Constraints on Long Range Interactions Mediating Contour Detection

MARK W. PETTET,*† SUZANNE P. MCKEE,* NORBERTO M. GRZYWACZ*

Received 18 July 1996; in revised form 3 February 1997; in final form 15 May 1997

Contour detection may be mediated by lateral interactions between neighboring cortical neurons whose receptive fields have collinear axes of preferred orientation. This hypothesis was tested in psychophysical experiments and computer simulations using a contour detection task in which observers searched for groups of Gabor patches that followed spatially extended contour paths embedded in noise consisting of several hundred Gabor patches with random positions and orientations. The orientation-selective units in the simulated neural network were linked by facilitatory interconnections whose strength depended on the geometry (distance, curvature, change in curvature) of smooth curves connecting the orientation axes of units in a pairwise fashion. Psychophysical detection performance was much higher for contour signal groups that followed closed rather than open-ended paths. However, just two sudden changes in orientation of neighboring Gabor patch elements in closed-path contours reduced detection performance to the same levels obtained with open-ended contours. These psychophysical data agreed with the results of the neural network simulations. Furthermore, the simulations also accounted for previous findings that removal of a single Gabor patch element from a closed-path contour group significantly degraded detection performance. We conclude that closure alone is not sufficient to enhance the visibility of a contour. However, if a closed contour meets certain geometric constraints, then lateral interactions based on these constraints can generate facilitation that reverberates around the closed path, thereby enhancing the contour's visibility. © 1998 Elsevier Science Ltd. All rights reserved.

INTRODUCTION

Proposed mechanisms for recognition of objects in a static scene often presuppose the segmentation of the scene into regions corresponding to putative objects. In many cases, a region consists of a finite, contiguous area where the statistics of the stimulus pattern in some multidimensional parameter space are roughly uniform. This area of uniformity is enclosed by a border, where the statistics of the stimulus suddenly shift to a different configuration. In more concrete terms, the borders of these putative object regions tend to occur along lines of transition between areas of different luminance, color, texture, or stereoscopic depth.

This paper develops a border-based object segmentation scheme. The general goal of this model was to allow individual, spatially localized, orientation-selective units to facilitate each other, in order to maximize the activity corresponding to spatially extended contours arising from object borders. Briefly, units facilitated each other in a pairwise fashion when their axes of preferred orientation were approximately co-aligned along a common path. The specific goal of this model was to implement a physiologically plausible computer simulation optimized for comparison with the results of psychophysical experiments.

Specialized mechanisms for detection of spatially extended contours can be tested using the following paradigm. Sparsely sampled contours embedded in dense noise consisting of randomly scattered samples having similar local characteristics (Fig. 1) are usually quite difficult to see. Detection performance in search tasks using these stimuli have revealed several properties important for visibility of contours (Uttal, 1983; Smits, Vos & Van Oeefelen, 1985; Beck, Rosenfeld & Ivry, 1989; Field, Hayes & Hess, 1993; Moulden, 1994; McIlhagga & Mullen, 1996). These studies showed that contour element groups were easier to find when the contour path was straight, rather than curved or angled. Detectability also depended on the spacing of contour elements relative to the noise density in a scale-invariant manner. Variation of individual contour element orientations with respect to the overall path of the contour also

*Smith-Kettlewell Eye Research Institute, 2232 Webster St, San Francisco, CA 94115, U.S.A.
†To whom all correspondence should be addressed [Email: pettet@skivs.ski.org].
degraded visibility. Chromatic contrast alone is sufficient to support contour detection, but the combination of chromatic and luminance information for contour detection is contrast dependent.

Recently, using this paradigm, Kovacs and Julesz (1993) demonstrated what they called a “closure enhancement effect”. Using a fixed noise element density, they varied the spacing between Gabor patch elements (see Methods below for definition of “Gabor patch elements”) in various contour signal groups to determine the critical spacing for criterion detectability. This critical spacing was larger for signal groups whose elements fell along closed contour paths rather then open-ended paths. In other words, closed contours were easier to see than open-ended ones. In fact, with closed contour element spacing set to generate 75% correct performance, the removal of just one element from the contour group reduced performance to near chance levels.

Kovacs and Julesz (1993) suggested that a collinear excitatory mechanism might enhance responses to Gabor patch stimuli that follow an extended contour path. They also predicted that sudden changes in orientation, or “kinks” in the contour would disrupt these interactions and degrade visibility. This latter proposal agrees with the orientation specificity of the long-range, “association field” grouping hypothesis proposed by Field et al. (1993). Yet, even as they specifically disavow topological closure as the fundamental principle driving the enhancement effect, Kovacs and Julesz still express some ambivalence, writing, “closure has a global binding effect which makes an otherwise undetectable dashed curve pop out from the background” (1993, p. 7496).

The current study seeks to clarify the mechanism of the closure enhancement effect in contour detection. First, we present psychophysical evidence confirming the disruptive effect of sharp corners on the visibility of closed contours. Second, we verify that a simple computational model based on facilitatory interactions between neighboring oriented filters can explain both the closure enhancement effect and the disruption caused by corners and gaps. We conclude that a contour completely enclosing a subregion of visual space will only have enhanced visibility if it meets certain geometric constraints, and that knowledge about the closure of the contour path is not necessary to generate this enhanced visibility.

**METHODS**

**Psychophysical experiments**

A typical stimulus pattern consisted of several hundred Gabor patch elements, scattered over a viewing area on the screen of a 17” color monitor (Nanao) controlled by a microcomputer (Commodore) equipped with a high-speed graphics card (Merlin). Observers viewed the screen at a viewing distance of 97 cm so that the pixel spacing in the 1024 × 768 viewing area subtended 1 arcmin of visual angle. The overall dimensions of the viewing area at this observation distance were 17.0 deg × 12.8 deg.

