Bandpass channels, zero-crossings, and early visual information processing

D. Marr and S. Ullman

Psychology Department, Massachusetts Institute of Technology, 79 Amherst Street, Cambridge, Massachusetts 02139

T. Poggio

Max-Planck-Institut fuer biologische Kybernetik, 74 Tuebingen 1, Spemannstrasse 38, Germany

(Received 11 September 1978; revised 15 December 1978)

Under appropriate conditions zero-crossings of a bandpass signal are very rich in information. The authors examine here the relevance of this result to the early stages of visual information processing, where zero-crossings in the output of independent spatial-frequency-tuned channels may contain sufficient information for much of the subsequent processing.

In most approaches to computer vision, an important preliminary computation is the localization of discontinuities in image intensity. This can be achieved by finding peaks in the first directional derivative of intensity, or equivalently, zero-crossings in the second directional derivative. The latter quantity may be obtained by convolving the image with a bar-shaped mask, which approximates the second directional derivative at its particular scale. By using a range of mask sizes, one can begin to deal with the wide range of scales over which changes take place in a natural image.[1]

 0030-3941/79/060914-03$00.50

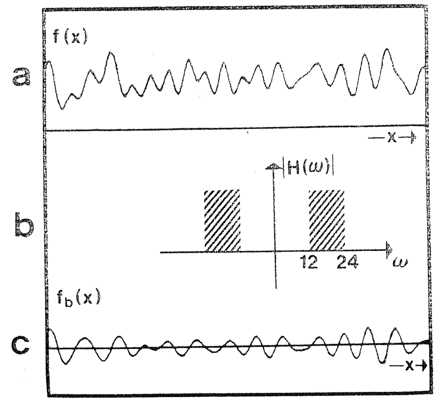

FIG. 1. The meaning of Logan's theorem. (a) shows a stochastic band-limited gaussian signal $f(x)$, and (c) exhibits the result $f_b(x)$ of filtering (a) through an ideal one-octave bandpass filter. The modulus of its transfer function is shown in (b). Since (c) has a bandwidth of one octave, and it has no zeroes in common with its Hilbert transform, Logan's theorem tells us that (c) is determined, up to a multiplicative constant, by its zero-crossings alone. The aspect of Logan's result that is important for this article is that under the right conditions, zero-crossings alone are very rich in information.

These ideas begin to account, on purely information processing grounds, for the presence of spatial-frequency-tuned channels in early human vision.[2] Recent work by Wilson and Giese[3] shows that such channels can be realized by linear units with bar-shaped receptive fields, reminiscent of the simple cells that Hubel and Wiesel[4] have described. Marr and Poggio's[5,6] recent theory of human stereopsis is, for example, conceived within this framework, and assumes that the elements that are matched between the two images are equivalent to the zero-crossings in bar-mask outputs. The object of this note is to point out that very recent advances in information theory provide fascinating additional theoretical support for this framework.

The advance in question is a theorem due to Logan,[7] who showed that if a one-dimensional analytic function is (a) bandpass of bandwidth less than one octave, and (b) has no free zeroes, i.e., complex zeroes in common with its Hilbert transform, then the function is completely determined (up to an overall multiplicative constant) by its (real) zero-crossings (see Fig. 1). Condition (a) is critical, but condition (b) can for practical purposes be ignored, since it is almost always satisfied except by pathological signals.

If one translates this result into the context of early visual processing, its meaning is this. We have already seen that the basic idea, of using zero-crossings in bar-mask convolutions from which to generate a primitive description of the image, has a strong physical motivation. Logan's result tells us that, if the bar-mask operators are bandpass with a bandwidth of less than one octave, then the zero-crossings alone are so rich in information that they determine essentially completely the convolution values (taken along a scan line perpendicular to the mask orientation). This method of recovering scan lines is the most obvious way of applying Logan's inherently one-dimensional result to two-dimensional images. It requires that each one-dimensional function along a scan line be bandpass with one octave bandwidth. Note that this is not necessarily satisfied by filtering the image through a two-dimensional bandpass filter [like a ring in the (ω_x, ω_y) plane of width one octave]. For instance, an image filtered through a (bandpass) bar-shaped mask is bandpass on each scan-line perpendicular to the mask's orientation; an image filtered through a (bandpass) circularly-symmetric mask is band-limited but not bandpass along any scan-line. This follows from the fact that the Fourier transform along (for instance) the x-axis of an image filtered through a bandpass "ring" is essentially the projection of the two-dimensional Fourier transform on ω_x, and is therefore not bandpass.

Within this framework, of a rather direct application of Logan's theorem, one is committed to orientation-dependent operators applied along scan lines across the image. Figure 2 illustrates how a two-dimensional image can be reconstructed from zero-crossings along scan-lines in the x- and y-directions. This method relies upon transforming the two-dimensional image into a set of functions to which Logan's one-dimensional theorem can then be applied. Although this is an existence proof that, under the appropriate conditions, zero-crossings specify an image completely, there are probably more powerful ways of extending Logan's result to two dimensions. Their existence remains however an open question.

