Effects of element separation and carrier wavelength on detection of snakes and ladders: Implications for models of contour integration

Keith A. May
McGill Vision Research, Department of Ophthalmology, McGill University, Montréal, Québec, Canada

Robert F. Hess
McGill Vision Research, Department of Ophthalmology, McGill University, Montréal, Québec, Canada

In this paper, we examine the mechanisms underlying the perceptual integration of two types of contour: snakes (composed of Gabor elements parallel to the path of the contour) and ladders (with elements perpendicular to the path). We varied the element separation and carrier wavelength. Increasing the element separation impaired detection of snakes but did not affect ladders; at high separations, snakes and ladders were closely matched in difficulty. One subject showed no effect of carrier wavelength, and the other showed a decline in performance as the wavelength increased. We discuss how these results might be accommodated by association field models. We also present a new model in which the linkage results from overlap in the filter responses to adjacent elements. We show that, if 1st-order filters are used, the model’s performance on widely spaced snake contours deteriorates greatly as the carrier wavelength of the elements decreases, in contrast to our psychophysical results. To integrate widely spaced contours with short carrier wavelengths, the model requires a 2nd-order process, in which a nonlinearity intervenes between small-scale 1st-stage filters and large-scale 2nd-stage filters. This model detects snakes when the 1st and 2nd stage filters have the same orientation, and detects ladders when they are orthogonal.

Keywords: contour, snakes, ladders, association field, filter, 2nd-order


Introduction

Objects in the real world are bounded by contours that can twist and turn across considerable distances within the visual field. Local orientations can also vary greatly along the length of a contour. Since the receptive fields of cells in the early visual cortex are spatially localized and orientation-tuned (Hubel & Wiesel, 1959, 1962, 1968), the outputs of these cells must be linked in some way in order to achieve perceptual integration of contours. In this paper, we examine the mechanisms that underlie the integration of two types of contour: snakes and ladders.

These contour stimuli were introduced by Field, Hayes, and Hess (1993). They both consist of smooth paths of Gabor elements embedded in a background of identical elements with random position and orientation. In snake stimuli, the contour elements form tangents to the path of the contour whereas, in ladder stimuli, the contour elements are perpendicular to the path. The terms “snake” and “ladder” were introduced by Bex, Simmers, and Dakin (2001).

Ladders have generally been found to be harder to detect than snakes (Bex et al., 2001; Field et al., 1993; Hess, Ledgeway, & Dakin, 2000; Ledgeway, Hess, & Geisler, 2005; May & Hess, 2007a, 2007b). However, each study has used only a small set of stimulus parameters. In the experiment described here, we undertook a systematic sampling of the parameter space.

Methods

Subjects

Two male observers, BCH and KAM, participated in the experiment. KAM had corrected-to-normal vision and had not participated in a contour integration experiment before; BCH had normal vision without correction and had participated in one previous study on contour integration (Hansen & Hess, 2006).

Apparatus

The experiments were run on a Dell PC with a VSG 2/5 graphics card (Cambridge Research Systems). Experiments
were controlled using software written in MATLAB (The MathWorks, Inc.). Images were generated using C routines called from MATLAB. The images were linearly scaled to fit the range 0–255 and stored in an 8-bit frame store on the VSG card. Stimuli were then scaled to the correct contrast and gamma corrected by mapping the 8-bit values onto 15-bit values. An analogue input to the monitor was generated from these 15-bit values using two 8-bit digital-to-analogue converters on the VSG card. Stimuli were displayed on a Sony G520 monitor at a frame rate of 120 Hz. The graphics card was configured to generate a screen display that measured 1024 pixels horizontally by 769 pixels vertically. Subjects viewed the screen binocularly from a distance of 60 cm.

**Stimuli**

On each trial, one interval contained a stimulus consisting of a path of odd-symmetric Gabor elements embedded in a grid of similar Gabor distractors; the other interval contained a stimulus consisting only of distractors. Examples of the contour stimuli are shown in Figure 1.

The Gabor elements were generated using Equation 1:

\[
L(x,y) = L_0(1 + cw),
\]

where \(c\) is the carrier, and \(w\) is the envelope, as defined in Equations 2 and 3, respectively:

\[
c = C \sin[2\pi(x \cos \theta + y \sin \theta)/\lambda],
\]

\[
w = \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right).
\]

\(L\) is the luminance at position \((x, y)\), measured from the center of the Gabor patch; \(L_0\) is the mean (background) luminance of 52 cd/m\(^2\); \(C\) is the Michelson contrast, which was set to 0.9; \(\lambda\) is the wavelength of the Gabor carrier; \(\sigma\) controls the “width” of the Gaussian envelope; and \(\theta\) is the orientation of the element from vertical.

We generated the stimuli using the same general algorithm as Beaudot and Mullen (2003). The stimulus area was divided into an invisible \(10 \times 10\) grid of squares. For stimuli containing no contour, each grid square was filled with one randomly oriented Gabor element, placed at a random location within the square, subject to the constraint that each element should be placed at a distance of at least \(h/2\) from its nearest neighbor, where \(h\) is the height (and width) of a grid square. This ensured that there was never more than a slight overlap between the elements. Overlapping values were simply added (note that this addition occurred before the contrast function, \(cw\), was converted to luminance using Equation 1). For stimuli containing a contour, the contour was first positioned randomly within the grid, and then each remaining empty grid square was filled with one element with random orientation and random position within the grid square, as with the no-contour stimulus.

Part of a snake contour is represented schematically in Figure 2. The contour was constructed along an invisible backbone of ten line segments, joined end-to-end. A contour element was placed at the center of each segment. The distance between adjacent elements along the contour is referred to as the element separation, \(s\). For snake contours, the element was oriented parallel to the segment; for ladders, the element was orthogonal to the segment.

The absolute difference in orientation between adjacent segments is referred to as the path angle. The sign of this difference was random for each pair of adjacent segments. For each pair, the path angle was jittered by adding a random value uniformly distributed between \(\pm 10^\circ\).
Table 1 gives a summary of all the stimulus parameters used in the experiment. The width of the grid squares was set to $2s / (1 + \sqrt{2})$, which ensured that the mean separation between adjacent distractors was close to the element separation, $s$ (Beaudot & Mullen, 2003). For more details of the contour generation algorithm, see Beaudot and Mullen (2003).