Each Gabor patch was a 16 × 16 pixel bitmap whose pixel brightness values, \( L \), at positions \((x, y)\) relative to the center of the patch were scaled by the following equation:

\[
L(x, y) = L_{\min} + \frac{L_0(L_{\max} - L_{\min}) - L_0}{2\pi^2} \times \cos\left(\frac{2\pi(x \sin \theta - y \cos \theta)}{p}\right),
\]

where

- \( L_{\min} \equiv \text{minimum screen luminance} = 30 \text{ Cd/m}^2 \).
- \( L_{\max} \equiv \text{maximum screen luminance} = 150 \text{ Cd/m}^2 \).
- \( L_0 \equiv \text{balancing factor (see below)} \).

The current study seeks to clarify the mechanism of the closure enhancement effect in contour detection. First, we present psychophysical evidence confirming the disruptive effect of sharp corners on the visibility of closed contours. Second, we verify that a simple computational model based on facilitatory interactions between neighboring oriented filters can explain both the closure enhancement effect and the disruption caused by corners and gaps. We conclude that a contour completely enclosing a subregion of visual space will only have enhanced visibility if it meets certain geometric constraints, and that knowledge about the closure of the contour path is not necessary to generate this enhanced visibility.
where

\[ a = \text{standard deviation of circularly symmetric gaussian envelope} = 4 \text{ arcmin}, \]

\[ p = \text{period of sinusoidal carrier} = 8 \text{ arcmin}, \]

\[ \theta = \text{orientation of carrier}. \]

By this definition, the orientation of the bright bar in the center of the Gabor patch was horizontal when \( \theta = 0 \), and the orientation rotated counterclockwise as \( \theta \) became more positive. The spatial frequency bandwidth of these Gabor patches was approx. 1.2 octaves, centered at 7.5 cpd. The background luminance was set to \( L_{\text{min}} + (L_{\text{max}} - L_{\text{min}})/2 \). The balancing factor, \( L_{\beta} \), was added to make the measured mean luminance of the Gabor patches equal the background luminance. The balance of luminance between the Gabor patches and the background was confirmed by viewing the stimulus patterns from a distance of approx. 10 m. All luminances were measured with a Pritchard Spectrophotometer.

Stimulus patterns were drawn by blitting Gabor patch bitmaps chosen from a library of 180 pre-calculated patches indexed by orientation in 1 deg increments. Positions and orientations of the elements in any given pattern were random, except that the individual Gabor patch bitmaps were non-overlapping. A subset of stimulus elements, designated the “signal group”, had positions and orientations fit to a spatially extended contour path (Fig. 1). The specific methods for generating the various signal groups are described individually for each experiment below. The remainder of elements in the pattern were designated “noise”.

The observers (two adult males with normal or corrected-to-normal vision) viewed the monitor from a headrest at a distance of 97 cm from the monitor. On each trial, two stimulus patterns were presented, each for 150 msec, separated by a blank screen for 750 msec. Observers were instructed to hold their gaze steady on the center of the screen and to choose, by pressing a button during the interval between trials while viewing a blank screen. They initiated the next trial by a second button press.

Responses were collected in sessions consisting of 4–6 runs of 100 trials each. Individual data points in the results below were collected from at least two sessions occurring on different days. For any given type of signal group, the percentage of correct responses were collected as a function of noise level (the total number of elements in the stimulus pattern). However, for any given run, the noise level was constant. In some cases (where noted), different types of signal groups were randomly interdigitated in a single run.

Model and computer simulations

The first stage of the model simulates the orientation-selective spatial filtering that occurs in primary visual cortex. For computational convenience, the first stage reduces the number of active units in the network to be equivalent to the number of Gabor patch elements in the stimulus pattern. In principle, this could be accomplished by applying a spatially localized winner-take-all selection process to the outputs of a topographically organized array of orientation-selective units. So, for a given Gabor patch element in the stimulus pattern, there will be one unit that is optimally tuned for the position, orientation, and spatial scale of that element. Neighboring units tuned to nearby positions or similar orientations or scales will also be partially stimulated by this Gabor patch, but only the output of the optimally tuned unit is selected for processing by the second stage of the model.

Thus, the processing in the second stage begins with a discrete set of units, each tuned for a particular position, \((x, y)\), and orientation, \(\theta\), exactly corresponding to the positions and orientations of the stimulus elements. These units were interconnected to form a single-layer feedback network. The responses of the units in the network were driven by the additive combination of three terms. The first term corresponded to the constant input from the previous orientation-selective stage. The second term was the weighted sum of facilitatory (i.e., positive) inputs from all the other units in the network. The third term was a normalizing auto-inhibitory shunt that forced the network to converge to a finite steady state (see Appendix for details). The resulting non-linear differential equation was solved by numerical integration to determine steady-state response values (Burdens & Faires, 1989).

The facilitatory interaction between any two second-stage units was weighted by a product of three factors, whose values depended on the preferred location and orientation of each unit’s receptive field. The first of these factors was simply a gaussian function of the Euclidian distance between the receptive field centers. The second and third factors approximated strength of interaction based respectively on the curvature and changes in curvature of a function fit through the orientation axes of the two units. Strictly speaking, we could have determined these interaction strengths by finding the best-fitting function that minimized these geometric properties (a calculus of variations problem). In practice, however, there are no simple analytical expressions for these curves (Nitzberg, Mumford & Shiota, 1991); thus, simulations based on this approach would be slow. Instead, we derived approximate penalties for total curvature and changes in curvature by fitting a spline through the orientation axes of the two units’ receptive fields (Fig. 2). An index for total curvature was calculated by integrating the square of the second derivative along the length of the spline. Interaction strengths between the units were a gaussian function of this total curvature index. An index for change in curvature was calculated from the third derivative of the spline. Again, interaction strengths were a gaussian function of this index.

