Vision Res. Vol. 30, No. 7, pp. 1111–1117, 1990
Printed in Great Britain.

0042-6989/90 $3.00 + 0.00

LETTER TO THE EDITORS

FUNDAMENTAL LIMITS OF LINEAR FILTERS IN THE VISUAL PROCESSING OF TWO-DIMENSIONAL SIGNALS

C. Zetzsche and E. Barth
Lehrstuhl für Nachrichtentechnik, Technische Universität München, Arcisstr. 21, D-8000 München 2, F.R.G.

(*Received* 14 *September* 1989; *in revised form* 29 *November* 1989)

Since Lettvin, Maturana, McCulloch and Pitts (1959), neurophysiologists have known that the visual system contains detectors that respond to stimulus features such as "bugs", line ends, bars, corners, etc. (Hubel & Wiesel, 1965). During the same period, the theory of linear systems has been applied successfully to the analysis and modelling of visual functions (DeValois & DeValois, 1980). Interestingly, however, there is a fundamental incompatibility between these two approaches that has not yet received adequate attention. Linear filtering, even if modified by common nonlinearities like thresholding or rectification, will generally confound straight signals with signals that show essentially two-dimensional variations. This principle deficiency is illustrated for a curvature detector recently suggested by Dobbins, Zucker and Cynader (1987, 1989), which is based on a nonlinear combination of linear filters. However, the problem can be solved by using the mathematical formalism of differential geometry. We employ the concept of "Gaussian curvature" of surfaces to derive a class of physiologically plausible operators for the detection of two-dimensional signal variations. Two essential properties of these detectors turn out to be necessary: the use of "*and*" operations, that are impossible with linear filters, and a specific "*compensation principle*" corresponding to inhibitory interactions between orientation selective filters.

One example for the encoding of essentially two-dimensional signal variations is the detection of curvature. According to a recent hypothesis by Dobbins et al. (1987, 1989), this can be accomplished by using the difference between the outputs of two simple cells with different receptive field sizes to generate "endstopped" responses that proportionally vary to stimulus length and curvature. It can be shown, however, that this particular model, as well as any essentially linear system, is subject to response ambiguities in that it is always possible to find a stimulus of zero curvature that erroneously elicits a response. Consider the stimulus configuration shown in Fig. 1a. While the curved lines and short bars give rise to appropriate responses of the Dobbins et al. detector, it also reacts erroneously to certain straight stimuli on the right side (Fig. 1b). The corresponding critical spectral area within which such false responses can occur is indicated in Fig. 2.

The very reason for the occurrence of such ambiguous responses has to be sought in a fundamental limitation of linear filters in the processing of two-dimensional signals. Such signals can be classified into three elementary categories: (1) constant signals that show no variation at all; (2) intrinsically one-dimensional signals that are constant along one orientation and can, therefore, be completely characterized by their variation along the orthogonal orientation (here: 1D-signals); (3) actually two-dimensional signals that vary along all orientations (here: 2D-signals). Obvious examples of 1-D signals are straight lines, straight edges, or sinusoidal gratings with arbitrary orientation. Curved lines, curved edges and junctions, intersections, terminations, etc. are typical 2D-signals (Marko, 1974; Julesz, 1981). An essential requirement for all detectors which encode 2D-signal properties is that they should not erroneously respond to 1D-signals. Curvature detectors, for example, should not respond to straight stimuli. We will show that such an unambiguous detection of 2D-signal properties necessarily employs "and" operations. Such

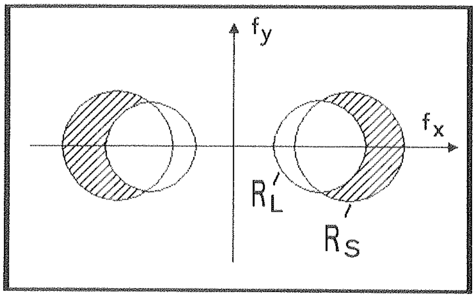

Fig. 2. Spectral representation of the different sized receptive fields used by Dobbins et al. (1987, 1989), providing an explanation for the responses of Fig. 1b. Although they are obviously not curved, certain sinusoidal gratings whose frequencies fall within the hatched area can pass the nonlinear filter combination. This becomes evident if the nonlinear equation given in Dobbins et al. (1987, 1989), is rewritten as a set of linear equations:

$R = [[c_s R_s] - [c_l R_l]]$ can be rewritten as:

$$R = \begin{cases} c_s R_s - c_l R_l & R_s > 0,\ R_l > 0,\ R_s > R_l c_l / c_s; \quad (1a) \\ c_s R_s & R_s > 0,\ R_l < 0; \quad (1b) \\ 0 & \text{otherwise}, \quad (1c) \end{cases}$$

with [.] denoting one-way rectification, c_s, c_l denoting constants, and R_s, R_l denoting responses from a small and a large receptive field. Since sinusoidal eigenfunctions pass both the R_s- and R_l-filter without a change in their basic form, their positive parts can fulfil equation (1a), provided R_s is great enough. Equation (1b) can be fulfilled by the sidelobes of the responses to dark straight lines. It is easy to show that, independent of the filter function, similar effects will occur for any conceivable "filter followed by threshold" operation.

operations have already been claimed to be of general importance for visual information processing (Marr & Hildreth, 1980; Reichardt & Poggio, 1981; Barlow, 1985; MacKay, 1985; Glünder, 1990). Signal processing in linear filters, however, can be seen as being restricted to an "or" combination of their inherently 1D basis functions. These basis functions, the eigenfunctions of linear shift invariant systems, are complex exponentials varying sinusoidally along one orientation, while being constant along the orthogonal one. Linear signal processing can be completely characterized via a decomposition of signals into these special 1D-signals (Gaskill, 1978). The action of a filter is to merely modulate the amplitude and phase of the input eigenfunctions, while it preserves their basic form. The final step within this descripton scheme is the additive recombination of these weighted eigenfunctions. It is now easy to see that for every linear filter, independently of its filter function, one can always find at least one 1D-signal causing a nonzero output. This is because the filter can be regarded as performing a logical "or" operation: on the one hand, the presence of a single input eigenfunction will produce a nonzero output signal (except for the trivial case of a zero filter coefficient). On the other hand, the additive interaction of several eigenfunctions can only change the form of the output signal but can never cause it to vanish completely for all (x, y), as is easily deduced from Parseval's Theorem (Gaskill, 1978).

It might be argued that this limitation of linear filters could be overcome by a combination of linear filters and nonlinear operations of the rectifying, clipping, thresholding, etc. type. However, if the nonlinear functions can be described, at least approximately, by piecewise linear transducer functions, the operation of the whole system can always be expressed by a set of linear equations, each being valid within a specific definition range determined by the nonlinearities (see the caption of Fig. 2 for an example). Since these equations are either zero, constant or linear filter operations, ambiguous responses to 1D-signals can only be avoided if the definition ranges are determined in such a way that the constant and linear subcases can not be obtained for any 1D-signal. It will turn out that this is only possible for the special case where the set of equations corresponds to some version of an "and" operation.

Our solution of the problem of response ambiguities is based on differential geometry. The approach is related to an earlier proposal by Koenderink and van Doorn (1987) and to techniques in image processing where the image intensity function is regarded as a surface onto which geometrical principles can be applied (Paton, 1975; Hsu, Mundy & Beaudet, 1978; Haralick, Watson & Laffey, 1983; Besl & Jain, 1988). A central concept of differential geometry is the approximation of a local two-dimensional surface by an "osculating" paraboloid. In the limit this paraboloid will degenerate to a plane: the surface is "planar". If one has to bend a plane, like bending a piece of paper, the surface is "parabolic" and the corresponding image intensity function is a 1D-signal (from this point of view, sinusoids are corrugated cardboard). For 2D-signals, like corners, the corresponding surface cannot be formed without elastic deformation (paper will crease). In this case the paraboloid is "elliptic" or "hyperbolic", and what we seek is the mathematical specification of this property. This specification is provided by the concept of "Gaussian