**Procedure**

Gabor element separation and carrier wavelength were varied in half-octave steps. The parameter values are given in Table 1. Each combination of carrier wavelength, element separation, and contour type (snake or ladder) was tested in two separate sessions. Within a session, only path angle was varied and took values of $0^\circ$, $10^\circ$, $20^\circ$, $30^\circ$, or $40^\circ$. Overall, there were 40 trials for each combination of parameter values. Each trial consisted of two stimuli, presented sequentially. One stimulus contained a contour, and the other contained only distractor elements. The order of stimuli was determined randomly. Each stimulus was presented for 500 ms, separated by an inter-stimulus interval of 1000 ms, during which a small fixation dot was displayed on a background of uniform luminance, equal to the mean (background) luminance of the contour stimuli (52 cd/m$^2$). Subjects were allowed to move their eyes during the presentation of the stimuli. On each trial, the subject had to indicate, using a button box, which interval contained the contour. After each trial, the subject received auditory feedback to indicate whether the response was correct or incorrect.

Since the snake and ladder conditions were tested in separate sessions, the subjects knew which type of contour they were supposed to be looking for. This avoided the potential dilemma that a subject would have had if they had detected the target but had also detected a randomly occurring contour of the other type in the other interval. If

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grid size (in terms of number of cells)</td>
<td>$10 \times 10$</td>
</tr>
<tr>
<td>Element contrast</td>
<td>0.9</td>
</tr>
<tr>
<td>Carrier spatial frequency (c/deg)</td>
<td>5.19, 3.67, 2.59, 1.83</td>
</tr>
<tr>
<td>Carrier wavelength, $\lambda$ (deg visual angle)</td>
<td>0.193, 0.273, 0.385, 0.545</td>
</tr>
<tr>
<td>$\sigma$ (deg visual angle)</td>
<td>0.136</td>
</tr>
<tr>
<td>$\lambda/\sigma$</td>
<td>1.41, 2, 2.83, 4</td>
</tr>
<tr>
<td>Separation, $s$ (deg visual angle)</td>
<td>1.09, 1.54, 2.18, 3.08</td>
</tr>
<tr>
<td>Width of a single square within the grid, for each separation value (deg visual angle)</td>
<td>1.28, 1.81, 2.55</td>
</tr>
<tr>
<td>$s/\lambda$</td>
<td>5.66, 4.00, 2.83, 2.00</td>
</tr>
<tr>
<td>$s/\lambda$</td>
<td>8.00, 5.66, 4.00, 2.83</td>
</tr>
<tr>
<td>$s/\sigma$</td>
<td>11.3, 8.00, 5.66</td>
</tr>
<tr>
<td>Path angle</td>
<td>$0^\circ$, $10^\circ$, $20^\circ$, $30^\circ$, $40^\circ$</td>
</tr>
<tr>
<td>Path angle jitter</td>
<td>Uniform probability between $\pm 10^\circ$</td>
</tr>
<tr>
<td>Orientation jitter</td>
<td>None</td>
</tr>
<tr>
<td>Separation jitter</td>
<td>None</td>
</tr>
<tr>
<td>Stimulus duration</td>
<td>500 ms</td>
</tr>
<tr>
<td>Inter-stimulus interval duration</td>
<td>1000 ms</td>
</tr>
</tbody>
</table>

Table 1. Stimulus parameters used in the experiment. In the table, the values of $\lambda$ are displayed in a row, and the values of $s$ are displayed in a column. This gives rise to a matrix of $s/\lambda$ values, where each row corresponds to a particular value of $s$, and each column corresponds to a particular value of $\lambda$. The rows and columns of this matrix correspond to the rows and columns of panels in Figures 3 and 4.
the subjects had been biased towards one type of contour, then the detection levels for the other contour type would have been underestimated.

Results

Figures 3 and 4 show the proportion of correct responses in each condition for subjects BCH and KAM, respectively. While performance on ladders was generally lower than on snakes, the difference between the two contour types was smaller for the higher separations (the bottom rows of Figures 3 and 4).

To quantify the effects of separation and carrier wavelength, we collapsed the data across path angle to give a single proportion-correct score for each panel of Figures 3 and 4. Since this score was determined from a range of path angles (and hence difficulty levels), it provided a sensitive measure of overall performance that was away from floor and ceiling in every condition for both snakes and ladders. This score is plotted as a function of separation in Figure 5, and as a function of carrier wavelength in Figure 6.

Figure 5 shows that performance on ladders was unaffected by separation, whereas performance on snakes declined substantially with increasing separation, so that...
snakes and ladders were quite closely matched at the highest separation. Figure 6 shows that increasing the carrier wavelength led to a slight decline in performance for KAM, but not BCH (except, perhaps, for snakes with the smallest two separations levels).

To confirm these observations statistically, we found the Pearson product moment correlation coefficients of performance level against separation and wavelength for snakes and ladders. Each correlation used the sixteen data points for each subject and contour type plotted in Figures 5 and 6. For each subject there were four correlations to carry out, and the criterion of significance for each correlation, \( a_1 \), was chosen such that the type I error probability across all four correlations, \( a_n \), was equal to 0.05. If there are \( n \) correlations, then

\[
a_1 = 1 - (1 - a_0)^{1/n}.
\]  

(4)

For \( n = 4 \), \( a_1 = 0.0127 \). For the correlations of performance against separation, we took each set of four scores for a particular contour type and wavelength (i.e., each set of connected data points in Figure 5) and divided each score within the set of four by the mean of the set. This normalization reduced any variance due to differences in performance across wavelength, leading to a purer measure of the effect of separation. Similarly, for the correlations of performance against wavelength, we took each set of four scores for a particular contour type and separation (i.e., each set of connected data points in Figure 6) and divided each score within the set of four by the mean of the set.

The correlation coefficients are given in Table 2 (second column from the right). Both subjects showed very strong negative correlations of performance against separation for snakes. For ladders, neither subject showed a significant correlation of performance against separation. The subjects differed in the effect of carrier wavelength. KAM’s performance declined significantly with increasing wavelength for both snakes and ladders, whereas BCH showed no significant effect of wavelength.