In total, the strength of interaction between two units in the second stage of the model was determined by the product of three standard gaussian functions: one for distance, one for total curvature, and one for change in curvature of a spline fit through the two receptive fields. The standard deviations of the gaussians \((\sigma_d, \sigma_c, \text{ and } \sigma_d')\),...
FIGURE 2. Spline-fitting algorithm for determining curvature interaction factors. Schematic representations of the receptive fields of two second stage units are labeled "i" and "j". The short, thin lines extending from the central subregion of each receptive field indicate the preferred orientation axis of the unit. \( \theta_0 \) and \( \theta_1 \) represent the orientation of these axes (for unit i and j, respectively) with respect to the line connecting the two receptive field centers. \( \theta_0 \) and \( \theta_1 \) are used to estimate the total curvature and the derivative of curvature for a spline fit through the orientation axes of the two units (shown as the dotted curve.) These curvature estimates are, in turn, used to compute the weight of interaction, \( w_0 \), between the two units.

respectively) were free parameters that controlled the relative importance these three factors. For example, a small value for \( \sigma_d \) would only allow for strong facilitatory interactions over short distances. Similarly, a small value for \( \sigma_o \) would only allow strong interactions between units whose orientation axes were nearly collinear. Finally, a small value for \( \sigma_d \) would only allow strong interactions between two units whose preferred orientation axes fell on a nearly circular path. The values for the free parameters of the model were chosen to achieve the best possible agreement between the psychophysical and the simulation results (see Discussion for more details).

The third stage of the simulation assessed whether the response distribution to a given stimulus pattern indicated the presence of a signal group. Examples of simulated response distributions to stimuli with and without a contour signal group are illustrated in Fig. 3. For each stimulus pattern, the third stage counted the number of responses that exceeded a non-parametric outlier criterion. In simulated two-alternative, forced-choice trials, the numbers of outliers were compared for 100 pairs of noise-alone and signal-plus-noise stimulus patterns. The proportion of correct responses was estimated using an algorithm based on the theory of signal detection (Green & Swets, 1966).

All simulations were run on a Sun SparcStation 2 stand-alone workstation.

RESULTS

Psychophysical experiments

In the first experiment, to repeat and confirm the basic findings of Kovacs & Julesz (1993), we compared the visibility of open-ended and circular contours. Both types of contour signal group consisted of 12 Gabor patch elements, each element separated from its neighbor (hereafter, local spacing) by 40 arcmin. In the open-ended contours, the difference in orientation between neighboring elements was randomly either +30 deg or –30 deg. Thus, the open-ended contour was characterized by random wiggles back and forth along its length. The difference in orientation of neighboring elements in the circular contour was constant at +30 deg, so that the contour closed back upon itself. As controls, we also tested the visibility of open-ended and circular signal groups, where the orientations of the individual elements were random with respect to the overall contour path. An

FIGURE 3. Network output response distributions. These examples were obtained from simulations, run on stimulus patterns containing 300 Gabor patch elements. The stimulus generating the top histogram contained noise alone, while the bottom histogram’s stimulus also contained a 12-element open-ended contour signal group. The vertical dotted line on the right side of the histograms represents the criterion level for the designation of outlier responses. In this example, note that the secondary peak on the right side of the outlier criterion level corresponded to the units responding to the contour signal elements.
example of an open-ended contour group occurs in the stimulus pattern illustrated in Fig. 1.

The results from two observers (Fig. 4) showed that the circular contours were much easier to see than the open-ended contours over a wide range of noise levels. The control groups were impossible to see except at very low noise levels, and proved that the visibility of the contours was due to the orientational alignment, and not the positional arrangement of the constituent elements.

Field et al. (1993) showed that contour visibility decreased when orientational variation along the length of the contour was increased. Our controls, with randomized orientations represent the most extreme case of such variation. Both Field et al. and Kovacs and Julesz speculated that orientational alignment between neighboring elements was a crucial constraint on local grouping mechanisms mediating contour visibility. Kovacs and Julesz extended this idea by suggesting that one or two sudden changes in local curvature would degrade contour visibility, even when the rest of the contour was fairly smooth, and that this degradation would even offset the closure enhancement effect.

To test these hypotheses, we compared the visibility of three different kinds of closed contours (Fig. 5 and Table 1). Each of the three shapes consisted of 11 elements. The first shape was a circular contour with 76 arcmin radius (spacing between neighboring elements = 43 arcmin, orientation change between neighboring elements 33 deg). The second shape was identical to the circle, except that the first three elements were reflected around the secant separating them from the other elements in the group. Thus, this contour formed a crescent moon shape with changes in orientation of 82 deg at the two corners. We also created a closed contour similar to the crescent moon shape, in which the corners had been smoothed to form a bean-like shape. From trial to trial, the position of the center of the signal group was randomized, but the global orientation of the signal group remained fixed.

Of these three shapes, the circular contours were the easiest to see, the crescent moons were the most difficult to see, and the beans were of intermediate visibility (Fig. 6). ANOVA indicated significant differences between the proportions of correct responses for the three shapes pooled over all observers and noise levels \( (n=108) \), degrees of freedom = 2,106; \( F=13.26; \) \( P=0.0001 \).