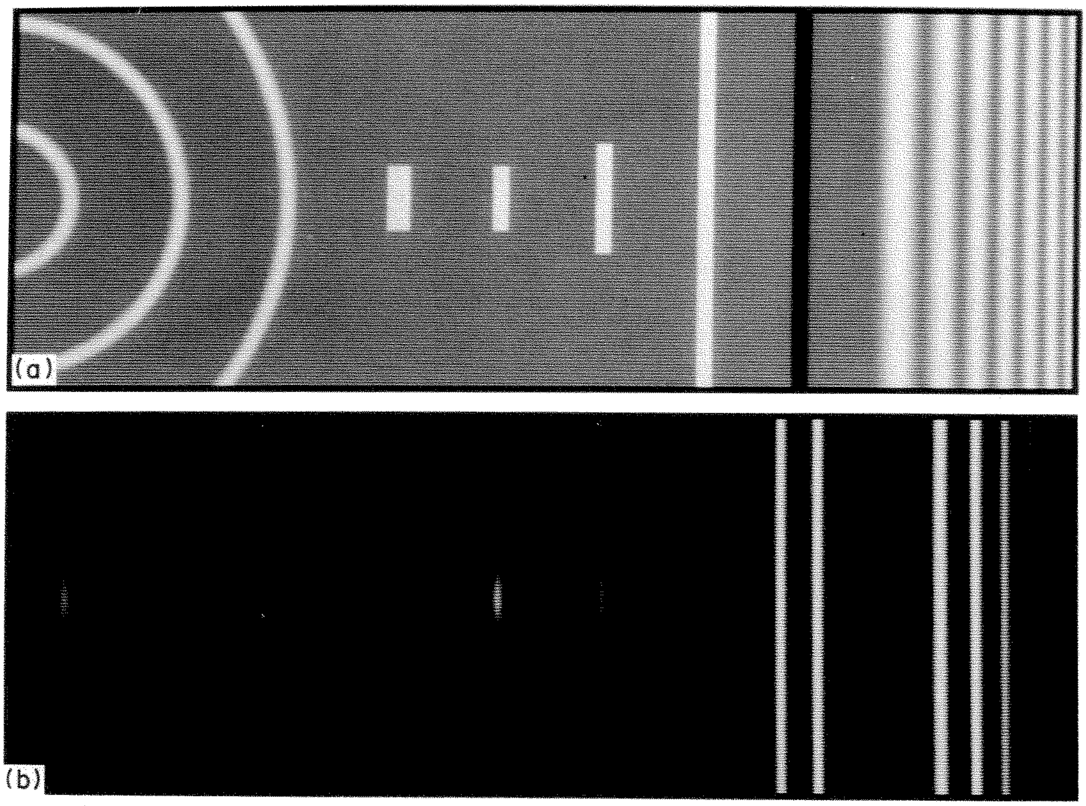

Fig. 1. (a) Input. Several 2D-signals are shown on the left side: three curved line segments with different degrees of curvature and three bars with different length and width. 1D-signals are on the right side: a bright straight line, a dark line, and a frequency modulated grating with spatial frequency increasing from left to right. (b) Output of the even-symmetry (ES) model of Dobbins et al. (1987, 1989). Note that the model responds as expected to the presence of curvature (left) and to the bar stimuli (center). Also as expected there is no response to the straight bright line. However, even the "optimal" stimulus (the middle bar) yields a weak response compared with the false-positive responses caused by suitable 1D-signals in the right side (dark line, grating with appropriate spatial frequency).

