

Perceptual sensitivity maps within globally defined visual shapes

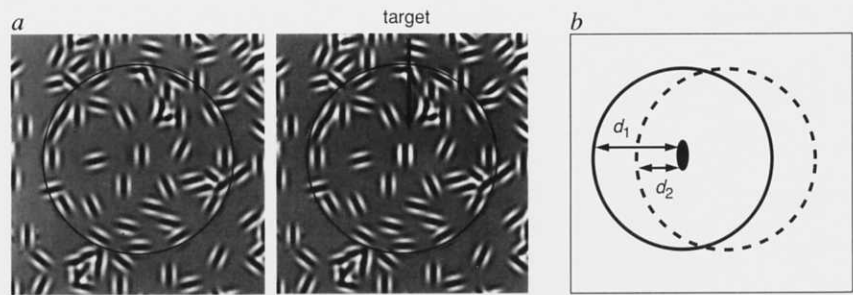
Ilona Kovács & Bela Julesz

Laboratory of Vision Research, Rutgers University,
Busch Campus-Psychology Building, Piscataway,
New Jersey 08854, USA

AN UNSOLVED problem of biology is the processing of global shape in natural vision. The known processes of early vision are spatially restricted (or local) operations, and little is known about their interactions in organizing the visual image into functionally coherent (or global) objects. Here we introduce a human psychophysical method which allows us to measure the effect of perceptual organization on the activity pattern of local visual detectors. We map differential contrast sensitivity for a target across regions enclosed by a boundary. We show that local contrast sensitivity is enhanced within the boundary even for large distances between the boundary and the target. Furthermore, the locations of maximal sensitivity enhancement in the sensitivity maps are determined by global shape properties. Our data support a class of models which describe shapes by the means of a medial axis transformation¹⁻³, implying that the visual system extracts 'skeletons' as an intermediate-level representation of objects. The skeletal representation offers a structurally simplified shape description which can be used for higher-level operations and for coding into memory.

Stimuli were composed of Gabor patches, which are gaussian-modulated sinusoid luminance distributions⁴. Gabor patches model the known receptive-field properties of neurons in the primary visual cortex⁵ (V1). Their oriented shape and small size ensures that the stimuli activate a limited set of early cortical neurons that respond to small parts of the visual field. We measured contrast sensitivity for a Gabor-patch target placed in the context of an enclosing contour and a dense background

FIG. 1 Stimuli for the contrast sensitivity measurements. *a*, Two sequential frames of one trial (only a small central portion of the stimulus frames is shown). Among the randomly positioned and oriented segments, there is an embedded circle (the continuous circle helps the reader to find the contour). The two frames are equivalent except for the contrast of the central element, which is higher in the second frame. The observer's task was to indicate which frame contained the high-contrast target. *b*, Target–contour distance (d_1, d_2) was varied during these experiments. The target was presented on the same central location (in the centre of fixation), and the circle was moved closer to or farther from it.



METHODS. Stimuli were composed of band-limited signals (Gabor patches; GP wavelength (λ) was 0.18° ; gaussian envelope size (standard deviation) was equal to λ ; GP amplitude was 16% of mean luminance). The target was a centrally positioned GP with a contrast increment from the background, and its orientation was parallel to the closest segment of an embedded contour. The background contained 2,500 randomly positioned and oriented GPs in a $16^\circ \times 16^\circ$ field. Contrast increment thresholds for the target were measured as a function of target–contour

distance in a two-alternative temporal forced-choice procedure. One trial consisted of two successive presentations of 170-ms stimulus frames (interframe interval was 500 ms), with the target presented in either the first or second frame. Contrast thresholds were estimated by a staircase procedure¹⁷, four to eight times for each relative target–contour distance. The relative change in sensitivity was calculated by comparing it with the sensitivity for a target with no contour presented. Six human observers participated in the experiments.

field (Fig. 1). To avoid uncertainties about the location of the target, it was always presented on the same central location (in the centre of fixation), and the contour was moved around it. Embedded in the random background, the contour segments are not salient locally. This contour becomes salient only when a chain of roughly colinear segments is connected across the gaps^{6,7}.