Two-tailed \( t \)-tests showed significant differences between these means (circles vs moons, \( P=0.0001 \); circles vs

<table>
<thead>
<tr>
<th>#</th>
<th>Circular</th>
<th>Crescent moon</th>
<th>Bean</th>
<th>Open moon</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31, 69, 155 deg</td>
<td>--</td>
<td>23, 68, 128 deg</td>
<td>--</td>
</tr>
<tr>
<td>2</td>
<td>64, 41, 122 deg</td>
<td>50, 41, 57 deg</td>
<td>38, 38, 90 deg</td>
<td>71, 63, 32 deg</td>
</tr>
<tr>
<td>3</td>
<td>76, 0, 90 deg</td>
<td>38, 0, 90 deg</td>
<td>38, 0, 90 deg</td>
<td>113, 76, 0 deg</td>
</tr>
<tr>
<td>4</td>
<td>64, 41, 57 deg</td>
<td>50, 41, 122 deg</td>
<td>38, 38, 90 deg</td>
<td>154, 63, 147 deg</td>
</tr>
<tr>
<td>5</td>
<td>31, 69, 24 deg</td>
<td>--</td>
<td>23, 68, 128 deg</td>
<td>--</td>
</tr>
<tr>
<td>6</td>
<td>10, 75, 17 deg</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>7</td>
<td>49, 57, 139 deg</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>8</td>
<td>72, 21, 106 deg</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>9</td>
<td>72, 21, 73 deg</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>10</td>
<td>49, 57, 40 deg</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>11</td>
<td>10, 75, 8 deg</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Note: The symbol, "--", indicates that \( (x, y, \theta) \) for the element is the same as for the corresponding element in the circular contour. Numbering of elements described in legend of Fig. 5.
FIGURE 5. Closed contour shapes used to test effects of local curvature changes. Exact coordinates and orientations of individual elements listed in Table 1. The elements designated "#1" in each column of Table 1 respectively correspond to the elements indicated by arrows. The numbering proceeds clockwise around the contour.

beans, $P = 0.0038$; beans vs moons, $P = 0.0033$). Thus, if contour signal group elements fall along a closed path with sharp corners, the closure enhancement effect is reduced.

Finally, we examined whether a common mechanism might explain the reduced visibility of open contours and closed contours with corners. We created an open-ended contour consisting of two circular arcs that were identical to the circular arcs in the crescent moon shape. In the open-ended version of the moon shape, the smaller arc was reflected around a line tangent to the larger arc at the top corner (Fig. 5 and Table 1).

The percentage of correct responses for the open-ended and closed moon shapes were virtually identical for both observers at all noise levels tested (Fig. 7). Sharp corners and overt breaks in closed contours have very similar effects on visibility, suggesting a common mechanism.

Computer simulations

To test whether our model of pairwise facilitatory interactions could explain the results above, we performed the following simulations. For each noise level, a set of 100 stimulus patterns containing noise alone was created. The network responses to these noise patterns were compared with responses to one or more sets of 100 stimulus patterns containing noise plus one of the signal groups tested in the psychophysical experiments. The algorithms for generating the stimulus patterns were the same as those used for the psychophysical experiments.

The results of the simulations agreed well with the qualitative trends seen in the psychophysical experiments. The circular contours were more visible than the open-ended contours [Fig. 8(A)], although the simulations generally predicted smaller differences than seen with human observers (Fig. 4). The circular shapes were easier to see than the bean shapes, which were, in turn, easier to see than the crescent moon shapes [Fig. 8(B)]. However, the differences between the curves for the three shapes were generally larger in the simulation results than in the psychophysical results (Fig. 6). Finally, there was relatively little difference between the open-ended and the closed moon shapes [Fig. 8(C)], although the two curves seemed to diverge slightly at higher noise levels.

We also ran simulations to test whether the model predicted the closure enhancement effect observed by Kovacs and Julesz (1993). In the latter study, psychophysical performance in detecting a circular contour dropped by an average of approx. 16% when one element in the contour was omitted. In contrast, removing the terminal element from an open-ended contour only degraded performance by approx. 1%. To reproduce the conditions of Kovacs & Julesz's experimental paradigm, a few modifications were made to our simulation. First, in the psychophysical studies, the number of noise elements
CONSTRANTS ON LONG RANGE INTERACTIONS MEDIATING CONTOUR DETECTION

Proportion Correct

<table>
<thead>
<tr>
<th>Obs MP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
</tr>
</tbody>
</table>

TABLE 6. Psychophysical visibility of the shapes illustrated in Fig. 5.

![Graph showing proportion correct for different shapes and number of noise elements.]

FIGURE 6. Psychophysical visibility of the shapes illustrated in Fig. 5.

was held constant while the spacing between the elements in the contour signal groups were adjusted so that both the circular and open-ended contours elicited 75% correct responses. In our simulation, the spacing between elements was set to 40 arcmin for circular contours and 35 arcmin for open-ended contours. Second, in the psychophysical study, to control for eccentricity effects, one element of the contour group was always constrained to fall at 1 deg from the fixation point. Thus, although the entire display was filled with Gabor patches, the entire circular contour always fell within 4 deg of the fixation point, reducing the effective search area to approx. 25% of the total display area (16 deg x 16 deg). For computational convenience, we constrained the contour signal group to fall in the center of a 8 deg x 8 deg area containing 500 total elements, thus equaling the density of elements in the psychophysical experiments.

