stimuli, they provide no evidence that the first spatial-frequency filtering stage involves oriented receptive fields. In fact there are reasons to believe that the initial filters are not oriented, and that orientation sensitivity is introduced only at the subsequent stage where the zero crossings are detected and represented.[7] If this is true, the values of w measured by Wilson and Bergen must be multiplied by $\sqrt{2}$ to obtain the diameter of the corresponding circularly symmetric center-surround receptive field. To see this, we write the channel as

$$G(r) = \frac{1}{\sigma_e^2} e^{-r^2/2\sigma_e^2} - \frac{1}{\sigma_i^2} e^{-r^2/2\sigma_i^2}, \tag{1}$$

where σ_e and σ_i are, respectively, the space constants for the excitatory and inhibitory distributions. Write $\sigma_i/\sigma_e = \gamma$: then in two dimensions, the diameter w_{2D} of the central (excitatory) portion of the receptive field is given by

$$w_{2D} = 2\sqrt{\left(\frac{\gamma^2}{\gamma^2 - 1} 2\ln\gamma^2\right)}\,\sigma_e. \tag{2}$$

The response of such a channel to a line stimulus is given by

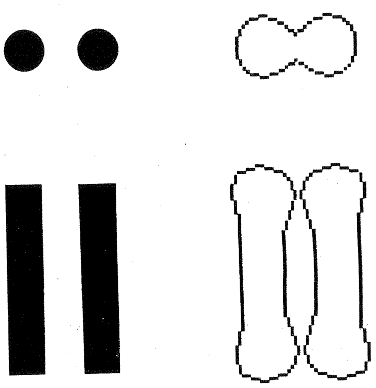
$$\int_{-\infty}^{\infty} G(r)dy,$$

which varies with

$$\left(\frac{1}{\sigma_e} e^{-x^2/2\sigma_e^2} - \frac{1}{\sigma_i} e^{-x^2/2\sigma_i^2}\right).$$

For such a distribution, the width w_{1D} of the central excitatory region is

$$w_{1D} = 2\sqrt{\left(\frac{\gamma^2}{\gamma^2 - 1} 2\ln\gamma\right)}\,\sigma_e = \frac{1}{\sqrt{2}} w_{2D}. \tag{3}$$



(a) (b)

FIG. 1. (a) Zero-crossings of the pattern on the left, filtered through a circularly symmetric receptive field (the difference of two Gaussians of equal area) with a central diameter of 2.1′. (b) Zero-crossings associated with a two-bar pattern filtered through the same receptive field. The angular gap in both cases is 1′. A slightly larger receptive field or a smaller separation leads to a zero-crossing profile practically indistinguishable from the zero-crossings of a one-bar pattern. Thus the criterion of separation of the zero-crossing profiles in the filtered output implies that line acuity is about half the diameter of the receptive field center (and 1.75 σ_e).

Hence the smallest of Wilson's channels, the N channel, will have a central diameter of $3.1\sqrt{2} = 4.38'$ which corresponds to about 9 foveal cones.

This figure cannot represent the smallest available channel. It is too large, for a number of reasons, of which the main ones are illustrated in Fig. 1. These are

(i) Two-point acuity is in fact around 1 min of arc.[8,9] Zero-crossings cannot separate two points as close as 1′ apart if the underlying receptive fields have a central diameter as great as 4′. In fact, a receptive field of at most about 2′ is required to provide this level of acuity, as shown by Fig. 1(a) (if the criterion relies on the separation of the zero-crossing profiles; see Waessle and Creutzfeld,[10] for instance, for another, more traditional criterion leading to about the same relation between acuity and receptive field diameter).

(ii) Two bars can be resolved at separations of the same order or smaller. Westheimer[11] reported a figure of 1′ at 75% confidence level, whereas Craik,[12] in an earlier and slightly different experiment, reported successful discrimination at separations as small as 30″.

(iii) Human sensitivity to gratings of high spatial frequency (up to 60 cpd[13]) and the reported receptive field sizes in the monkey[14–16] all suggest a minimum sized channel that has a central diameter smaller than 4′ (taking into account the fact that the monkey's acuity is probably a factor of two worse than that of humans).

(iv) Wilson made no observations above 16 cpd, so his experiments do not exclude the existence of a smaller channel.

For all these reasons, it is very likely that an additional, fifth channel exists, with a central diameter in the fovea of between 1′ and 2′.

It is therefore interesting to ask, what is the smallest possible channel consistent with the physiology? There are two important limiting factors: one is the optics of the eye; and, the other, the mosaic of retinal cones. Under optimal viewing conditions, the approximately Gaussian core (corresponding to the $[J_1(r)/r]^2$ diffraction image) of the eye's point-spread function has a σ of about 20″[19] (See also the line spread of Ref. 13, p. 571) (range 18″–23″). A single cone can be thought of as a cylinder 27″ in diameter[17,18] (range 25″–29″). Computations of the convolution between the optical point-spread function and the cone cross section show that the consequence of the cone's finite size is to increase the effective σ of the point-spread function, but by a surprisingly small amount, in fact to $\sigma = 22''$ (range 19″–25″).