Determination of the change in contrast sensitivity within a circular contour as a function of contour–target distance revealed a striking enhancement in sensitivity at the centre of the circle (Fig. 2*a*; see also Fig. 6 of ref 7). The enhancement peak in the centre, where contrast sensitivity for the target increased by a factor >2 , was at 1.44° , eight times the Gabor-patch wavelength (λ) distance from the perimeter. The local effect near the contour (peak at $0.36^\circ = 2\lambda$ distance from the perimeter) is consistent with psychophysically measured perceptual spatial interactions^{8,9}. Contrast threshold measurements for a central Gabor target flanked by two Gabor signals showed

that the maximal spatial range of interaction was $5-6\lambda$ in a colinear configuration⁸, but only $2-3\lambda$ when the three Gabor signals were orthogonal to the virtual line connecting them⁹. As the latter case applies to our configuration, the peak at the centre (at 8λ) results from interactions spreading much farther than the known psychophysically measured local interactions. These long-range interactions are specific for the interior of the contour (not shown in Fig. 2*a*; however, in ref. 7, Fig. 6 presents data for the outside region, where sensitivity enhancement peaked at 2λ distance and was zero at 4λ). Analogous results were reported in the primary visual cortex of the macaque monkey, revealing a strong asymmetry in the cells' responses depending on whether the receptive field was positioned inside or outside a figure^{10,11}. These neuronal correlates of figure–ground segregation suggest an early cortical locus for our long-range, interior-specific interactions.

Why does the remote enhancement peak lie at the centre of the circle? To test for the significance of the equidistant property

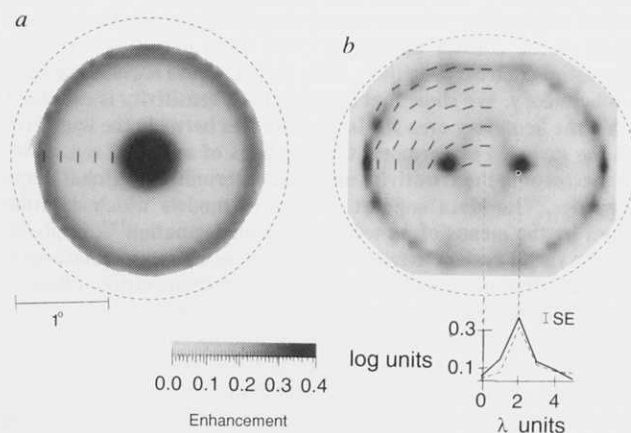
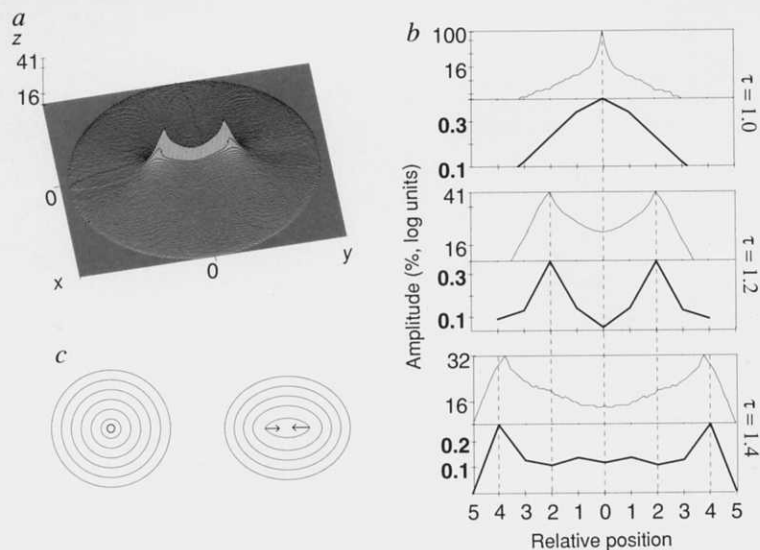