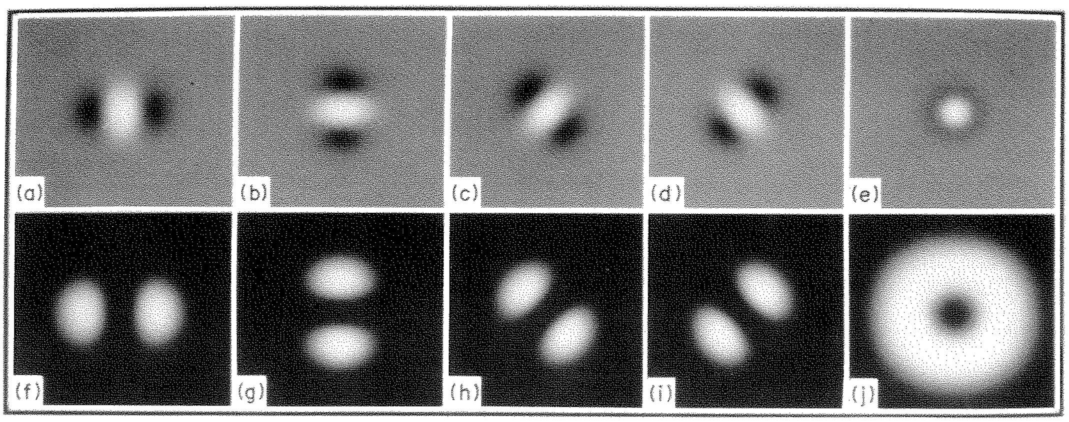

Fig. 3. Impulse responses (first row, a–e) and corresponding Fourier transforms (second row, f–j) of the terms appearing in the discriminant $D = l_{xx} l_{yy} - l_{xy}^2$. In addition to differentiation, an appropriate isotropic Gaussian band limitation is applied to ensure stability. (a) l_{xx}, (b) l_{yy}, (c) l_{uu}, (d) l_{vv}, (e) $D = l_{xx} l_{yy} - l_{xy}^2$. Obviously the spectra of l_{xx} and l_{yy} (f, g) have some overlap in the oblique areas of the frequency domain. This has to be compensated by substraction of $l_{xy}^2 = 1/4\ (l_{uu} - l_{vv})^2$, which has its spectral energy concentrated within the overlapping areas (h, i). Since the operator D is nonlinar, its impulse response (e) or its Fourier transform (j) can not be interpreted in the conventional sense. Nevertheless they illustrate the important fact that the nonlinear combination of anisotropic orientation-selective filters can lead to an *isotropic* behavior.