The results in Fig. 9 confirm the closure enhancement effect. The proportion of correct responses for circular contours dropped from 0.72 to 0.59 when one element was removed. In contrast, removing the terminal element from the open-ended contour reduced the proportion of correct responses from 0.69 to 0.64.

DISCUSSION

One reasonable approach to finding the edges of objects in an image is to find luminance contours. Many models in the literature extract this contour/edge information by transforming the input image with differential luminance operators (Marr & Hildreth, 1980; Haralick, Watson & Laffey, 1983). However, despite its theoretical elegance, this approach falters somewhat when applied to real images. Often, continuous contours in the input images are transformed into fractured contours in the outputs (Berzins, 1984). Furthermore, additional open-ended contours are often generated by surface texture. In their simplest form, filter-based contour mechanisms lack methods for binding disjointed contours into borders for correct segmentation.
To solve this binding problem, several strategies have proven useful. One of these strategies employs cooperative and competitive interactions between neighboring filter units. These interactions are spatially organized to enhance signals arising from extended contours, which are assumed to be good candidates for object borders. Examples of this border-based strategy are numerous in the computational literature (Grossberg & Mingolla, 1985; Shashua & Ullman, 1988; Parent & Zucker, 1989; Kellman & Shipley, 1991; Nitzberg et al., 1991; Heitger & von der Heydt, 1993). The second strategy involves grouping together, usually by some diffusive process, neighboring subsets of the image that have similar visual features (Cohen & Grossberg, 1984; Kumaran, Geiger & Gurvits, 1996). This region-growing process can be limited by the outputs of border-based algorithms described above, or can be used as an independent check of these outputs. However, in the human visual system, segregation is not likely to be driven strictly by region-growing processes, since such a scheme would not allow for the visibility of open-ended contours in our experiments.

Our model critically relies upon assumptions about functional connectivity between units separated by large retinotopic distances. Anatomical evidence for this connectivity has been reviewed elsewhere (Gilbert, 1992). Briefly, the topographic projection of a typical classical receptive field (CRF) in primate V1 spans approx. 1.5 mm (Hubel & Wiesel, 1974). This span approximately agrees with the combined spread of thalamic afferents to layer 4 and the dendritic fields of their putative targets (Lund, 1973; Freund, Martin, Soltész, Somogyi & Whitteridge, 1989). In contrast, axon collaterals from layer 5 pyramidal cells reach across 6–8 mm of cortex (McGuire, Gilbert, Rivlin & Wiesel, 1991; Rockland & Lund, 1982, 1983; Amir, Harel & Malach, 1993). These collaterals could allow excitatory monosynaptic interactions (Hirsch & Gilbert, 1991, 1993) between units spatially separated by four–six receptive field diameters. Since these collaterals seem to span even larger distances in extrastriate cortex (Amir et al., 1993), feedback from these other visual areas could enable interactions between units separated by even

FIGURE 7. Psychophysical visibility of closed and open-ended moon shapes.
larger distances. And, of course, polysynaptic relays within V1 could extend lateral connectivity indefinitely.

To accompany and complement this anatomical evidence, many physiological studies on single units in mammalian visual cortex have demonstrated the modulatory effects of stimuli located outside the CRF (Allman, Miezin & McGuinness, 1985). Many of these effects were suppressive and confined to the near periphery surrounding the CRF, although stimuli in the far periphery sometimes influenced the overall sensitivity

FIGURE 8. Simulation results from stimulus patterns used in psychophysical experiments. (A) Compare with Fig. 4. (B) Compare with Fig. 6. (C) Compare with Fig. 7.
contours agree with the closure enhancement effect this manipulation was much larger than expected. In fact, predicted by those authors, sharp corners reduced the stimulus patterns (Polat & Norcia, 1996). Have been observed in evoked potentials elicited by these collinear and separated by a distance 2-3-times the wavelength of the targets carrier wave. The effect seemed in a foveal Gabor target, provided that the target and the flanks were matched the target. Recently, comparable effects to peak when the orientation and the spatial scale of the preferred orientation axis of the CRF being studied would disinhibit or even facilitate the unit’s activity (Nelson & Frost, 1985; Kapadia, Ito, Gilbert & Westheimer, 1995). This latter finding suggests a tantalizing link with some of the anatomical data about the long-range horizontal axon collaterals described above. Specifically, terminal arbors from these collaterals tend to cluster in remote cortical regions populated by cells having the same preferred orientation as the cell projecting into the axon (Gilbert & Wiesel, 1989).

A possible psychophysical link to these physiological findings comes from studies on how flanking stimuli influence contrast sensitivity for static luminance Gabor patches (Polat & Sagi, 1993, 1994). Specifically, Gabor patch flanks enhanced contrast sensitivity for a foveal Gabor target, provided that the target and the flanks were collinear and separated by a distance 2-3-times the wavelength of the targets carrier wave. The effect seemed to peak when the orientation and the spatial scale of the flanks matched the target. Recently, comparable effects have been observed in evoked potentials elicited by these stimulus patterns (Polat & Norcia, 1996).

Our results with open-ended and circular closed contours agree with the closure enhancement effect described by Kovacs and Julesz (1993). Further, as predicted by those authors, sharp corners reduced the visibility of the closed contours. However, the effect of this manipulation was much larger than expected. In fact, closed contours with sharp corners and comparable open-ended contours were equally visible. Rounding off these corners partially compensates for this lost visibility. Our model, based on lateral interactions between oriented units, accounts for these observations.