The right-hand column of Table 2 shows analogous correlations performed on the unnormalized data (i.e., the raw scores plotted in Figures 5 and 6). The pattern of significance and nonsignificance is the same as for the normalized data, except for KAM’s correlation between performance and wavelength for snakes, which is not significant in the unnormalized data. Figure 6 indicates that this was because the large effect of separation gave rise to a high variance in the snake data, which would have masked any effect of wavelength. Normalizing the data for each level of separation removed this source of

\[
\begin{array}{|c|c|c|c|}
\hline
\text{Subject} & \text{Independent variable} & \text{Contour type} & \text{Pearson correlation of normalized performance vs. independent variable} & \text{Pearson correlation of unnormalized performance vs. independent variable} \\
\hline
\text{BCH} & \text{Separation} & \text{Snake} & r = -0.904 \ (p = 1.5 \times 10^{-6})^* & r = -0.866 \ (p = 1.4 \times 10^{-5})^* \\
 & \text{Separation} & \text{Ladder} & r = -0.554 \ (p = 0.026) & r = -0.529 \ (p = 0.035) \\
 & \text{Wavelength} & \text{Snake} & r = -0.572 \ (p = 0.021) & r = -0.271 \ (p = 0.31) \\
 & \text{Wavelength} & \text{Ladder} & r = 0.0988 \ (p = 0.72) & r = 0.0770 \ (p = 0.78) \\
\hline
\text{KAM} & \text{Separation} & \text{Snake} & r = -0.915 \ (p = 6.8 \times 10^{-7})^* & r = -0.837 \ (p = 5.2 \times 10^{-5})^* \\
 & \text{Separation} & \text{Ladder} & r = -0.0146 \ (p = 0.96) & r = -0.0140 \ (p = 0.96) \\
 & \text{Wavelength} & \text{Snake} & r = -0.801 \ (p = 1.9 \times 10^{-4})^* & r = -0.384 \ (p = 0.14) \\
 & \text{Wavelength} & \text{Ladder} & r = -0.818 \ (p = 1.1 \times 10^{-4})^* & r = -0.814 \ (p = 1.2 \times 10^{-4})^* \\
\hline
\end{array}
\]
variance, allowing the effect of carrier wavelength to be detected.

In the above analysis of the effect of separation, we used the absolute separation, in degrees of visual angle. An alternative would have been to use separation divided by wavelength. However, as shown in Figure 7, the graph of performance against this measure is greatly affected by the carrier wavelength, so it is not possible to define a single function that maps this measure onto performance. Figure 5 shows that (except for KAM’s performance on ladders), the graphs overlap quite well, indicating that detection performance is largely a function of absolute separation. Because of this, absolute separation was used in the analysis of the effect of separation on performance.

BCH’s lack of an effect of carrier wavelength mirrors the results of Dakin and Hess (1998, Experiment 1). Like us, they varied carrier wavelength ($\lambda$) for snake stimuli, while keeping the Gabor envelope size ($\sigma$) constant. Their four longest-wavelength stimuli had the same ratios $\lambda/\sigma$ as the stimuli in our four different wavelength conditions (i.e., our Gabor micro-patches were scaled copies of theirs). They found that, over this range, the carrier wavelength had little effect, although performance was impaired when the carrier frequency was very high (>8 c/deg). They speculated that this might be because the critical variable for determining contour detectability is separation expressed in units of $\lambda$. However, our results, plotted in Figure 7, rule out this idea.

Discussion

Our subjects’ ability to detect snakes gradually declined with increasing element separation, as found previously (Beaudot & Mullen, 2003; Field et al., 1993). In contrast, the separation between the elements made little difference to performance on ladders. At small separations, snakes were much easier to detect than ladders, whereas at large separations, performance levels on the two types of contour were quite closely matched. Another finding was that the performance was largely a function of absolute separation rather than separation expressed as a multiple of wavelength. In the following subsections, we discuss how these findings might be accommodated by two very different classes of contour integration model: association field models and filter-overlap models.

Association field models

Field et al. (1993) proposed that, at each point in the visual field there is an association field that links features together. There have been few attempts to implement an association field model that can integrate both snake and ladder contours. One such model, which accounted for a wide range of psychophysical data, was described by Yen and Finkel (1996, 1997, 1998). Their model contained two sets of facilitatory links, one favoring co-axial (snake) configurations and the other favoring trans-axial (ladder) configurations. Within each set, the association strength fell as Gaussian functions of element separation and deviation from co-circularity. Mutually facilitated units developed synchronized temporal oscillations, and the model grouped temporally synchronized units into a single contour.

Yen and Finkel (1998, p. 728) stated that “As the average separation between all elements increases, the degree of facilitation is decreased. However, since inputs from background elements also decrease, the signal-to-noise ratio is not altered.” In Appendix A, we show that this is unlikely to be precisely true; indeed, Yen and Finkel’s model did show a slight decline in snake-detection performance with increasing separation, approximately in line with Field et al.’s (1993) data. Yen and Finkel only tested their model up to a separation of 0.9° visual angle; it is likely that the model’s performance on snakes would deteriorate noticeably as the separation increased to around 3° visual angle, as in our experiment, although only a full simulation would establish this for certain.

A possible problem with Yen and Finkel’s (1996, 1997, 1998) model is that the strength of the facilitatory connections between trans-axial units fell more sharply with distance than for co-axial units. Our finding that only snake detection was affected by element separation suggests that the opposite may be true. Yen and Finkel did...
not give any specific reasons why trans-axial association strength should fall more sharply with distance—the parameters were chosen to optimize performance on just one example stimulus. It is not clear whether or not the model’s performance would be badly compromised by making the trans-axial association strength fall more gradually with distance.

A relatively slow fall in facilitatory connection strength with distance between trans-axial units may be a general feature of early visual coding. In an investigation of the flanker facilitation effect, Cass and Spehar (2005a, 2005b) varied the target-flanker separation and, for each separation, found the shortest stimulus duration for which flanker facilitation was found. They found that this critical duration increased with separation, with a much more gradual increase for the trans-axial than co-axial configurations. This was interpreted as indicating that the facilitatory signal propagates across the cortex far more quickly for the trans-axial configuration. However, as suggested by Li Zhaoping (personal communication to May & Hess, 2007a), Cass and Spehar’s results could have resulted from a more gradual fall-off in strength of facilitatory connection with distance for the trans-axial configuration. We do not wish to suggest that contour integration and flanker facilitation are mediated by the same mechanism since there is ample evidence against this view (Huang, Hess, & Dakin, 2006; Williams & Hess, 1998). However, both systems may show a more gradual drop in connection strength between trans-axial than co-axial units.

In Yen & Finkel’s model, the fall-off in association strength with distance was fixed, giving fixed association field sizes, whatever the element separation. An alternative view, proposed by Pelli, Palomares, and Majaj (2004), is that, at each point in the visual field, there are integration fields of different sizes; these integration fields serve a similar purpose to Field et al.’s (1993) association fields, and the visual system chooses the integration field with the most appropriate size for the given task and stimulus. Pelli et al. argued that the minimum integration field size is proportional to eccentricity so that, when observing a stimulus in the periphery, the visual system may be forced to choose an inappropriately large integration field. They argued that this accounts for many of the characteristics of crowding, whereby identification of a peripherally located target letter is disrupted by the presence of flanking letters.

