

FUNCTIONAL SIZE INVARIANCE IS NOT PROVIDED BY THE CORTICAL MAGNIFICATION FACTOR

PATRICK CAVANAGH

Département de Psychologie, Université de Montréal, Montréal, Québec, Canada H3C 3J7

(Received 21 July 1981; in revised form 22 March 1982; in final form 14 May 1982)

Abstract—Schwartz (1980) proposes that the log polar or “complex log” mapping of the striate cortex that gives rise to the cortical magnification factor also serves to provide size invariance in vision. It is shown that, in fact, the cortical magnification factor cannot support functional size invariance. The essential flaw in Schwartz’ proposal is the lack of position invariance in the striate representation.

Schwartz (1980) proposes that the representation of visual space on the surface of the striate cortex may be approximated by a log polar mapping (he uses the term “complex log” mapping). He states that the properties of this representation have direct functional utility for perceptual processes and specifically “the complex logarithmic remapping of the visual pattern at the local level of the striate cortex provides a possible mechanism for size and rotation invariances in vision”. Moreover, he states that the size invariance obtained in transform sequences proposed by myself and others (Cavanagh, 1978; Brousil and Smith, 1967; Casasent and Psaltis, 1976) is entirely due to the log polar constituent of the sequence and so is a derivative of his proposal. In making these claims, Schwartz has vastly overstated the potential contribution of the cortical magnification factor to size invariance and has seriously misrepresented the size invariant transform sequence that I proposed. I will show first that the cortical magnification factor cannot support functional size invariance and second that the fundamental difference between Schwartz’s proposal and my own—that of a preliminary position invariant transform—is absolutely essential if true size invariance is to be achieved.

According to Schwartz, changes in the size and orientation of a visual stimulus will, for a given fixation point, produce simple shifts in the position of an invariant representation on the striate surface. Schwartz claims that this constancy of striate representation serves as a basis of a simple form specific encoding for pattern recognition processes. I will demonstrate, however, that the conditions for which the spatial mapping of the striate cortex can provide size and rotation invariances are too restrictive to have any functional role in perception. In addition, I will compare pattern recognition data for those situations satisfying the restrictions necessary for invariances to hold on the striate representation against those situations violating the necessary conditions. Additional perceptual capacities should be evident for those situations meeting the required conditions. If no

differences are revealed it must be assumed that the invariances inherent in the striate mapping do not contribute functionally to perceptual processing.

There are three important limitations to the size and rotation invariances of the striate mapping proposed by Schwartz.

(1) The proposed log polar mapping is an approximation which is only reasonably accurate outside the central 4° of the visual field (Schwartz, 1980, p. 649). Size and rotation invariances should therefore only hold for figures whose contours are entirely outside this 4° central region. It should be noted here that the central 4° of visual field are of utmost importance to pattern vision. Any process that cannot account for size invariance within this area can only be playing a secondary role at best in the overall size invariance mechanism.

(2) The mapping does not possess position invariance. Schwartz claims that the striate pattern of activity for a large or small stimulus or for an upright or rotated stimulus is invariant except for shifts of its position on the striate surface due to changes in size and orientation—but only as long as a given fixation point is maintained. Recognition of a shape at its original locus of presentation should therefore call upon a qualitatively different process than does its recognition at a new locus.

(3) The mapping is only invariant to size and rotation changes if these changes are centered at the origin. That is, a log polar striate representation is invariant to size and rotation changes only if they can be represented by simple magnifications or rotations of the entire plane about its center—the fovea in the case of the human visual system. Any object that doubled in size would also have to double its distance from the fovea to maintain an invariant striate representation. Recognition for shapes undergoing size or rotation changes centered at the fovea should therefore be qualitatively different from recognition involving changes centered elsewhere.

An examination of the various studies concerning the effects of size changes on recognition reaction

time reveals no evidence of the qualitative differences predicted by Schwartz's hypothesis.

Observers are quite able to recognize a random shape presented centrally at a different size than its original presentation even when the shape is too small (0.86° to 2.58° , Howard and Kerst, 1978) for Schwartz' size invariant process to be effective. The recognition reaction time in this experiment increases as a fairly linear function of the size ratio between the test and target stimuli. Similar data are also seen for letter (Larsen and Bundesen, 1978) and rectangle (Sekuler and Nash, 1972) stimuli. It might be argued that a second size invariant process is at work for shapes smaller than 4° in extent. However, the patterns of reaction times for recognitions when both the original and the test are within 4° are qualitatively and quantitatively similar to the patterns seen when both stimuli are greater than 4° (Larsen and Bundesen, 1978).