Although we designed our model to explore how contour visibility was influenced by curvature, we have found that it explains the properties of the closure enhancement effect, as follows. First, the strength of interaction between any two units in the model is determined by the retinotopic distance between them. Consider the response of a unit responding to an element in a closed, circular contour. This unit will receive strong facilitation from two nearly co-aligned units, one on each side, corresponding to the nearest two elements in the contour. Because of the uniform spacing and curvature around the contour, this pattern of bidirectional facilitation is reiterated for each element in the group. Also, since the interaction between a pair of units is recursive \([A(1)]\), facilitation communicated to a neighboring unit is reflected in the facilitatory input returned from that unit. Given a closed loop with appropriate geometry, such reverberation could propagate around the contour. Now, if one of the elements is omitted from the contour, the response levels of units neighboring the gap will drop. Despite the residual facilitation between units on either side of the gap, the reverberation of activity is disrupted and the visibility of the contour drops. By the same reasoning, the exceedingly large gap between the terminal elements of an open-ended contour prevents any reverberation at all. Corners have a similar effect on interactions between units. In this case, the sudden change in curvature reduces interactions between the two units responding to elements on either side of the corner. Again, the overall extent of facilitation is disrupted, and the contour becomes less visible than a circular one. While rounding off the corners will partially store some of the facilitatory interactions, there are still penalties from changes in curvature at the corners of the bean shape that do not occur in the circular contour.

Even though our model describes the data fairly well, there are alternative interpretations of the results. One of these interpretations concerns the role of overall shape. When we added corners to the circle to create our crescent moon, we also changed the shape of the contour. Likewise, the bean stimulus was different in shape from either the circle or the moon. Thus, while our model attributes differences in visibility to the differences in curvature at the corners of the three stimulus patterns, the differences in overall shape could also partially account for the data.

Although our model is somewhat sensitive to overall shape, the effect on interactions between units is small and somewhat unpredictable. If the standard deviation for the distance factor is large enough to allow interactions between elements in the contour that are three or four elements removed from each other, then changes in shape will certainly cause changes in the interactions between these distant neighbors. Of course, this explanation
predicts a very small contribution of overall shape, since the weight of interaction between such distant neighbors will be much smaller than for the nearest neighbors. Furthermore, a cubic spline fit through two distant elements will only roughly follow the true path of the contour between them. Thus, some information about the shape of the true path will be lost.

There are additional design features of our implementation worth discussing when interpreting the simulation results. First, for computational convenience, the simulation only calculates the interactions between a limited subset of units in the network. Each Gabor patch element in a given stimulus pattern will elicit a distribution of responses in a local population of first stage units tuned to a range of stimulus parameter values. To reduce the number of calculations, first stage responses are fed forward to the second stage only if they are optimally stimulated by one of the Gabor patch elements in the stimulus pattern. In essence, the second stage only computes interactions between units whose preferred orientations and positions correspond to the elements in the stimulus pattern. We assume that, while first stage units tuned to nearby positions and similar orientations will also be partially activated by any given stimulus element, the activity level of these neighboring units will always be less than the activity level of the optimally tuned unit.

The second feature to note about the simulation regards the architecture of the second stage of the model. We have implemented the interactions in this stage using a single-layer, feedback network. Again, this choice was motivated by computational convenience, and was not meant to imply any predictions about the actual cortical connectivity subserving the behavior described by our model. In particular, all the connections between units in the second stage were monosynaptic, which is almost certainly untrue in the corresponding cortical architecture. A more plausible cortical architecture would probably include connectivity between different layers and even between different cortical areas. However, it is still quite plausible that the effects of this inter-layer and inter-areal connectivity would be manifest in the response properties of the unit located in the very early stages of processing.

Third, the simulation does not include architecture corresponding to the short-range excitatory and inhibitory connectivity observed in primary visual cortex (Levitt, Lund & Yoshioka, 1996; Stemmler, Usher & Niebur, 1995). These interactions probably shape the local tuning properties of individual cortical neurons. Further, the simulation lacks any provision for facilitation or suppression of neighboring units falling along an axis orthogonal to a unit’s preferred orientation. Thus, our simulation does not reproduce some of the lateral interactions demonstrated by Polat and Sagi (1994) and Kovacs and Julesz (1993). Our model also excludes factors related to eccentricity and spatial scale. Although these omissions diminish physiological realism, our simulation was specifically designed to test how long-range interactions might be related to extended contour geometry. Inclusion of additional architecture would enlarge the dimensionality of the model’s parameter space. While additional free parameters might allow better agreement between psychophysical and simulation results, finding the parameter set that minimizes this discrepancy would be more difficult. Also, the behavior of a more complex model would not necessarily be more stable to perturbations in parameter settings than the model presented here. Thus, the architecture in the model represents a compromise between physiological realism and computational parsimony.

Fourth, aside from response changes during convergence to steady-state, the simulation did not provide for temporal modulation of facilitatory interactions in the network. Thus, we cannot address whether propagation of reverberatory facilitation around a closed contour generates any dynamic behavior relevant to the closure enhancement effect. However, in a model proposed by Yen and Finkel (1996), contour visibility was determined by the degree of dynamic response synchronization between spatially remote local mechanisms. Their model also predicted enhanced visibility for circular contours compared with open-ended ones. Although our model and theirs used similar geometric constraints for computing interactions between local orientation-selective filters, their model was designed for natural image processing rather than psychophysical simulation. Thus, the extent of quantitative agreement between the two models is unknown.