May and Hess (2007b) borrowed this idea to explain their finding that, unlike snakes, ladders are undetectable at very small eccentricities. May and Hess’s model formed strong associations between co-axial (snake) elements and weak associations between trans-axial (ladder) elements. In the periphery, the larger association fields led to greater interference, which was much more disruptive to the weak ladder associations than to the stronger snake associations.

If there was no upper limit on the size of the association fields, then increasing the element separation should not have a disruptive effect. The different effects of separation on snakes and ladders might therefore indicate that the upper limit was being reached only for snakes in our experiment.

Another possible reason why ladder detection was not disrupted by an increase in the element separation is that increasing the separation may have released the elements from the crowding effect proposed by May and Hess (2007b). In their model, the disruption of ladder detection in the periphery is a crowding effect, caused by the fact that the elements are too close together to be processed properly by the large peripheral association fields. Increasing the separation between the elements might stop this crowding effect, and this could compensate for any deterioration in integration performance due to the increased separation. Since ladders were more severely affected by crowding than snakes in the model, ladder detection might benefit the most from the release from crowding, explaining why only ladder detection was immune to any disruptive effects of increased separation.

One aspect of our current data that we have not discussed yet is the finding that increasing the carrier wavelength of the stimulus elements had either no effect (BCH) or led to a small decrease in performance (KAM): Increasing the width of the bars in the Gabor elements did not seem to increase the distance over which the contours could be integrated. The lack of an effect of feature width on linking performance shown by BCH fits well with the proposal that contour integration and crowding are mediated by similar mechanisms. A key characteristic of crowding is that the critical target-flanker spacing necessary for crowding to occur is independent of the size of the target or flanker (Pelli et al., 2004). This suggests that the range of integration/association field sizes available is not determined by the size or scale of the features to be integrated.

In May and Hess’s (2007b) association field model, it was assumed that the orientation of each element was determined prior to the linking process. This model could accommodate the difference between BCH and KAM in the effect of carrier wavelength if we assume a greater level of noise in KAM’s orientation channels. Increasing the carrier wavelength while keeping the envelope constant has the effect of increasing the orientation bandwidth of the elements. In a system with little noise, the peak of response across orientation would still be close to the veridical orientation but, in a noisy system, the widening of the orientation bandwidth could substantially increase the error in the estimation of element orientation. In May and Hess’s model, this would have the same effect as jittering the actual orientations of the elements, and this is known to disrupt contour detection performance (Bex et al., 2001; Field et al., 1993; May & Hess, 2007a).
In both May and Hess’s (2007b) model and that of Yen and Finkel (1996, 1997, 1998), there is competition between snake and ladder associations. In Yen and Finkel’s model, the sum of inputs to a unit from co-axial connections is compared with that from trans-axial connections, and the weaker set of inputs is suppressed. In May and Hess’s model, links between elements are inserted in order of association strength, so that the stronger associations tend to capture the elements, preventing weaker associations from doing so; in this way, associations compete with other associations of the same type as well as associations of different type. Both of these models can be considered to process the stimulus using two types of association field, one for snakes and one for ladders. This raises the question of whether there is any top-down control of the type of association field that is employed. In May and Hess’s (2007b) experiments, snake and ladder stimuli were randomly interleaved within a session, so the visual system had to process the stimulus with both snake and ladder association fields. In the current experiment, snake and ladder stimuli were blocked into separate sessions. If the observer was able to use that knowledge to selectively suppress the inappropriate type of association field, then that would yield a performance benefit, because it would prevent interference from irrelevant associations. On the other hand, it might be that the observer was unable to suppress inappropriate associations; in this case, any benefit derived from blocking the stimuli would occur at the decision stage (when the observer can ignore randomly occurring contours of the wrong type), rather than at the processing stage. Zhaoping and May’s (2007) finding that task-irrelevant stimulus components can substantially disrupt performance on visual search and segmentation tasks suggests that the extent of top-down influence on low-level grouping processes is limited.

Filter-overlap models

In filter-overlap models, the linkage occurs purely because the filter responses to adjacent elements overlap. Hess and Dakin (1997, 1999) implemented a filter-overlap model that linked filter responses within individual orientation channels. The output of each filter was thresholded, giving rise to zero-bounded regions (ZBRs), and the model performed the contour integration task by looking for the longest ZBR. Hess and Dakin showed that the model was much worse at detecting curved contours than human subjects viewing the stimuli in the fovea.

The filter-overlap model’s ability to detect highly curved contours can be improved by allowing it to link spatially overlapping ZBRs from adjacent orientation channels. The filtered images from the different orientation channels can be visualized as a single 3D representation, which has a dimension representing the filter orientation, in addition to the two spatial dimensions of the image. ZBRs within this representation can be defined in the same way as in Hess and Dakin’s (1997, 1999) model, except that the ZBRs extend across orientation, as well as space. The visual cortex seems well suited to implementing such a mechanism because orientation preferences of neurons vary systematically across the cortex so that cells from adjacent orientation channels are physically adjacent (Bartfeld & Grinvald, 1992; Blasdel, 1992; Blasdel & Salama, 1986; Bonhoeffer & Grinvald, 1991, 1993; Grinvald, Lecce, Frostig, Gilbert, & Wiesel, 1986; Hubel & Wiesel, 1974, 1977; Swindale, Matsubara, & Cynader, 1987).

The next two subsections describe implementations of orientation-linking filter-overlap models for snake and ladder detection. These models are specific examples of the general class of grouping algorithm outlined by Rosenholtz, Twarog, and Wattenberg (2007). In this class of algorithm, the image is represented in a multidimensional space containing the two spatial dimensions of the image, along with other dimension(s) representing local orientation, luminance, color, etc. The multidimensional representation is then blurred so that strongly active points with similar positions within this space join up to form continuous regions, which can then be projected back into the original image space to form grouped regions within the image. This gives rise to grouping by proximity and similarity.