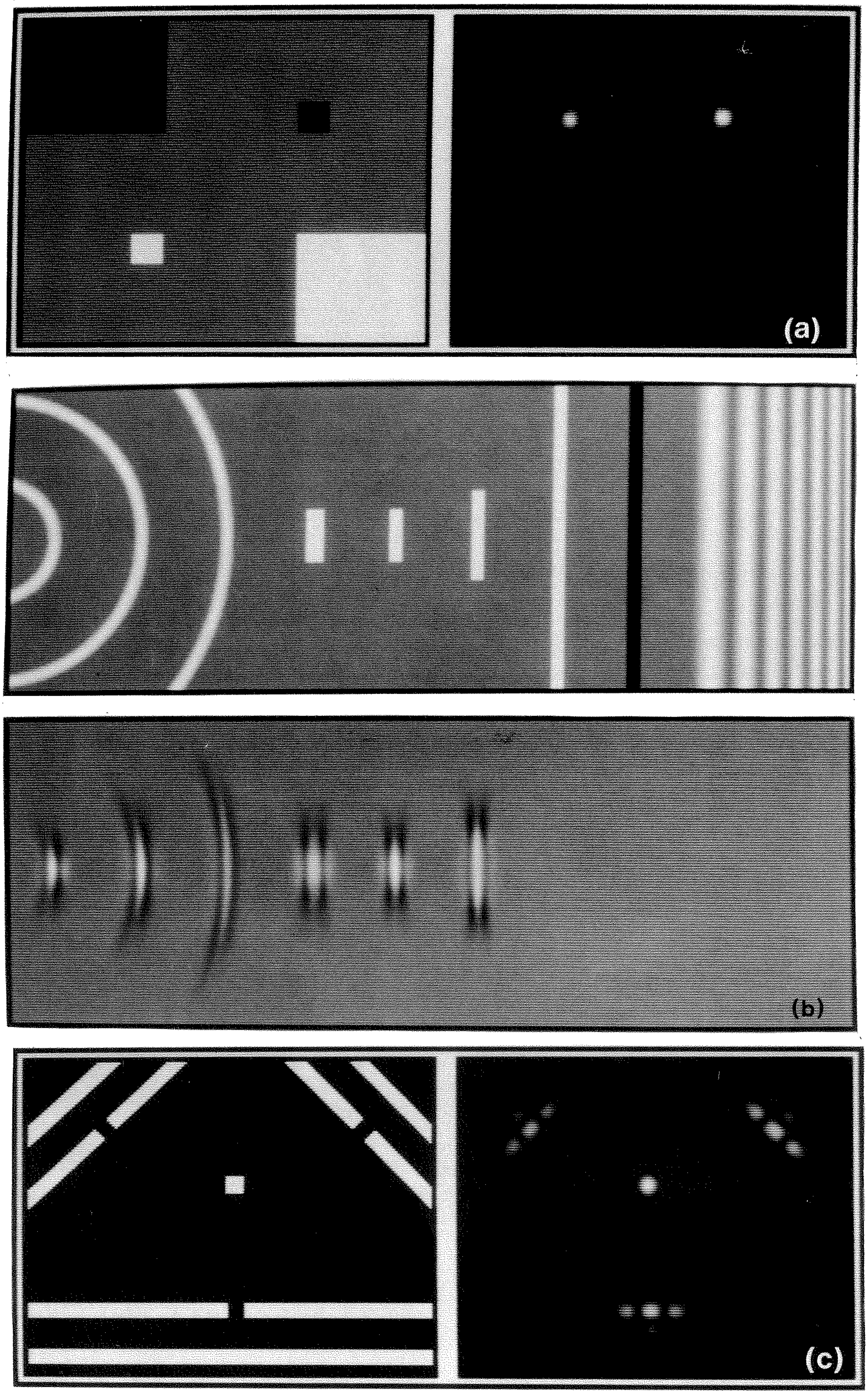

Fig 6(a–c). *Caption opposite.*

curvature" or, more precisely, by the discriminant of the osculating paraboloid (also known as determinant of the "Hessian") (Spivak, 1975). It can be written as:

$$D = l_{xx} \cdot l_{yy} - l_{xy}^2 \quad (2)$$

where $l = l(x, y)$ is the image intensity function, and the subscripts denote partial differentiation in the respective directions. D is zero for planar and for parabolic points, and hence a corresponding operator can already be regarded as a detector for 2D-signals (Beaudet, 1978). However, the above formula is of more general importance, since it allows the detection of the essential requirements for any unambiguous detection of 2D-signals, and it can easily be translated into terms of neural signal processing.

The impulse responses of the appropriately band-limited second derivtives l_{xx} and l_{yy} are shown in Fig. 3(a, b). Differentiation being a linear operation, they can be interpreted as oriented even-symmetric receptive field profiles or, equivalently, orientation selective filters (Young, 1985; Koenderink & van Doorn, 1987). Since any one of them will give an unwanted response to 1D-signals, their nonlinear multiplicative combination must be the essential operation for the computation of D. From the definition of 1D- and 2D-signals, it is clear that appropriate detectors should not respond if there is no variation at all, i.e. if both orientation filters yield zero outputs. Likewise, they should not respond if the signal is constant along one orientation (one filter yields a zero, the other, a nonzero output). They should respond, however, if (and only if) both filters indicate a signal variation. It is obvious that the required nonlinear combination of filter outputs corresponds to a logical "and" operation, and we claim that such an "and" has to be the necessary basis for any detector of two-dimensional signal variation.

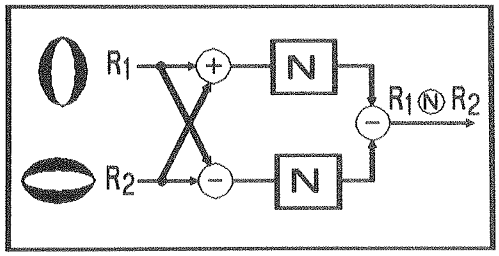

Fig. 4. Explicit multiplication can be avoided by the old Babylonian trick (Resnikoff & Wells, 1973),

$$R_1 \cdot R_2 = 1/4\,[(R_1 + R_2)^2 - (R_1 - R_2)^2].$$

This can be modified to:

$$\min(R_1, R_2) = 1/2\,(|R_1 + R_2| - |R_1 - R_2|);$$

or, even more general:

$$R_1 \circledN R_2 = N(R_1 + R_2) - N(R_1 - R_2);$$

where Ⓝ indicates a generalized "and" operation, and "N" can be any even-symmetric nonlinearity. A general test for the correctness of 2D-signal detectors would consist of a power series expansion of the nonlinearities and an analysis of the resulting terms for the appearance of products of complex exponentials whose two-dimensional frequencies do *not* differ in orientation. If such products exist, it is definitely possible to find a corresponding 1D-signal that will cause a false-positive response.