A change in the location of the test stimulus or a change in its size about a center other than the fovea should render the size invariance property of the striate mapping inoperative. Nevertheless, subjects are able to recognize that two random shapes presented on adjacent sides of the fovea are the same even though they differ in size (Bundesen and Larsen, 1975)—a presentation that violates both the position and the centering restrictions. Not only were the subjects able to perform with very low error rates but the data here strongly resemble the data for presentations conforming to the requirements for striate size invariance (same presentation location, size change centered at the fovea, Larsen and Bundesen, 1978).

In sum, these data are most simply explained by a single-size invariant mechanism capable of recognizing objects after arbitrary size and position changes. Such general size invariance is impossible for the striate mapping mechanism and since there was no evidence of two distinct size invariant mechanisms in the data considered here, the striate mapping hypothesis would appear to be ruled out all together.

The phenomenology of the constancy of object perception also argues against Schwartz' proposal. Consider the following trajectories of objects in the visual field. For a shape in the periphery to maintain a constant striate representation as it followed a circular path around the fovea, it would also have to rotate as it orbited (i.e. as if painted on a disk turning with the fovea as its center). However, the phenomenology is quite the opposite. A triangle orbiting in this fashion is alternatively seen as a delta and a triangle. If the triangle maintains a fixed orientation around the orbit it appears to remain the same shape although its striate representation is continuously deforming (see Fig. 1). Similarly, a rectangle moving from the periphery to the fovea while remaining the same size maintains a constancy of shape as much as a square starting out large in the periphery and decreasing steadily in size as it approaches the fovea. It is the second trajectory that maintains an invariant striate representation however; the representation for the first

trajectory changes from being a closed object when not covering the fovea to being an open contour when it encloses the fovea (Fig. 2).

Finally, if an observer were moving through an environment, shapes would undergo both magnifications and shifts as they grow larger and move into the periphery. Schwartz (1981) has maintained that this is a naturally occurring trajectory that maintains a fixed striate representation for each shape. However, even if the observer constantly maintains fixation along the axis of movement (a necessary restriction), objects moving into the periphery will be seen increasingly in side view. Thus in fact, only shapes having the same side and front views (e.g. spheres) will maintain invariant striate representations. Of all possible trajectories in depth only that directly along the axis of gaze will provide an invariant striate representation (and here only as long as all contours remain outside the central 4 degrees of the visual field).

There therefore appears to be no evidence supporting the proposal of a functional, size and rotation invariant representation on the surface of the striate cortex. Moreover, the severe restrictions outlined above would make the potential role of such a representation of minimal importance.

The transform proposed by Schwartz fails to provide true size and rotation scaling because of the necessity for size and rotation changes to be centered at the origin. Schwartz is thus mistaken in claiming that the size and rotation invariances demonstrated by Brouil and Smith (1967), Casasent and Psaltis (1976) and myself (Cavanagh, 1978) are due to a "complex log preprocessing". The very success of these transform sequences is due exactly to *not* using the log polar stage with its various restrictions as the first level. The first stage of all the transform sequences is a Fourier amplitude transform. There is no necessity, however, that this first step be a Fourier transform (Cavanagh, 1978, p. 168). I have pointed out that the representation reported by Maffei and Fiorentini (1977) of orthogonal axes of preferred spatial frequency (or bar size) and orientation is similar enough to the log polar frequency transform that I proposed to suggest the possibility that this organization may subservise size invariance in the visual system. A recent study by Tootell *et al.* (1981) also suggests a local, two-dimensional, spatial frequency by orientation analysis, although organized parallel to the surface of the cortex rather than perpendicular to the surface as Maffei and Fiorentini (1977) reported. The requirements necessary for either of these potential local organizations to support size invariance have been described previously (cf. Cavanagh, 1978, Fig. 3). The data of Tootell *et al.* (1981; see also Thompson and Tolhurst, 1980a,b) do not have sufficient resolution for a full evaluation of the requirements for size invariance but are not inconsistent with it. A more recent study by Berardi *et al.* (1982) appears to support the necessary logarithmic spacing of preferred frequency as well as orthogonal axes of