An orientation-linking filter-overlap model for snake detection

The model first filters the stimulus with elliptical Gabor kernels with a general form described by Equation 5:

\[
 f(x, y) = \exp\left(-\frac{u^2}{2\sigma_u^2} - \frac{v^2}{2\sigma_v^2}\right)\cos(2\pi u/\lambda). (5)
\]

\(\lambda\) is the carrier wavelength; \(x\) and \(y\) represent the horizontal and vertical displacement from the center of the kernel; and \(u\) and \(v\) represent displacement perpendicular to and parallel to the bars of the sinusoidal carrier, respectively. If the carrier has an orientation of \(\theta\) from vertical, then \(u\) and \(v\) are given by

\[
 u = x \cos \theta + y \sin \theta, \\
 v = y \cos \theta - x \sin \theta. (6)
\]

\(\sigma_u\) and \(\sigma_v\) are the standard deviations of the envelope perpendicular to and along the bars of the carrier, respectively.
Figure 8. An orientation-linking filter-overlap model for snake detection. The stimulus (top) contains a snake contour, which is highlighted in red for presentational purposes only. The separation is 1.09° visual angle, the carrier wavelength is 0.39°, and the path angle is 30°. The contour wiggles left and right along a vertical trajectory before bending round in an arc. The stimulus was processed with the 1st-order filter-overlap model for snake detection described in the text. The carrier wavelength of the Gabor filter kernels was half an octave above that of the elements. Four differently orientated kernels are illustrated to scale above the corresponding filtered images (labeled “Filter output”). The next row shows the ZBRs obtained by thresholding the filtered images. ZBRs in adjacent orientation channels were then joined together across orientation if they overlapped spatially. The bottom image shows the orientation-linked ZBRs collapsed across orientation. The color of each ZBR in the bottom image reflects the number of pixels in the ZBR after collapsing across orientation (so that pixels in the overlap between orientation channels were not counted twice): The largest ZBR is colored red, and the remaining ZBRs are colored in shades of gray, such that the brightness increases with increasing size. The ZBRs in the individual orientation channels are all colored white except for those that form part of the longest orientation-linked ZBR, which are colored red. Note that only the zig-zagging part of the contour produces a long ZBR within an individual orientation channel: The smoothly curved section gives rise to small ZBRs which cannot be discriminated from the background if processing occurs separately within each orientation channel, as previously noted by Lovell (2002, 2005). But when spatially overlapping ZBRs are linked across orientation, all of these small ZBRs become part of one long ZBR that traces out the shape of the contour and is easily discriminated from the background on the basis of its size alone.
To allow the filter to respond strongly to the stimulus elements, the wavelength of the filter kernel’s carrier should be close to that of the elements. To give the filter kernel the best chance of bridging the gap between the stimulus elements (which was up to 16 times the carrier wavelength of the stimulus elements), it had to be quite long for a given carrier wavelength, so we set the ratio $\sigma_v/\lambda$, to approximately the highest physiologically plausible value. To arrive at this value, we examined the data of Jones and Palmer (1987), who fitted a Gabor model to the receptive fields of simple cells in the cat’s striate cortex. We selected the cell with the highest ratio $\sigma_v/\lambda$ and, for this cell (labeled 0811 by Jones & Palmer, 1987), we noted the values of the ratios $\sigma_u/\sigma_v$ and $\sigma_u/\lambda$, which were 2.41 and 0.340, respectively. We fixed the corresponding ratios in our filter kernels at these values so that, apart from the orientation, $\lambda$ was the only free parameter of the kernel.

The stimulus image was filtered with Gabor Kernels with 24 different orientations, $\theta$, from 0° to 172.5°, differing in steps of 7.5°. Examples of the filtered images are shown in Figure 8. A threshold was set at 2.2 standard deviations above the mean of the filtered image values across all orientations (the expected value of this mean was zero). Values above the threshold were set to 1; all other values were set to 0. This created a set of ZBRs in each orientation channel, also shown in Figure 8.

At this point, the model was virtually identical to that of Hess and Dakin (1997, 1999). The difference was that our model had an extra stage, in which ZBRs in adjacent orientation channels were linked together across orientation if they overlapped spatially. In this stage of the algorithm, the first orientation channel (0° from vertical) was considered to be adjacent to the last (172.5° from vertical). This is because orientation is a naturally cyclic property: Stepping 7.5° from 172.5° gives 180°, which is the same as 0° for even-symmetric filters. Again, the organization of the visual cortex provides an ideal basis for implementing such a mechanism because cells with different orientation preferences are laid out in a circular “pinwheel” rather than a linear arrangement: Preferred orientation changes gradually from 0° to 180° in a full 360° rotation about the pinwheel, so the “first” and “last” orientation channels are physically adjacent (Bartfeld & Grinvald, 1992; Bonhoeffer & Grinvald, 1991, 1993). Figure 8 shows that this orientation-linking procedure allows the model to form a continuous ZBR that traces the outline of a highly curved snake contour.

The ZBRs in the orientation-linking filter-overlap model can be considered to occupy a 3D space formed by taking the 2D ZBR images from the different orientation channels, and stacking them on top of each other: The resulting space has the two spatial dimensions of the image, and a third dimension representing the filter orientation. In the bottom image of Figure 8, this 3D space is viewed from “above,” so that it is collapsed across orientation. Figure 9 shows some different views of the space. This figure shows a schematic representation of two ZBRs from Figure 8: one (in red), which corresponds to the target contour, and another (in green) that spatially overlaps it. Figure 9D shows that the reason that these spatially overlapping ZBRs do not join up to form a single ZBR is that, although they overlap spatially, they are well separated along the orientation axis, and never come into contact. To aid visualization, Movie 1 shows a flyby movie of these two ZBRs.

The ability to segment spatially overlapping contours is a major difference between the model presented here and

Figure 9. (A) The largest ZBR from the bottom image of Figure 8 (in red), and another ZBR (in green), which spatially overlaps it. (B) A schematic representation of the two ZBRs in A, formed by joining the centers of the elements that are overlapped by each ZBR. (C, D) Two different views of the 3D space in which the ZBRs are created. In C and D, the vertical axis represents the local orientation. These views show that, although the two ZBRs overlap spatially, they are well separated along the orientation axis. The thick red and green lines join the centers of the elements, as in B, and provide schematic representations of the ZBRs. The thin red and green lines parallel to the orientation axis show the projections of the element positions onto the 2D image plane. The orientation axis spans the range of values of orientation, $\theta$, such that 0° $\leq \theta < 360°$. Because orientation is a cyclic property, the “top” layer of the space is considered to be adjacent to the “bottom” layer, and the choice of which orientation should correspond to the bottom layer is arbitrary. In this figure, the bottom layer of the space corresponds to 24.7°, which is the orientation of the last-but-one element on the smoothly curved end of the red contour. At the location of this penultimate element, the contour dives into the floor of the space and reappears from the ceiling. The continuity of the ZBR is maintained because the floor and ceiling are considered to be adjacent.
the adaptive filtering model proposed by Dakin (1997). Dakin’s model filters the image with a range of differently oriented filters, as in our model, but then selects, at each image location, the response of the most active filter. The resulting response image is thresholded to form ZBRs but, by this stage, the information about filter orientation has been lost and cannot be used to segment the ZBRs. In Figure 9A, every location within the overlapping red and green ZBRs is above threshold in at least one filter, so these locations would form a continuous above-threshold region in Dakin’s adaptive filtering model, giving rise to a single ZBR.