FIG. 2 Maps of contrast sensitivity change for a target within embedded circular and ellipsoid contours. The contours are represented by the dashed lines, and contrast sensitivity change is indicated by different levels of grey, where black stands for sensitivity enhancement >0.3 log units. *a*, Sensitivity change within a circle. The most salient effect was that the circle induced increased sensitivity at its centre, which was at 1.44° distance from the perimeter. Enhancement peak was also observed at 0.36° distance from the contour. The effect of target–contour separation was sampled with 1 target period (λ) resolution (vertical bars represent relative target positions). Averaged results of three observers. (*a* was enhanced from ref. 5, Fig. 6. As we discussed in ref. 5, closed contours are more readily seen than non-closed ones. Here we use perfectly closed contours, and it is not known whether closure is a crucial factor in determining the sensitivity maps.) *b*, Contrast sensitivity change within an ellipse (aspect ratio = 1.2). Note that the centre shows no change in sensitivity, while two peaks can be localized 0.36° away from the centre. Measurements were taken with a resolution of one target period (λ) in one quadrant of the ellipse (superimposed black bars), and the data were reflected symmetrically to the other quadrants to generate the full map. The line graphs show sensitivity change along the major axis for ellipses of different sizes. Ordinate, sensitivity change in log units; abscissa, distance between the target and the centre of the ellipse in λ units. λ was scaled with the size of the ellipse. Dashed line, $\lambda = 0.12^\circ$; 2 observers. Thick line and grey-level map, $\lambda = 0.18^\circ$; 3 observers. Thin line, $\lambda = 0.24^\circ$; 2 observers. Major axes of the ellipses were 16.8λ . Standard error (SE) was averaged across observers. As long as contour–probe distance is expressed in λ units, the overall shape of the sensitivity curves is the same for the different sizes, indicating that the map scales with the figure.

FIG. 3 Peaks of contrast sensitivity can be predicted by the D -function. *a*, D -function (within an ellipse of 1.2 aspect ratio). x and y are spatial positions. D is defined by the percentage of the boundary points equidistant from (x, y) within a tolerance of 1% of the boundary length. The two maxima indicate those locations equidistant from the largest portion of the boundary. *b*, Equidistant points predicted by the D -function (thin lines) for a circle (aspect ratio, $\tau = 1.0$) and two ellipses ($\tau = 1.2$ and 1.4), and corresponding contrast sensitivity data (thick lines). Abscissa, locations along the major axes (the centre of the conic sections is at 0). Ordinate, percentage of equidistant points on log scale (thin lines) and amplitude of sensitivity change in log units (thick lines). There is very strong conformity between the data and the computed equidistant points. Data are means of 3 observers for each conic section. (The D -function was computed by the DAVID visualization software, Fluid Sciences Inc. 1992.) *c*, Propagation of grassfire¹⁻³ within a circle and an ellipse. The grassfire transformation can be imagined as spreading waves of a fire lit along the boundary. Parallels show the succession of fire. A perfect circle collapses into a single point (left panel). A large portion of the waves within an ellipse (right panel) converge in two points along



the major axis simultaneously (corners in the central plum-stone), and the residual waves converge later (arrows pointing towards the centre).

of the centre, we performed measurements on ellipses. Figure 2*b* shows the sensitivity map of an ellipse (aspect ratio = 1.2). When the centre is not equidistant from all the contour segments there is no enhancement in the centre, and two displaced peaks arise along the major axis (at $0.36^\circ = \lambda$ distance from the centre). The relative peak locations were independent of the absolute size of the ellipse (within a range of 2° – 4° diameter; Fig. 2*b*, line graphs). This indicates that the peak locations do not depend on fixed visual angles or tissue separations in the visual cortex.

Obviously, an ellipse contains no location equidistant from all the boundary points. There are, however, singularities approximately equidistant from a large proportion of the boundary points. Figure 3*a* shows one possible way to estimate these equidistant locations within an ellipsoid contour: at every internal point, a ' D -function' is defined as the percentage of boundary points equidistant within a tolerance of 1% of the boundary length. The maxima in this space are displaced from the centre along the major axis. Clearly, the D -function predicts a single peak within the circle. To test whether the D -function can also predict maxima in the ellipse sensitivity maps, we repeated the measurements for an ellipse with 1.4 aspect ratio. Figure 3*b* shows the predicted singularities and the measured maxima for conic sections of 1.0 (circle), 1.2 and 1.4 aspect ratios. Note that the D -function predicts sensitivity maxima for the three conic sections within measurement error.

