

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A. I. Memo No. 541

August, 1979

EVIDENCE FOR A FIFTH, SMALLER CHANNEL
IN EARLY HUMAN VISION

D. Marr, E. Hildreth, and T. Poggio

ABSTRACT Recent studies in psychophysics and neurophysiology suggest that the human visual system utilizes a range of different size or spatial frequency tuned mechanisms in its processing of visual information. It has been proposed that there exist four such mechanisms, operating everywhere in the visual field, with the smallest mechanism having a central excitatory width of 3' of arc in the central fovea. This note argues that there exists indirect evidence for the existence of a fifth, smaller channel, with a central width in the fovea of 1.5'.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-75-C-0643 and in part by National Science Foundation Grant MCS77-07569.

© MASSACHUSETTS INSTITUTE OF TECHNOLOGY 1979

The idea that the human visual system may use a range of different size or spatial frequency tuned mechanisms was originally introduced on the basis of psychophysical evidence by Campbell and Robson (1968). This led to a virtual explosion of papers dealing with spatial frequency analysis in the visual system. Recently, Wilson and Giese (1977), Wilson and Bergen (1979) integrated these and other anatomical and physiological data into a framework consisting of (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 receptive field sizes; (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 bandwidths of the N and S channels is about 1-3 octaves, but may be slightly larger for the T and U channels. (3) The receptive field shape of these channels is the difference of two gaussian distributions. (4) In the fovea, and using line stimuli, the widths w of the central excitatory regions of the receptive fields have the following values: W-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 size increases linearly with eccentricity, being about double at 4° eccentricity.

Essentially all of the psychophysical data on the detection of spatial patterns below 16 cycles/degree 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 in the channels.

This note argues that there exists indirect evidence from psychophysics and neurophysiology for the existence of a fifth, smaller channel, which in the central fovea would have a central excitatory width of roughly 1.5'.

Our current theory of these channels is that they are the first step in the detection of intensity changes in the image (Marr & Poggio 1979, Marr, Poggio, & Ullman 1979, Marr & Hildreth 1979). The critical idea is that the sustained channels can be regarded as spatial second derivative operators acting on the image at two scales. Sharp intensity changes correspond to zero-crossings (fast transitions from positive to negative values) in the output from these channels. It has been shown (a) that the optimal differential filter has a shape very similar to that found by Wilson and Bergen (Marr & Hildreth 1979), and (b) that the zero-crossings together provide rich information about the image (Marr, Poggio, & Ullman 1979).

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 we 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. If this is true, the values of w measured by Wilson and Bergen must be multiplied by .2 to obtain the diameter of the corresponding circularly symmetric centre-surround receptive field. Hence the smallest of their channels, the N-channel,

will have a central diameter of $3.1 \cdot 2 = 4.38'$; which corresponds to about 9 foveal cones.

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

(1) Two-point acuity is in fact somewhat more than 1' of arc, for an ideal subject under ideal conditions at the 75% confidence level, and this corresponds well to the diffraction limits imposed by the optics of the eye (Westheimer 1976, Snyder & Miller 1977). 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 least about 2' is required to provide the desired acuity, as shown in figure 1a.

(2) In a similar vein, two bars can be resolved at the 75% confidence level when they are separated by about 1' (Westheimer 1977). The same arguments apply here, and are illustrated in figure 1b.

(3) Human sensitivity to gratings of high spatial frequency (up to 60 c.p.d. (Campbell & Gubisch 1966)), and the reported receptive field sizes in the monkey (Hubel & Wiesel 1974, Schiller, Finlay, and Volman 1976, Poggio & Fischer 1978) all suggest a minimum sized channel that has a central diameter smaller than 4'.