If the central excitatory region of the channel's receptive field were driven by a single cone, its central diameter w_{2D} obtained from Eq. (2) (with $\gamma = 1.6$) is 77″ (range 67″–88″). This size channel leads (via Fig. 1) to a two-bar separation acuity of around 40″. It is therefore not inconsistent with the psychophysical data that the smallest (foveal) channel corresponds to a (midget) ganglion cell receiving excitatory input from a single cone (the corresponding inhibitory distribution has $\sigma_i = 35''$ and can be synthesized by summating neighboring cones, see later).

Interestingly, we have found (numerically) that coupling of a cone with its neighbors on a hexagonal lattice (perhaps via gap junctions) does not much increase the space constant of the effective point-spread function. Of course the effective point-spread function would become increasingly Gaussian-like as a consequence first, of the effect of the finite size of the photoreceptors and second, of the coupling between the cones (because of the central limit theorem). If a cone is coupled with weight 1 to its six neighbors, the value of σ increases from 22″ to only 32″ (range 29″–35″). Thus summation of neighboring cones leads to only a small loss in acuity (for these parameter values), while increasing the signal-to-noise ratio by a factor of up to $\sqrt{7}$. It is not clear to us, however, that this would usually be a desirable trade-off for foveal vision.

Finally, interpolation of the sampled values represented by these ganglion cells could locate the zero-crossings with a precision in the hyperacuity range.[19,20] Recent computer experiments show that even simple linear interpolation of the values of center-surround receptive fields preserve the positions of zero-crossings essentially as well as the ideal reconstruction schemes required by the sampling theorem.

ACKNOWLEDGMENTS

We are grateful to W. Richards and H. Waessle for useful discussions. This work was conducted at the Artificial Intelligence Laboratory, a Massachusetts Institute of Technology research program supported in part by the Advanced Research Projects Agency of the Department of Defense, and monitored by the office of Naval Research under contract number N00014-75-C-0643. D.M. was also supported by NSF contract number 77-07569-MCS.

[1] F. W. Campbell and J. Robson, "Application of Fourier analysis to the visibility of gratings," J. Physiol. **197**, 551–566 (1968).
[2] H. R. Wilson and S. C. Giese, "Threshold visibility of frequency gradient patterns," Vision Res. **17**, 1177–1190 (1977).
[3] H. R. Wilson and J. R. Bergen, "A four mechanism model for spatial vision," Vision Res. **19**, 19–32 (1979).
[4] J. R. Bergen, H. R. Wilson, and J. D. Cowan, "Further evidence for

four mechanisms mediating vision at threshold: Sensitivities to complex gratings and aperiodic stimuli," J. Opt. Soc. Am. **69,** 1580–1587 (1979).

[5] D. Marr and T. Poggio, "Theory of human stereopsis," Proc. R. Soc. London Ser. B **204,** 301–328 (1979).

[6] D. Marr, T. Poggio, and S. Ullman, "Bandpass channels, zero crossings, and early visual information processing," J. Opt. Soc. Am. **69,** 914–916 (1979).

[7] D. Marr and E. Hildreth, "A theory of edge detection," Proc. R. Soc. London **207,** 187–217 (1980).

[8] F. Tonner, "Die Sehschaerfe," Pflugers Arch. **247,** 183–193 (1943).

[9] G. Westheimer, "Diffraction theory and visual hyperacuity," Am. J. Optom. Physiol. Opt. **53,** 362–364 (1975).

[10] H. Waessle and O. D. Creutzfeld, "Spatial resolution in visual system: A theoretical and experimental study on single units in the cat's lateral geniculate body," J. Neurophysiol. **26,** 13–27 (1973).

[11] G. Westheimer, "Spatial frequency and light-spread descriptions of visual acuity and hyperacuity," J. Opt. Soc. Am. **67,** 207–212 (1977).

[12] J. W. Craik, "The effect of adaptation upon visual acuity," Brit. J. Psych. **29,** 252–266 (1939).

[13] F. W. Campbell and R. W. Gubisch, "Optical quality of the human eye," J. Physiol. **186,** 558–578 (1966).

[14] D. H. Hubel and T. N. Wiesel, "Uniformity of monkey striate cortex: a parallel relationship between field size, scatter and magnification factor," J. Comp. Neurol. **158,** 295–306 (1974).

[15] P. H. Schiller, B. L. Finlay, and S. F. Volman, "Quantitative studies of single-cell properties in the monkey striate cortex. III. Spatial frequency," J. Neurophysiol. **39,** 1334–1351 (1977).

[16] G. F. Poggio and B. Fischer, "Binocular interaction and depth sensitivity of striate and prestriate cortical neurons of the behaving rhesus monkey," J. Neurophysiol. **40,** 1392–1405 (1978).

[17] B. O'Brien, "Vision and resolution in the central retina," J. Opt. Soc. Am. **41,** 882–894 (1951).

[18] A. W. Snyder and W. H. Miller, "Photoreceptor diameter and spacing for highest resolving power," J. Opt. Soc. Am. **67,** 696–698 (1977).

[19] H. B. Barlow, "Reconstructing the visual image in space and time," Nature **279,** 189–190 (1979).

[20] F. H. C. Crick, D. C. Marr, and T. Poggio, "An information processing approach to understanding the visual cortex," in Cerebral Cortex Colloquium, edited by F. O. Schmitt, F. G. Worden, S. G. Dennis (MIT, Cambridge, to be published).