Finally, our model and its implementation were developed to analyze human performance in the contour detection task described in the previous section. Strictly speaking, this was a search task, since the primary source of uncertainty was in the position of the signal group. However, we have made no attempt to model visual search, per se. Instead we assumed that success in the psychophysical task was directly related to the relative strength of responses elicited by the stimulus elements in the signal groups and the noise. In this context, “search” can be viewed as the process employed to sample from the overall population of responses. Although we did not specify the details of this process, we assumed that it was uniform across the different experiments we conducted, and thus could be factored out of our analysis. However, for the simulation of Kovacs & Julesz’s closure enhancement effect (Fig. 9), the change in size of the search window might account for the differences in the two data sets.

As noted in the description of the results above, there are some differences between the psychophysical and the simulation results. Data from several noise-free simulations (data not shown) account for how quantitative performance of the simulation varies between the different contour types. Although an exhaustive review of these findings is beyond the scope of this paper, several performance trends were noteworthy.

For example, changes in the interaction parameters ($\sigma_d$, $\sigma_\alpha$, and $\sigma_\omega$) will push apart the curves for open-
ended and closed contours [Fig. 8(A)]. Specifically, raising $\sigma_\epsilon$ and lowering $\sigma_{\epsilon'}$ will place a heavy penalty on interactions between open-ended contour elements that are three or more steps removed from each other. Since open-ended contours tend to turn back and forth, widely separated elements on these contours are more likely to be fit by splines having changes in curvature. Thus, a heavier penalty on changes in curvature will have a greater effect on open-ended contours than circular ones.

However, heavy penalties for changes in curvature have a severe impact on the visibility of the bean-shaped contours, rendering them less visible than the moon shapes. The bean shapes differed from the other shapes in a way that accounted for this observation. The network units with the largest response values to the bean stimulus corresponded to the elements in the flat part of the bean. This was mostly due to the perfect collinearity of the three elements in this part of the contour. The units in the flat part of the bean were also partially driven by elements in the corners of the bean. However, this interaction was extremely sensitive to the penalty for changes in curvature. Increasing this penalty virtually eliminated interactions with the units in the flat part of the bean. The result was a sudden drop in visibility relative to the other shapes.

To preserve the order of the circles, beans, and moons, we set the penalty for changes in curvature to relatively low levels. However, this parameter setting generates another discrepancy between the simulation and the psychophysics. Careful inspection of the data shows that the open-ended contours (Fig. 4) were much harder to see than the moon shapes (Fig. 6). However, the simulation results show the converse relationship (Fig. 8). In general, adjusting interaction parameters to improve fits for a specific set of curves tended to degrade fits for other curves. This suggests that our simulation was not a good model for shape discrimination. Perhaps there are different algorithms for estimating geometric penalties that better capture the shape dependencies of contour viability.

Despite its limitations as a shape discriminator, our model does answer some of the basic questions we have posed about the importance of closure and its relationship to contour geometry. More generally, our model instantiates a general class of models based on lateral interaction. It shares with these models several properties that make it quite intriguing in the context of object segmentation. It enhances the detectability of spatially extended contours at an early stage of processing, allowing extraction of potential object border information before higher-level mechanisms come into play. Furthermore, the degree of enhancement can be used as a cue to the geometric structure of the contour (i.e., whether it has gaps, or corners, or free ends). This latter property could be quite useful in the processing of occluded surfaces. According to our model, in the immediate vicinity of a T-junction between two surfaces, the visible part of the occluded surface is delimited by a contour with a sharp corner occurring at the point of occlusion. Although the contour of the occluding surface is partially coextensive with the contour of the occluded surface, the occluding contour will lack the sudden change in curvature that occurs at the corner in the occluded contour. Thus, facilitation along the length of the occluded contour is disrupted at the occlusion point, while the facilitation along the occluding contour is unaffected by the occlusion point. The resulting difference in responses corresponding to the two contours provides a direct, low-level cue to designate which surface occludes which.

In conclusion, lateral interactions between neighboring units in early stages of cortical visual processing may be a powerful mechanism for quickly extracting perceptually relevant information about the spatial structure of static images. Even though more sophisticated models may provide better quantitative fits to the data, the current results strongly suggest that the very limited constraints imposed by ours should be part of any model (no matter how realistic) that attempts to describe these visual processes. However, given the enormous potential of this approach, we are confident that continued psychophysical and computational studies along these lines will advance our understanding of human visual processing.

REFERENCES


Greysware, N. M., Harris, J. M. & Amthor, F. R. (1996). Computational...


Lund, J. S. (1973). Organization of neurons in the visual cortex, area 17, of the monkey (Macaca mulatta). Journal of Comparative Neurology, 147, 455–496.


Acknowledgements—This work was supported by NIH EY06672-01 and AFOSR F49620-95-1-0265.

APPENDIX
The computer simulation

Processing in the second stage begins with n units (where e = the number of Gabor patch elements in the stimulus pattern), each tuned for a particular position, (x,y), and orientation, θ, corresponding to the positions and orientations of the elements in the stimulus pattern. These units were interconnected to form a single-layer feedback network. The response, r, of unit i in the network was driven by the following time-differential equation:

\[ \frac{dr_i}{dt} = 1 + \alpha \sum_{j=1}^{n} w_{ij} \beta \theta^2, \]  

where and are positive free parameters, and \( w_{ij} \) are the positive-valued connection weights between unit i and unit j. The first term on the right side of the equation corresponds to the constant input from the previous orientation-selective stage. The second term is the weighted sum of facilitatory (i.e., positive) inputs from all the other units in the network. The third term is a normalizing auto-inhibitory shunt that forces the network to converge to a finite steady state. Steady state response values were estimated by integrating \( [A(1)] \) using the Runge-Kutta algorithm (Burden & Faires, 1989). This algorithm automatically reduced the magnitude of the time steps to insure that all responses remained positive as they approached steady state.