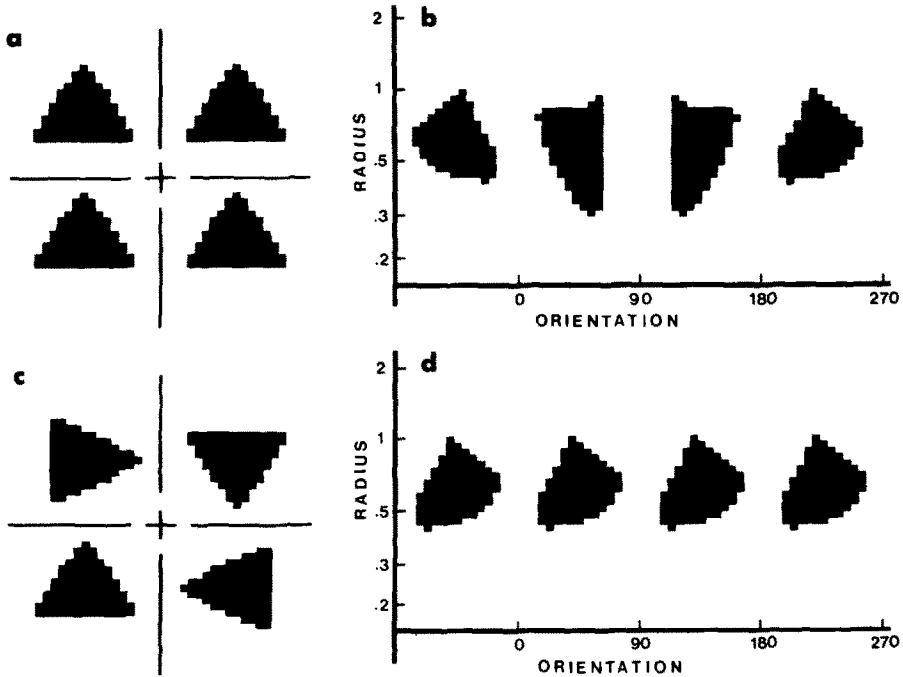


Fig. 1. (a) An upright triangle at four equidistant points about the visual origin (fovea); all four triangles are perceptually equivalent. (b) The log polar representation of the four triangles in (a) the patterns have lost their equivalence of shape. (c) Four triangles orbiting clockwise about the origin; the four triangles are not perceptually equivalent. (d) The log polar representations of the four triangles in (c) retain the same shape regardless of orientation.