In its extreme form, our thesis may be summarized as follows. In order to construct a faithful representation of the image using only zero-crossings, it is necessary to filter it through a set of independent bandpass channels with one octave bandwidth. Hence the masks (or receptive fields) that approximate the second directional derivative operator should, as closely as possible, be bandpass with one octave

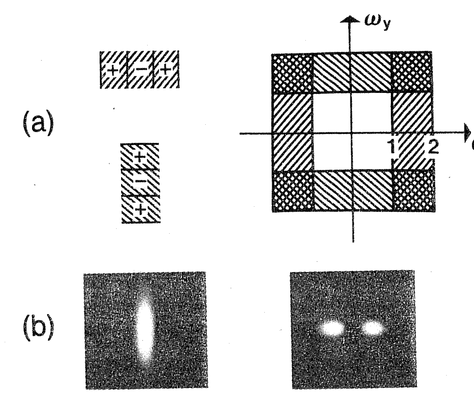

FIG. 2. (a) On the left are shown short bar-shaped masks at the vertical and horizontal orientations, and on the right, the amplitude of their (idealized) transfer functions. The bandwidth shown here is one octave, the maximum value for which Logan's theorem applies. (In practice, an ideal one-octave bandwidth requires side-lobes in the "receptive field"). If for each mask, zero-crossings are found along scan-lines lying perpendicular to the mask's orientation, these zero-crossings contain full information about that part of the image whose spectrum falls within the shaded region (on the right) of the Fourier plane. The remaining regions of the Fourier plane can be covered by similar masks of different sizes. Provided that there is sufficient overlap in the Fourier domain, information from different masks can in principle be combined to give the original image up to a single scaling factor. (b) If the masks are more elongated, the support of their Fourier transform will become smaller. In order to cover the Fourier plane a set of such masks of several orientations will be required. Figure 2(b) shows an elongated mask, whose cross section is the difference of two gaussians, together with its Fourier transform. Interestingly, if one uses masks constructed from the difference of two gaussian curves,[3] their Fourier transforms behave like ω^2 for values of ω that are small compared to σ. In other words, they approximate a second derivative operator.

bandwidth. Such a system would allow the recovery of sharp intensity changes directly from the mask outputs, while providing the necessary basis for the recovery of the information contained in arbitrary intensity profiles.

What experimental evidence is there that our thesis is relevant to biological visual systems? As we mentioned earlier, Logan's free zero condition will almost always be satisfied in practice. The critical condition concerns the bandwidth. There is ample evidence for the existence in the human visual system of independent, spatial-frequency-tuned bandpass channels, of about one octave bandwidth. Precise estimates of the bandwidth vary considerably, however, ranging from very narrow (0.5 octaves[8]) to very large values.[9,10] More recent approaches based on spatial probability summation allow most of the existing psychophysical data to be fitted using medium bandwidth channels. The especially convincing estimates of Wilson and Giese[3] hover around an octave and a half.[11] In any case, the channels are not the ideal one-octave bandpass filters that Logan's theorem requires. There is unfortunately little available information about channel characteristics in their normal (suprathreshold) conditions, although there are hints that their bandwidth may then be somewhat narrower.[12]

The important point is that Logan's theorem shows that zero-crossings of a one-octave bandpass signal contain complete information and the evidence is that they remain rich in information even when the one-octave condition is relaxed.

Experiments with 1.5 octave Fourier polynomials with randomly chosen coefficients indicate that only in 8% of the cases are there insufficient zero-crossings to determine the signal. For ergodic bandpass Gaussian processes Rice's formula (1945, p. 60) shows that Logan's conclusion holds up to 1.67 octave (K. Nishihara, personal communication). Alternatively, we conjecture that the allowed bandwidth may be increased by adding extra information, for example the derivative of the signal at the zero-crossings. In any case it becomes of considerable interest to determine the channel bandwidths under suprathreshold conditions.

Finally, observe that the physiological detection of zero-crossings need not depend on the detection of cells with zero response. For instance, near an intensity edge the zero-crossing in the bandpass signal is flanked by two peaks of opposite sign. Detection of zero-crossings can thus be performed on the basis of peaks, rather than zero responses.

ACKNOWLEDGMENTS

D. M. thanks Werner Reichardt and the Max Planck Society for their kind hospitality. We are grateful to I. Geiss for typing the manuscript and to L. Heimburger for drawing the figures. This work was conducted at the Max-Planck-Institut fuer biologische Kybernetik in Tuebingen, and 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. and S.U. were also supported by NSF contract number 77-07569-MCS.

[1]D. Marr, "Early processing of visual information," Phil. Trans. R. Soc. Lond. B **275,** 483–524 (1976), p. 488.

[2]F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings," J. Physiol. (London) **197,** 551–556 (1968).

[3]H. R. Wilson and S. C. Giese, "Threshold visibility of frequency gradient patterns," Vision Res. **17,** 1177–1190 (1977).

[4]D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," J. Physiol. (London) **160,** 106–154 (1962).

[5]D. Marr and T. Poggio, "Theory of human stereopsis," J. Opt. Soc. Am. **67,** 1400 (1977).

[6]D. Marr and T. Poggio, "A computational theory of human stereo vision," Proc. Roy. Soc. London B, (in press).

[7]B. F. Logan Jr., "Information in the zero-crossings of bandpass signals," Bell Sys. Techn. J. **56,** 487–510 (1977).

[8]M. B. Sachs, J. Nachmias, and J. G. Robson, "Spatial-frequency channels in human vision," J. Opt. Soc. Am. **61,** 1176–1186 (1971).

[9]J. J. Kulikowski and P. E. King-Smith, "Spatial arrangement of line, edge and grating detectors revealed by subthreshold summation," Vision Res. **13,** 1455–1478 (1973).

[10]R. M. Shapley and D. J. Tolhurst, "Edge detectors in human vision," J. Physiol. (London) **229,** 165–183 (1973).

[11]G. E. Legge, "Sustained and transient mechanisms in human vision: temporal and spatial properties," Vision Res. **18,** 69–81 (1978).

[12]J. D. Cowan, "Some remarks on channel bandwidths for visual contrast detection," Neurosciences Res. Prog. Bull, **15,** 492–517 (1977), Fig. A12.