So far in our description of the orientation-linking filter-overlap model, we have only considered a positive polarity channel, which forms ZBRs corresponding to bright regions of the filtered image. A negative channel, which forms ZBRs corresponding to dark regions, can be created by reversing the sign of all the filter outputs and applying the same thresholding and ZBR-finding procedure as before.

The threshold value of 2.2 was determined, by trial and error, to be low enough to allow linkage to occur, but high enough to prevent large numbers of spurious elements from linking to the target contour. With this threshold level, about 97% of the pixels were set to zero within each orientation channel in Figure 8. This might seem excessive but, when the ZBRs were collapsed across orientation, to project them back into the image space (as shown at the bottom of Figure 8), only about 80% of the pixels took a zero value. If the stimulus itself is thresholded at 1% of its maximum value, around 80% of the pixels are set to zero, so the amount of activity in the ZBR image is not inconsistent with the amount of activity in the stimulus. Figure 10 shows the effect of reducing the threshold to 1.5 standard deviations: Many of the spurious contours join up to the target contour.

**An orientation-linking filter-overlap model for ladder detection**

We also developed a filter-overlap model to detect ladders. This model was the same as the snake-detection model, except that the 1st-order filters were replaced with 2nd-order filters of the kind used in many models of texture segregation (e.g., Graham, 1991; Graham, Beck, & Sutter, 1992; Graham & Sutter, 1998; Graham, Sutter, & Venkatesan, 1993; Sutter, Beck, & Graham, 1989; Lin & Wilson, 1996; Wilson, 1993). In these models, a small-scale linear filter is followed by a nonlinearity (usually squaring or full-wave rectification), followed by a large-scale linear filter with orientation orthogonal to that of the small-scale filter. Such a mechanism gives a strong response at a texture border (such as a border between two areas of different orientation) but a weak response elsewhere. This kind of mechanism can be used to detect ladders: A small-scale filter kernel aligned with the elements will give strong positive and negative responses along the path of the ladder contour; if these responses are rectified or squared, then there will be a positive response along the contour, which will be picked up by a large-scale filter orthogonal to the small-scale filter.

In the model, all the filter kernels were parameterized in the same way as for the snake detection model, so that the

---

**Figure 10.** Outputs of filter-overlap models with the threshold set to 1.5 standard deviations above the mean, rather than the value of 2.2, which was used in all the other figures in this paper. (A) The 1st-order snake detector, from Figure 8. (B) The ladder detector, from Figure 11. (C) A 2nd-order snake detector, described in more detail later; in this case, the stimulus was the same as that in Figure 8 (and panel A of this figure), and the model was the same as in Figure 11 (and panel B of this figure), except that the 2nd-stage filter in each orientation channel had the same orientation as the 1st-stage filter.

---

**Figure 11.** An orientation-linking filter-overlap model for ladder detection. The stimulus (top) is the same as in Figure 8, except that the elements of the target contour (highlighted in red) are rotated by 90°. The 1st-stage filter kernels had a carrier wavelength half an octave below that of the elements. These are illustrated to scale above the corresponding 1st-stage filter outputs. These filtered images were then squared and filtered with Gabor kernels with a carrier wavelength half an octave above that of the elements, i.e., identical filters to those in Figure 8. In each channel, the 2nd-stage filter was orthogonal to the 1st-stage filter. ZBRs were then formed and linked across orientation as before, and these are displayed in the same way as in Figure 8.
Stimulus

1st-stage filter output

2nd-stage filter output

ZBRs

ZBRs linked across orientation
\[ \lambda_{\text{element}} = 0.19^\circ \]

\[ \lambda_{\text{filter}} = 2^{1} \times \lambda_{\text{element}} \]

\[ \lambda_{\text{filter}} = 2^{1.5} \times \lambda_{\text{element}} \]

\[ \lambda_{\text{filter}} = 2^{2} \times \lambda_{\text{element}} \]
only free parameters were the carrier wavelengths of the
1st-stage and 2nd-stage filter kernels. The image was
filtered with small-scale filters with the same 24 orienta-
tions that were used in the snake detection model, and
then the output of each filter was squared, and filtered with
a large-scale filter orthogonal to the small-scale filter.
ZBRs were found in the same way as before, by setting a
threshold of 2.2 standard deviations above the mean.
Unlike in the 1st-order snake-detection model, the
expected mean was slightly above zero because neither
the input to the 2nd-stage filters nor the even-symmetric
Gabor filter kernels had a zero mean: In both cases, the
mean was positive. A further difference between the
models was that, in the ladder-detection model, there was
only one polarity channel, since the responses to the
contour after squaring were always positive.

Figure 11 shows that this method successfully traces out
the shape of the contour; some background elements are
also linked to the contour, but an inspection of the
stimulus reveals that these background elements do indeed
form randomly occurring ladders that adjoin the target
contour, so it is not inappropriate for the model to link
these elements to it.

The effect of element separation on the filter-overlap
model

In the description of the snake-detection model, it was
noted that the ratio of \( \sigma_e \) (filter kernel length) to kernel
carrier wavelength was set at the highest physiologically
plausible value, to maximize the model’s ability to
integrate the contour when the element separation was
much larger than the element carrier wavelength. But even
though we maximized this ratio, Figure 12 shows that,
when the element wavelength is set at the smallest value
used in the experiment (0.19°), the model can only
integrate the contour with the smallest separation, and
for that, requires a filter with a carrier wavelength 1.5
octaves above that of the elements. If the filter kernel size
is increased by another half-octave, then its carrier
wavelength is so large that it cannot respond significantly
to the elements, and the result is largely noise.

Figure 13 shows what happens when the element carrier
wavelength is doubled. Again, it is the filter kernel with a
carrier wavelength 1.5 octaves above the element wave-
length that does best at integrating widely spaced
contours. Because the element carrier wavelength in
Figure 13 is twice that in Figure 12, the filter kernels in
Figure 13 can be enlarged to about twice the size before
they stop responding significantly to the elements, so they
can integrate elements separated by twice the distance. In
other words, the largest separation that can be integrated is
an approximately constant multiple of the carrier wave-
length (it is not an exact multiple because the element
envelope did not increase with the carrier wavelength). It
is clear that the model would predict that (1) integration of
snake contours should be very poor at the large separa-
tions, and (2) integration performance should be largely a
function of element separation expressed as a multiple of
carrier wavelength: If separation is expressed in absolute
units (e.g., degrees of visual angle), performance should
improve with increasing carrier wavelength. But Figures 5
to 7 show that this is not the case. Firstly, performance is
still substantially above chance even at the highest
separation; secondly, Figure 6 shows that increasing the
wavelength either has no effect (BCH) or leads to worse
rather than better performance (KAM), and a comparison
of Figures 5 and 7 shows that performance is largely a
function of absolute separation rather than separation
expressed as a multiple of carrier wavelength.