The facilitatory input from unit j to unit i, was weighted by the coefficient, \( w_{ij} \), whose value depended on the preferred location and orientation of each unit’s receptive field. This value was calculated using the following formulae:

\[ w_{ij} = G(z_{ij} \times G(z_{ij} \times G(z_{ij})), \]  

where

\[ G(z) = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{z^2}{2} \right), \]

\[ z_{ij} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \]

\[ \sigma_k = \tan^{-1}(\theta_i) + \tan^{-1}(\theta_i) + \tan^{-1}(\theta_i). \]
than \( o \) is normally distributed with mean \( 0 < \omega < \omega_m \). For sufficiently large \( n \), the number of inputs stronger
\[
np(1 - p) \leq \sigma^2 \text{.}
\]
Let the constant \( q > 1 \) be chosen so that the strength of facilitatory input from any other unit is larger than
\[
\omega_m - (2n)^{3/2} \left[ A(2) \right].
\]
As \( n \to \infty \), \( m \to \infty \), and the right-hand side of \( A(9) \) will become positive, causing \( p \) to grow without bound. Now consider the behavior of a network whose responses obey \( A(8) \). Each unit will have more than \( m \) inputs, and the strength of these inputs will be larger than \( \omega \).

Thus, \( dp/dt > dp/dt \), and \( r_i > \omega \). Therefore, \( r_i \) grows without bound.

In contrast, the non-linear shunt in \( A(1) \) will lead to bounded responses over time, even when \( n \) increases. Consider the network that obeys:
\[
dp/dt = 1 + \rho \omega (n - 1) \text{max} - \beta p \text{.}
\]
As \( A(10) \) is integrated over time, \( p \) increases until the value of \( dp/dt \) eventually reaches zero. Thus \( p \) is bounded. Units in a network obeying \( A(1) \) will have fewer than \( n - 1 \) inputs, and the weights of these inputs will be less than \( \omega_{\text{max}} \). Thus, \( dp/dt \) and \( r_i \) must also be bounded.

Finally, until the network converges, the individual responses, \( r_i \), in the network will always be increasing, and thus, positive. At time \( t = 0 \), \( dr_i/dt = 1 \) and \( r_i \) will increase. If, at time \( t = T, dr_i/dt = 0 \), then the curve will not change. However, if, at time \( t = T, dr_i/dt > 0 \) for any \( i \neq j \), then \( r_i \) will increase, making \( dr_i/dt \) positive, causing \( r_i \) to increase again. This positive feedback will only cease when all the units converge to steady-state responses, so that \( dr_i/dt = 0 \) for all \( i \).

The relative distribution of responses in the network were largely unaffected by the specific values of the parameters \( \alpha \) and \( \beta \) [Eq. \( A(1) \)], provided the following conditions were met. First, \( z \) had to be large enough relative to the input to overcome the effect of the shunt. Second, \( \beta \) had to be small enough relative to \( \alpha \) to allow convergence in a reasonable number of integration steps. Since the decision algorithm in our model depended only on the relative distribution of responses, not their absolute values, network performance was robust over a wide range of values for \( \alpha \) and \( \beta \).

The third stage of the simulation assessed whether the response distribution to a given stimulus pattern indicated the presence of a signal group in the stimulus. A response, \( r_i \), was designated an outlier if it met the following criterion [Spriet, 1993]:
\[
r_i - Md(R)/Md(R - Md(R)) > 4.0 \text{.}
\]
where \( Md(R) = \text{median of the set } R = \{r_1, r_2, \ldots, r_n\} \).

In the simulation, the third stage counted the number of outliers in the responses to noise-alone and signal-plus-noise stimulus patterns. To estimate the proportion, \( C \), of correct responses for \( m \) trials, the simulator used the following modification of Green and Swets’ [Green and Swets, 1966] Eq. (2.6):
\[
C = \sum_{j=1}^{m} \frac{1}{m} \sum_{j=1}^{m} P(k_n, k_s) \text{,}
\]
where \( k_n \equiv \text{number of outliers in noise trial } j \) and \( k_s \equiv \text{number of outliers in signal trial } i \).

The original Green and Swets’ formula expressed the proportion of correct responses in terms of likelihood that the internal response of the observer was elicited by a stimulus containing a signal. These likelihoods were estimated by the ratio, \( k \), of the gaussian-distributed probabilities that the internal response was generated by signal + noise or noise alone. The proportion of correct responses for \( m \) APF trials is the integral over \( k \) of the joint probabilities of obtaining likelihoods.
from the signal + noise intervals, while obtaining likelihood less than \( k \) from the noise alone intervals. Our formula used the number of outliers, as an estimate of \( k \). In essence, the simulator runs a series of trials in which a single signal + noise stimulus is compared with a single noise alone stimulus. For each trial, the stimulus with the larger number of outliers is called the signal + noise stimulus. If both stimuli have the same number of outliers, the simulator chooses randomly between the two. The proportion of correct responses is calculated for \( m \times m \) trials consisting of all possible pairings of \( m \) signal + noise and \( m \) noise alone stimuli.

In summary, the free parameters for all the simulations in the text were:

- \( z = \) weight of lateral interactions = 10,
- \( \beta = \) weight of non-linear shunt = 0.1,
- \( \sigma_d = \) standard deviation of gaussian factor for distance = 50.0 arcmin,
- \( \sigma_k = \) standard deviation of gaussian factor for total curvature = 1.0,
- \( \sigma_c = \) standard deviation of gaussian factor for change in curvature = 0.6,
- outlier criterion = 4.