The success of the D -function in predicting sensitivity peaks suggests an axis-based shape-processing in human vision. As opposed to conventional boundary tracing, an axis-based shape description is derived in a direction orthogonal to the boundary, where a 'skeleton' (medial axis or 'stick-figure') of a shape is made explicit. Blum's¹⁻³ 'grassfire' transformation was the first biologically motivated implementation of this idea. One way to describe the medial-axis transformation is by analogy to how a

grassfire would transverse from the object's boundary into its interior (see Fig. 3*c*). At the last moments, when the 'grass' burns down, the original shape collapses into a skeleton, which consists of quench points equidistant from the boundary. The central peaks in the sensitivity maps we report correspond to the major quench points of the simple, smooth boundaries we used. Small perturbations of more complex boundaries, however, may change the skeleton drastically. To overcome this difficulty, and still provide both the appropriate details and the general form of an object, it has been suggested that the representation should be obtained at various resolutions of different spatial scales¹²⁻¹⁵. Although our sensitivity map data seem to be the first direct evidence for an axis-based shape representation at a given scale, some recent psychophysical data, based on observers' ability to bisect distances between contours, support the multiscale nature of the representation¹⁵.

We have shown for regular two-dimensional planar shapes that local contrast sensitivity is enhanced on particular locations within them. The simplest model that agrees with our results and predicts the locations of maxima in the maps is probably the grassfire model¹⁻³. There are indications that the responsible neural substrates are at the level of the primary visual cortex^{10,11}, but recurrent pathways from higher cortical areas may also participate. As propagation velocities of corresponding cortical areas become known, it will be possible to map our findings from the spatial domain into temporal propagation of neural activity across an extensive medium. Recent findings using real-time optical imaging already provide parameters of a massive cortical activity spread in macaque monkey VI.¹⁶ Our psychophysical findings raise the possibility that long-range lateral interactions behind this activity spread play crucial roles in figure synthesis shape analysis and the extraction of simple spatial relations. □

Received 25 April; accepted 12 July 1994.

- Blum, H. in *Models for the Perception of Speech and Visual Form* (ed. Wathen-Dunn, W.), 362-380 (MIT Press, Cambridge, USA 1967).
- Blum, H. *Perspect. Biol. Med.* **10**, 381-408 (1967).
- Blum, H. *J. theor. Biol.* **38**, 205-287 (1973).
- Gabor, D. *J. IEE Lond.* **93**, 429-457 (1946).
- Marcelja, S. *J. opt. Soc. Am.* **70**, 1297-1300 (1980).
- Field, D. J., Hayes, A. & Hess, R. F. *Vision Res.* **33**, 173-193 (1993).
- Kovács, I. & Julesz, B. *Proc. natn. Acad. Sci. U.S.A.* **90**, 7495-7497 (1993).
- Polat, U. & Sagi, D. *Vision Res.* **33**, 993-999 (1993).
- Polat, U. & Sagi, D. *Vision Res.* **34**, 73-78 (1994).

- Lamme, V. A. F. *Invest. Ophthalm. vis. Sci.* **35**, 1489 (1994).
- Zipser, K., Lee, T. S., Lamme, V. A. F. & Schiller, P. H. *Invest. Ophthalm. vis. Sci.* **35**, 1973 (1994).
- Blum, H. & Nagel, R. N. *Pattern Recogn.* **10**, 167-180 (1978).
- Koenderink, J. J. & van Doorn, A. J. *Biol. Cybern.* **53**, 383-396 (1986).
- Kimia, B. B., Tannenbaum, A. R. & Zucker, S. W. *Int. J. comp. vis.* (in the press).
- Burbeck, C. A. & Pizer, S. M. *Invest. Ophthalm. vis. Sci.* **35**, 1626 (1994).
- Grinvald, A., Lieke, E. E., Frostig, R. D. & Hildesheim, R. J. *Neurosci.* **14**, 2545-2568 (1994).
- Tolhurst, D. & Barfield, L. *Vision Res.* **18**, 951-958 (1978).

ACKNOWLEDGEMENT. We thank Á. Fehér, B. Anderson and J. Nelson for helpful comments.