The specific implementation of the "and", however, can take place in various ways. Thresholding of the sum is the trivial solution for the restricted case of binary signals. Multiplication seems to be the obvious method in the case of analog signals but may be considered difficult as a neural operation. A physiologically plausible circuit for a generalized "and" operation is shown in Fig. 4. The most general case,

Fig. 6. (a) A "bug detector". (Input left, output right side.) "Off"-type filter outputs can be defined as: $R^- = -R$ if $R < 0$ and $R^- = 0$ otherwise. "And" combinations of such filters respond preferably to dark convex objects. Here even-symmetry, orientation-selective DOG's with orthogonal orientation axes have been used (cf. Fig. 5a). (b) Hypercomplex cell. (Input upper half, output lower half). The same input as for the Dobbins et al. operator is used (cf. Fig. 1). The hypercomplex cell is wired according to the scheme in Fig. 5c. For ease of computation, polar separable orientation filter functions have been used for all three filters. Orientation half-bandwidth is 26 deg, radial bandwidth is 2 octaves. Angular separation of the two "and" combined orientation filters is 26 deg. Note the hyperacuity property: a slight change from a straight to a curved line raises the response from zero to a considerable value. Due to the filter configuration employed, this hyperacuity property is confined to a limited range of orientations, as has been suggested by Watt and Andrews (1982). (c) Dot-responsive cell. (Input left, output right side.) While these cells do not respond to 1D-signals of arbitrary orientation, they do respond to interruptions in straight lines and to isolated dots. An essential property is their isotropic behaviour. It is interesting to note, however, that this isotropy need not necessarily be based on isotropic operations, but can result from an appropriate nonlinear combination of anisotropic orientation-selective filters, as in the discriminant formula. The combination of four orientation-selective filters shown in Fig. 3a–d has been used. Responses are two-way rectified.

however, comprises all those combinations of linear and nonlinear operations that can be guaranteed to produce solely mixed terms, i.e. products of eigenfunctions with different orientations. The approach of Koenderink and Richards (1988) can be regarded as a special case where the "and" is implicitly used in the form of a "gating" operation. A response occurs if and only if one orientation selective filter has the maximal output *and* another filter with a different orientation selectivity yields a nonzero output.

However, the "and" combination of orientation filter outputs cannot ensure the avoidance of false-positive responses to 1D-signals, unless the filters are strictly independent, i.e. have no spectral overlap. This condition can be fulfilled for orthogonally oriented filters in the human visual system (Fig. 5a), but restricts the sensitivity of detectors to a limited class of 2D-signals. We suggest, however, that various "and" combinations of orientation filters differing in size, even/odd-symmetry and angular separation are necessary to capture the variety of 2D-signals in the natural environment (Fig. 5a–c). In such combinations, depending on angular separation and orientation bandwidth, both filters may overlap spectrally, and therefore will respond to a suitable 1D-signal. Hence the "and" will lead to a false response. Here the second important property of the discriminant D comes into play. We call it the "*compensation principle.*" False responses to 1D-signals due to overlapping filter ranges can be avoided by subtraction of a compensation term, like the squared mixed partial derivative l_{xy} in (2), which can be written as:

$$l_{xy} = 1/2\,(l_{uu} - l_{vv}); \qquad (3)$$

with u and v being the oblique orientations (Koenderink, 1988). l_{uu} and l_{vv} can also be seen as orientation-selective filter operations (cf. Fig. 3c, d). If the filters that are "and" combined have a relatively sharp orientation tuning, as has been reported from neurons in the visual cortex (DeValois, Albrecht & Thorell, 1982), compensation is necessary for small angular separations only and can be achieved through inhibition from a single filter with intermediate orientation (Fig. 5c). Inhibitory interactions between orientation filters are a well-established fact and might even play a role in perceptual illusions caused by certain 2D-signals (Hess, Negishi & Creutzfeldt, 1975; Blakemore, Carpenter & Georgeson, 1970).