The reason that the snake detection model fails so badly
when the element separation is much greater than the
carrier wavelength is that it is not physiologically
plausible to have a large 1st-order receptive field that
responds to high spatial frequencies. This restriction does
not apply to the 2nd-order ladder detection model. In this
model, the outputs of small-scale filters are squared, and
then a large-scale filter is applied. This allows the model
to integrate high-frequency elements over a large distance,
as shown in Figure 14.

The success of the ladder-detection model at integrating
contours in which the element separation is much larger
than the carrier wavelength suggests that a similar
approach could be used for snakes. We set up a model
identical to the ladder detection model, except that the
large-scale filter had the same orientation as the small-
scale filter instead of being orthogonal to it. A similar
arrangement explains many psychophysical findings relat-
ing to edge detection and the perception of edge blur and
contrast (Georgeson, May, Freeman, & Hesse, 2007; May
& Georgeson, 2007a, 2007b), and it seems reasonable to
suppose that this kind of sequence of processes might
also underlie the related task of contour integration. The

---

Figure 12. The top row shows example snake stimuli with the four
different element separation values used in the experiment. The
contour elements are highlighted in red for display purposes only.
The carrier wavelength of the elements (\( \lambda_{\text{element}} \)) was the smallest
value used in the experiment (0.19° visual angle). The orienta-
tions of the elements are identical for the four stimuli: Only the
separation differs. A small amount of Gaussian white noise was
added to the image, to prevent the filters responding to the
elements in cases where the signal-to-noise ratio in a biological
system would be very low. The noise had a standard deviation of
0.01, and the signal values (i.e., the stimulus values before adding
the noise), which are given by \( \text{cw} \) in Equation 1, ranged from
−0.94 to +0.94. The bottom three rows show the results of
processing these slightly noisy stimuli with the 1st-order snake
detection model illustrated in Figure 8. Each row shows the
results of using a different carrier wavelength for the filter kernels
(\( \lambda_{\text{filter}} \)). These values were 1, 1.5, and 2 octaves above \( \lambda_{\text{element}} \).
The output images are displayed in the same way as in the bottom image of Figure 8.
Figure 13. The same as Figure 12, except that the carrier wavelengths of the stimulus elements and the filter kernels are twice as large.
2nd order filters in this model are sensitive to high-frequency elements collinear with the contour, and are large enough to integrate them over long distances, as shown in Figure 15.

In summary, snake detection can be supported by both the 1st- and 2nd-order mechanisms at low element separations and by the 2nd-order mechanism alone at high separations. Ladder detection requires a 2nd-order mechanism at all separations. Since the 2nd-order mechanisms can easily accommodate a wide range of element separations, this goes some way towards explaining why ladder detection is unaffected by separation, while snake detection is best at small separations (when both 1st-and 2nd-order mechanisms can function), but declines to approximately the performance level on ladders at high separations (when only a 2nd-order mechanism carry out the task).

Although this explanation has its attractions, it would still predict that performance would improve with increasing
stimulus element carrier wavelength, in contrast to our findings shown in Figure 6. This is because, as the carrier wavelength increases, the 1st-order mechanism can support performance over a wider range of separations (compare Figure 13 with Figure 12). This suggests that 1st-order mechanisms are not involved at all. It seems that, if filter-overlap mechanisms are used for contour integration, both snakes and ladders are integrated using 2nd-order mechanisms of the kind used in Figures 14 and 15. In this case, the different effects of separation on snakes and ladders could be accounted for by a similar explanation to that which we used for association field models: The upper limit on the scale of the 2nd-order filter may be lower for the snake mechanism.

Figure 15. The stimuli (top row) are the same snake stimuli as in Figure 12. The carrier wavelength of the elements ($\lambda_{\text{element}}$) is 0.19° visual angle. The bottom two rows show the results of processing these stimuli with a model the same as the ladder detector in Figure 14, except that the large-scale filter in each orientation channel has the same orientation as the small-scale filter. As in Figure 14, the smaller 2nd-order filter performs best on the two lower-separation stimuli, while the larger filter performs best at the higher separations.

Explaining the effects of carrier wavelength with a filter-overlap model

Figure 6 shows that performance either stays the same (BCH) or declines (KAM) with increasing carrier wavelength. The lack of an effect of wavelength is easily accommodated by the 2nd-order model: For a given separation, the visual system can keep the 2nd-stage filter...
contours using a filter-overlap mechanism, it is unlikely that the poorer performance on long-wavelength stimuli resulted from a reduced tolerance to noise for these stimuli.

Another possible reason for the reduced performance on the long-wavelength stimuli is that KAM might not have properly selected the wavelength of the 1st-stage filter to match the element profile. If this is the case, we can only speculate as to why it might have occurred. Perhaps, because the elements were still small, despite their larger wavelength, KAM used a 1st-stage filter that was inappropriately small for these elements. Another possibility is that it gets harder to estimate the appropriate scale as the number of visible bars of the carrier decreases.

The issue of scale selection will probably have to be addressed before a full implementation of the filter-overlap model can be achieved. As the simulations in Figures 12–15 and 17–19 show, different filter scales are appropriate for different stimuli, so it may be necessary for the model to have a way of choosing which scale to use in a given context. Alternatively, the problem of scale selection might be avoided by combining the outputs across scale, as Watt and Morgan proposed for edge coding (Morgan & Watt, 1997; Watt & Morgan, 1985). Another possibility is that the problem of scale selection or combination is dealt with prior to contour integration: Dakin and Hess’s (1999) finding that contours consisting of alternating Gabor and step-edge micro-patterns are hard to detect suggests that contour integration does not occur independently within separate spatial frequency channels.

Finally, we should consider how the model selects the threshold that gives rise to the ZBRs. All the simulations in this paper used the same threshold level of 2.2 standard deviations above the mean response level. It might be possible to improve the model’s performance, and accommodate top-down effects, by allowing the model to alter the threshold over space and time (see Geisler, Perry, Super, & Gallogly, 2001, p. 721, for a similar suggestion).