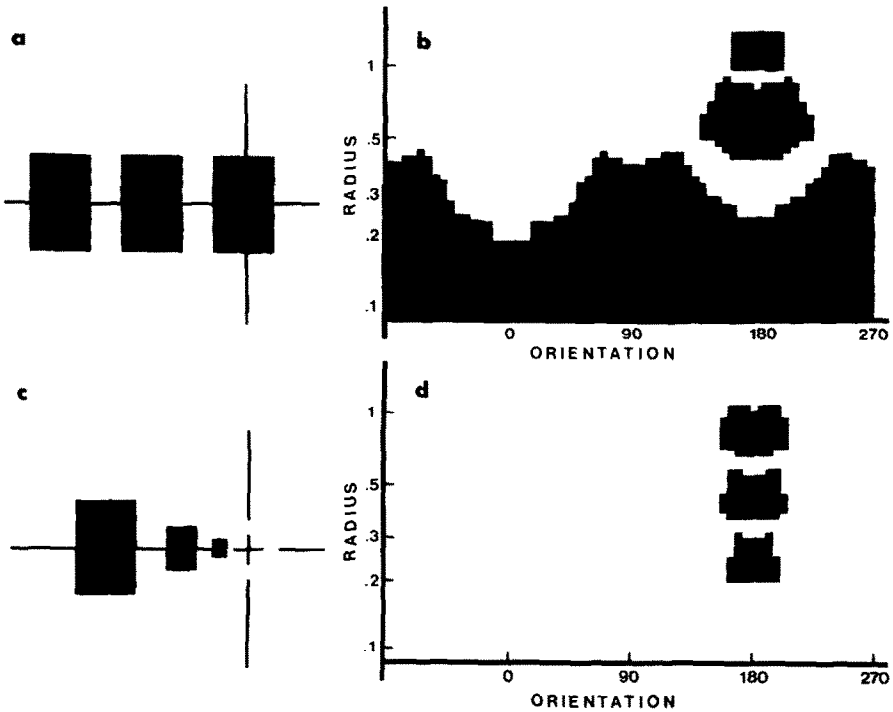


Fig. 2. (a) Three identical rectangles equally spaced from the fovea to the periphery. (b) The log polar representation of the three rectangles in (a). (c) Three rectangles whose sizes decrease in direct proportion to their distance from the origin. (d) The log polar representation of the rectangles in (c)—differences among the three shapes are due to sampling errors inherent in the discrete matrix representation.

frequency and orientation. Thus, although the nature of the local representation in the striate cortex remains controversial, there is no evidence that would suggest that the local log polar frequency transform hypothesis (Cavanagh, 1978) should be rejected.

Schwartz raises a number of other subjects concerning the size invariant transformation that I have proposed: specifically, the integration of local transforms, the insensitivity of the inferotemporal cortex to periodic stimuli, and the encoding of phase information. These questions have been dealt with in detail elsewhere (Cavanagh, 1978; 1981).

In conclusion, I have argued that there is no evidence that the architecture of the retinal encoding on the striate surface is in any way functionally involved in size invariance. On the other hand, it is possible that the local organization of orientation and size dimensions in the striate cortex (Maffei and Fiorentini, 1977; Tootell *et al.*, 1981) may contribute directly to a size invariant encoding (Cavanagh, 1978; 1981).

Acknowledgements—This research was supported by NSERC grant A8606 and by the Ministère de l'Éducation du Québec. I am grateful to Dan Pollen for helpful comments and suggestions.

REFERENCES

- Berardi N., Bisti S., Cattaneo A., Fiorentini A. and Maffei L. (1982) Correlation between the preferred orientation and spatial frequency of neurones in visual areas 17 and 18 of the cat. *J. Physiol.* **323**, 603–618.
- Brousil J. K. and Smith D. R. (1967) A threshold logic network for shape invariance. *I.E.E.E. Trans. Comput. EC-16*, 818–828.
- Bundesen C. and Larsen A. (1975) Visual transformation of size. *J. exp. Psychol., Hum. Percept. Perform.* **1**, 214–220.
- Casasent D. and Psaltis D. (1976) Position, rotation, and scale invariant optical correlation. *Appl. Opt.* **15**, 1793–1799.
- Cavanagh P. (1978) Size and position invariance in the visual system. *Perception* **7**, 167–177.
- Cavanagh P. (1981) Size invariance: Reply to Schwartz. *Perception* **10**, 469–474.
- Howard Jr J. H. and Kerst S. M. (1978) Directional effects of size change on the comparison of visual shapes. *Am. J. Psychol.* **91**, 491–499.
- Larsen A. and Bundesen C. (1978) Size scaling in visual pattern recognition. *J. exp. Psychol., Hum. Percept. Perform.* **4**, 1–20.
- Maffei L. and Fiorentini A. (1977) Spatial frequency rows in the striate visual cortex. *Vision Res.* **17**, 257–264.
- Schwartz E. L. (1980) Computational anatomy and functional architecture of striate cortex: A spatial mapping approach to perceptual coding. *Vision Res.* **20**, 645–669.
- Schwartz E. L. (1981) Cortical anatomy and spatial frequency analysis. *Perception* **10**, 455–468.
- Sekuler R. and Nash D. (1972) Speed of size scaling in human vision. *Psychon. Sci.* **27**, 93–94.
- Thompson I. D. and Tolhurst D. J. (1980a) Optimal spatial frequencies of neighbouring neurones in the cat's visual cortex. *J. Physiol.* **300**, 57 P.
- Thompson I. D. and Tolhurst D. J. (1980b) The representation of spatial frequency in cat visual cortex: a ¹⁴C-2-deoxyglucose study. *J. Physiol.* **300**, 58–59 P.
- Tootell R. B., Silverman M. S. and De Valois R. L. (1981) Spatial frequency columns in primary visual cortex. *Science* **214**, 813–815.