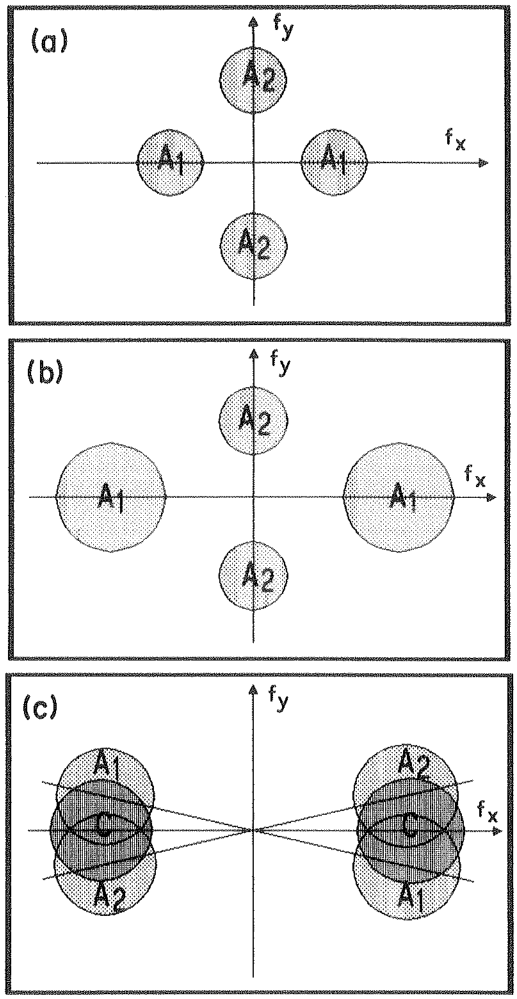

Fig. 5. Three examples of combinations of orientation filters differing in orientation, scale, and angular separation, respectively. In (c), orientation filters A_1 and A_2 are inhibited by C to compensate the overlap. The compensation filter function can be derived from the Fourier transform of a generalized version of the discriminant equation:

$$\mathscr{F}\{D'(x,y)\} = E(\rho,\theta)\,A_1(\rho,\theta) * E(\rho,\theta)\,A_2(\rho,\theta) - E(\rho,\theta)\,C(\rho,\theta) * E(\rho,\theta)\,C(\rho,\theta);$$

where ρ, θ are polar frequencies, E is the Fourier transform of an arbitrary input, A_1, A_2 are the transfer functions of the "and" combined filters, and C, of the compensation filter. * denotes convolution with respect to ρ. This equation can be easily solved analytically for C if the filter functions are polar separable $[A(\rho,\theta) = A_\rho(\rho)A_\theta(\theta)]$ and have the same radial frequency tuning $[A_{1\rho}(\rho) = A_{2\rho}(\rho) = C_\rho(\rho)]$. The correct angular tuning of the compensation filter is then given by $C_\theta(\theta) = \sqrt{\{A_{1\theta}(\theta)\,A_{2\theta}(\theta)\}}$.

In conclusion, we have shown that the unambiguous encoding of 2D-signals is a nontrivial capability achievable neither with linear filters nor by the common nonlinearities. Differential geometry, however, provides the framework for a wiring of visual cells that enables such a capability. The essential aspects of this framework are a nontrivial (Reichardt & Poggio, 1981), nonlinear combination of orientation-selective responses by an "and" operation and a suitable inhibition from intermediate orientations. It is interesting to note that cells which are obviously specified in the analysis of 2D-signals and which, therefore, resist any linear analysis, have been known for as long as 30 yr, namely the famous "bug detectors" of Lettvin et al. (1959) (Fig. 6a). For higher vertebrates, hypercomplex cells (Hubel & Wiesel, 1965) (Fig. 6b), and the recently discovered "dot responsive" cells (Saito, Tanaka, Fukada & Oyamada, 1988) (Fig. 6c), are also typical 2D-signal detectors. Although the latter two represent a considerable part of the visual cortex (Orban, 1984; Saito et al., 1988), exploration and modelling of spatial vision has been clearly dominated by linear filter theory in the last twenty years. From the logical point of view, the latter theory can only provide one half of the truth, namely "or" operations. The inclusion of "and" operations and geometrical principles in the theoretical concepts might bring us one step closer to understanding biological vision.

Acknowledgements—We thank I. Rentschler, K. O'Regan and an anonymous reviewer for valuable suggestions. The research was supported by grants of DFG and BRAIN to C.Z.

REFERENCES

Barlow, H. B. (1985). Cerebral cortex as a model builder. In Rose, D. & Dobson, V. G. (Eds.), *Models of the visual cortex* (pp. 37–46). New York: Wiley.

Beaudet, P. R. (1978). Rotationally invariant image operators. *4th International Joint Conference on pattern recognition*, Kyoto, Japan, pp. 578–583.