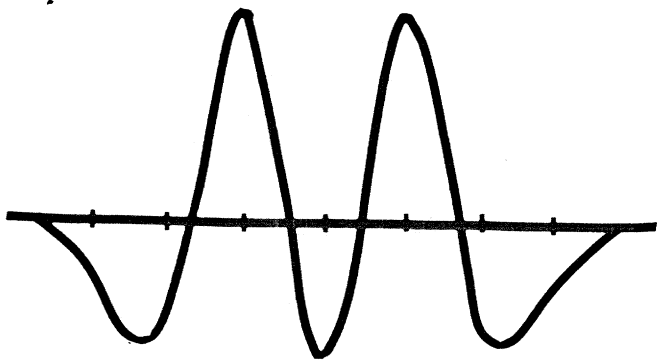
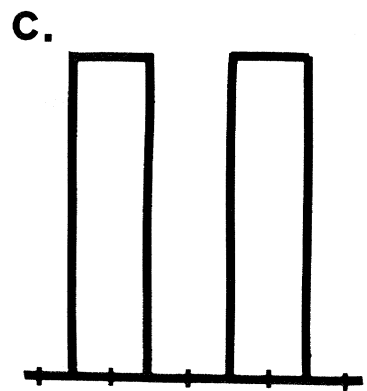
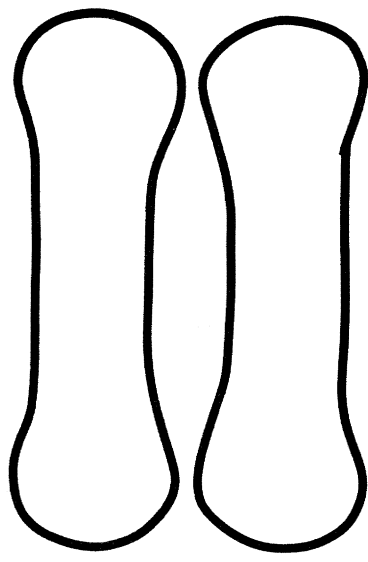
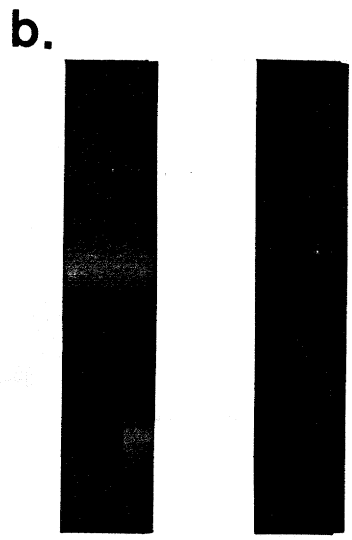
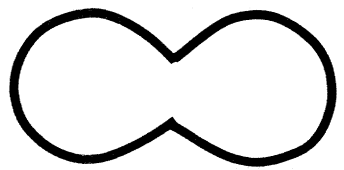
(4) Wilson made no observations above 16 c.p.d., so his experiments do not exclude the existence of a smaller channel.

For all these reasons, we predict that an additional, fifth channel

exists, with a central diameter in the fovea of about 2'. Using line stimuli in the manner of Wilson and Bergen, the measured value of w for this new, presumably sustained channel should be about 1'30". This fifth channel may be present only in the fovea.

Interpolation of the sampled values represented by these ganglion cells could locate the zero-crossings with a precision in the hyperacuity range (Barlow 1979, Crick, Marr, and Poggio 1979). 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.

Figure 1a shows the 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 $2.1'$. Figure 1b shows the zero-crossings associated with the two bars pattern filtered through the same receptive field. The angular separation in both cases is $1'$. A slightly larger receptive field or a smaller separation lead to a zero-crossings profile practically indistinguishable from the zero-crossings of a one-bar pattern. The output of the filtering stage is shown (in cross-section) in Figure 1c.



References

- Barlow, H. B. (1979) Reconstructing the Visual Image in Space and Time. *Nature* 279, 189-190.
- Campbell, F. W. & Gubisch, R. W. (1966) Optical and Retinal Factors Affecting Visual Resolution. *J. Physiol. (Lond.)* 186, 558-578.
- Campbell, F. W. & Robson, J. G. (1968) Applications of Fourier Analysis to the Visibility of Gratings. *J. Physiol. (Lond.)* 197, 551-556.
- Crick, F. H. C., Marr, D. & Poggio, T. (1979) in Cerebral Cortex Colloquium, Neuroscience Research Program, in press.
- Grimson, W. E. L. & Marr, D. (1979) in Proc. Image Understanding Workshop SRI.
- Hubel, D. H. & Wiesel, T. N. (1974) Uniformity of Monkey Striate Cortex: a Parallel Relationship Between Field Size, Scatter, and Magnification Factor. *J. Comp. Neurol.* 158, 295-306.
- Marr, D. & Hildreth, E. (1979) Theory of Edge Detection. *Proc. Roy. Soc.* (submitted).
- Marr, D. & Poggio, T. (1979) A Theory of Human Stereopsis. *Proc. Roy. Soc. Lond. B.* 204, 301-328.
- Marr, D., Poggio, T. & Ullman, S. (1979) Bandpass Channels, Zero-crossings, and Early Visual Information Processing. *J. Opt. Soc. Am.*, (in the press).
- Poggio, G. F. & Fischer, B. (1978) Binocular Interaction and Depth Sensitivity of Striate and Prestriate Neurons of the Behaving Rhesus Monkey. *J. Neurophysiol.* 40, 1392-1405.
- Snyder, A. W. & Miller, W. H. (1977) Photoreceptor Diameter and Spacing for Highest Resolving Power. *J. Opt. Soc. Am.* 67, 696-698.
- Schiller, P. H., Finlay, B. L. & Volman, S. F. (1976) Quantitative Studies of Single-cell Properties in Monkey Striate Cortex. *J. Neurophysiol.* 39, 1334-1351.
- Westheimer, G. (1976) Diffraction Theory and Visual Hyperacuity. *Am. J. Optometry and Physiological Optics* 53, 362-364.
- Westheimer, G. (1977) Spatial Frequency and Light-Spread Descriptions of Visual Acuity and Hyperacuity. *J. Opt. Soc. Am.* 67, 207-212.
- Wilson, H. R. & Giese, S. C. (1977) Threshold Visibility of Frequency Gradient Patterns, *Vision Res.* 17, 1177-1190.
- Wilson, H. R. & Bergen, J. R. (1979) A Four Mechanism Model for Spatial Vision. *Vision Res.* 19, 19-32.