Besl, P. J. & Jain, R. C. (1988). Segmentation through variable-order surface fitting. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 10(2)*, 167–192.

Blakemore, C., Carpenter, R. H. S. & Georgeson, M. A. (1970). Lateral inhibition between orientation detectors in the human visual system. *Nature, London 228*, 37–39.

DeValois, R. L. & DeValois, K. K. (1980). Spatial vision. *Annual Reviews in Psychology, 31*, 309–314.

DeValois, R. L., Albrecht, L. G. & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research, 22*, 545–559.

Dobbins, A., Zucker, S. W. & Cynader, M. S. (1987). Endstopped neurons in the visual cortes as a substrate for calculating curvature. *Nature, London, 329*, 438–441.

Dobbins, A., Zucker, S. W. & Cynader, M. S. (1989). Endstopping and curvature. *Vision Research, 29*, 1371–1387.

Gaskill, J. D. (1978). *Linear systems, Fourier transforms, and optics*. New York: Wiley.

Glünder, H. (1990). $\Sigma\pi$-networks for motion and invariant form analysis. In Eckmiller, R., Hartmann, G. & Hauske, G. (Eds.), *Proceedings of the international conference on parallel processing in natural systems and computers*. Amsterdam: Elsevier. In press.

Haralick, R. M., Watson, L. T. & Laffey, T. J. (1983). The topographic primal sketch. *International Journal of Robotics Research, 2(1)*, 50–72.

Hess, R., Negishi, K. & Creutzfeldt, O. D. (1975). The horizontal spread of intracortical inhibition in the visual cortex. *Experimental Brain Research, 22*, 415–419.

Hsu, S., Mundy, J. L. & Beaudet, P. R. (1978). Web representation of image data. *4th international joint conference on pattern recognition*, Kyoto, Japan, pp. 675–680.

Hubel, D. H. & Wiesel, T. N. (1965). Receptive fields and functional architecture in two non-striate areas (18 and 19) of the cat. *Journal of Neurophysiology, 28*, 229–289.

Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature, London, 290*, 91–97.

Koenderink, J. J. (1988). Operational significance of receptive field assemblies. *Biological Cybernetics, 58*, 163–171.

Koenderink, J. J. & van Doorn, A. J. (1987). Representation of local geometry in the visual system. *Biological Cybernetics, 55*, 367–375.

Koenderink, J. J. & Richards, W. (1988). Two-dimensional curvature operators. *Journal of the Optical Society of America, A5*, 1136–1141.

Lettvin, J. Y., Maturana, H. R., McCulloch, W. S. & Pitts, W. H. (1959). What the frog's eye tells the frog's brain. *Proceedings of the IRE, 47*, 1940–1951.

MacKay, D. M. (1985). The significance of "feature sensitivity". In Rose, D. & Dobson, V. G. (Eds.), *Models of the visual cortex* (pp. 47–53). New York: Wiley.

Marko, H. (1974). A biological approach to pattern recognition. *IEEE SMC, 4*, 34–39.

Marr, D. & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society, London, B207*, 181–217.

Orban, G. A. (1984). *Neuronal operations in the visual cortex*. Heidelberg: Springer.

Paton, K. (1975). Picture description using Legendre polynomials. *Computer Graphics and Image Processing, 4*, 40–54.

Reichardt, W. & Poggio, T. A. (1981). *Theoretical approaches in neurobiology* (pp. 186–187). Cambridge: MIT Press.

Resnikoff, H. L. & Wells, R. O. (1973). *Mathematics in civilization*. New York: Holt, Rinehart & Winston.

Saito, H., Tanaka, K., Fukuda, Y. & Oyamada, H. (1988). Analysis of discontinuity in visual contours in area 19 of the cat. *Journal of Neuroscience, 8(4)*, 1311–1143.

Spivak, M. (1975). *A comprehensive introduction to differential geometry* (Vols I–V). Boston, Mass.: Publish or Perish.

Watt, R. J. & Andrews, D. P. (1982). Contour curvature analysis: Hyperacuties in the discrimination of detailed shape. *Vision Research, 22*, 449–460.

Young, R. A. (1985). The Gaussian derivative theory of spatial vision: Analysis of cortical receptive field lineweighting profiles. General Motors Research Publication, GMR-4920, Warren, Michigan.