Conclusions

Our experiment produced three main findings:

1. increasing the separation between the elements had a disruptive effect on the detection of snakes but had no effect on ladders, so that as separation increased, performance on the two contour types converged;
2. in most cases, performance was largely a function of absolute separation rather than separation expressed as a multiple of the carrier wavelength of the stimulus elements; and
3. increasing the carrier wavelength had no effect on one subject and caused a decline in performance for the other.
Both association field models and filter-overlap models should be able to accommodate these findings, but the results constrain the sets of possible models. In particular, the finding that performance did not improve with increasing carrier wavelength shows that the visual system was not using a 1st-order filter-overlap mechanism to integrate these contours. A 2nd-order filter-overlap model looked much more promising. Our model begins with a standard “filter-rectify-filter” sequence of operations: Within each orientation channel, the image is filtered with a small-scale filter, the output of which is squared, and then filtered by a large-scale filter that is either parallel to the small-scale filter or perpendicular to it.

---

Figure 17. The stimuli (top row) are the same snake stimuli as in Figure 15, except that the carrier wavelength is 0.55° visual angle, the largest value used in the experiment. The bottom two rows show the results of processing these stimuli with a model the same as that in Figure 15, except that the 1st-stage filters have a longer carrier wavelength—they are half an octave below the carrier wavelength of the elements, in order to provide a good match to the stimulus element profile (see Figure 16). The 2nd-stage filters in this figure are the same as in Figure 15. With the 1st-stage filters adjusted to match the larger element carrier wavelength, the results in this figure are quite similar to those in Figure 15, in which the element carrier wavelength was only 0.19°. The main difference is that, in this figure, the smaller 2nd-order filter just manages to integrate the contour with the second-highest separation whereas, in Figure 15, the 2nd-order filter with this size just fails to integrate it. A small change in the threshold would make both filters give the same pattern of success or failure for each stimulus. Overall, the 2nd-order model is not greatly affected by the carrier wavelength of the stimulus elements.
The maintenance of orientation information at least until the ZBRs have been formed is crucial in allowing segmentation of differently oriented but spatially overlapping ZBRs, as shown in Figure 9. This is the key difference between our model and Dakin’s (1997) adaptive filtering model, which discards information about filter orientation before constructing the ZBRs. In this respect, our model appears superior from a computational standpoint. However, Dakin’s model provided a very good parameter-free fit to psychophysical data on orientation judgment of Glass patterns (Dakin, 1997), and it remains to be seen which kind of model will provide the best overall match to human performance.

Dakin’s (1997) adaptive filtering model is similar to our orientation-linking filter-overlap model in many ways, and our conclusion that the filters must be 2nd-order applies just as much to Dakin’s model as it does to ours. Like our model, Dakin’s model relies on the spatial overlap between filter responses to different elements, so it would predict

![Figure 18](image-url) The result of processing the highest-separation stimulus in Figure 15, with added noise. The element wavelength ($\lambda_{\text{element}}$) was 0.19° visual angle. The model was the 2nd-order snake-detection mechanism described in Figure 15. The wavelength of the 1st-stage filter kernel ($\lambda_1$) was equal to $\lambda_{\text{element}}$. The wavelength of the 2nd-stage filter ($\lambda_2$) differs between the rows, and is given at the side. The best value of $\lambda_2$ is 3.5 octaves above $\lambda_{\text{element}}$. Integration starts to break down after the noise standard deviation exceeds 0.4.

![Figure 19](image-url) Similar to Figure 18, except that ($\lambda_{\text{element}}$) was 0.55° visual angle. Again, $\lambda_1 = \lambda_{\text{element}}$. The best value of $\lambda_2$ is 1.5 octaves above $\lambda_{\text{element}}$. As in Figure 18, contour integration starts to break down after the noise standard deviation exceeds 0.4.

(filter for snake detection) or orthogonal to it (for ladder detection). The output is then thresholded at a fairly high level (about 2 standard deviations above the mean) to give a set of ZBRs in each orientation channel. Any ZBRs in adjacent orientation channels that overlap spatially are linked together to create 3D ZBRs within the 3-dimensional space formed from the two spatial dimensions of the image, along with another dimension representing orientation. To see how these ZBRs delineate the contours within the image, we can collapse them across orientation. However, this final stage is not necessary for performing contour integration tasks, and it may be desirable to maintain the orientation information in the ZBR representation.
that, if 1st-order filters are used, performance on widely separated contours should improve when carrier wavelength is increased, and larger-scale filters can be used.

We showed that, unlike 1st-order filter models, 2nd-order models have the flexibility to cope with a wide range of carrier wavelengths and element separations: The 1st-stage filter can be set to match the wavelength of the stimulus elements, and the 2nd-stage filter can be set to a scale appropriate for the element separation. The 2nd-stage filter in the filter-overlap model therefore plays a similar role to the association field but carries out the integration by blurring along the length of the contour rather than by means of horizontal connections linking disparate receptive fields with positions and orientations consistent with a smooth curve, as has often been assumed in association field models.

For both association field models and filter-overlap models, we suggest that the different effects of element separation on snake and ladder contours might be explained if the upper limit on the size of the integration mechanism is smaller for snakes than ladders. Alternatively, we offer the speculative proposal that the relative sparing of ladder detection at high separations might result from a release from the crowding effect proposed by May and Hess (2007b).

Appendix A

In this Appendix, we examine the effect of an increase in element separation on the signal-to-noise ratio in Yen and Finkel's (1996, 1997, 1998) association field model. The association strength falls as a Gaussian function of distance between the elements. To avoid unnecessary clutter in the equations, we assume that the distance across the visual field is measured in spatial units whose size causes the standard deviation of this Gaussian function to have a value of $1/\sqrt{2}$. Then, the association strength between two elements will be proportional to

$$e^{-x^2},$$

where $x$ is the distance between them. Suppose, for a particular contour element, there is another contour element (the “signal”) at a distance $s$ from the element, and a distractor element (the “noise”) at a distance $n$. Then the signal-to-noise ratio, $\text{SNR}_1$, for that pair of elements (assuming equal association strength in other respects due to equal levels of co-circularity, etc.) will be given by

$$\text{SNR}_1 = e^{-s^2} / e^{-n^2} = e^{n^2-s^2}.$$  

Now, if we increase the separation between all elements by a factor $m > 1$, then the new signal-to-noise ratio, $\text{SNR}_2$, is given by

$$\text{SNR}_2 = e^{m^2(n^2-s^2)}.$$  

The ratio of these SNRs is given by

$$\frac{\text{SNR}_2}{\text{SNR}_1} = e^{(m^2-1)(n^2-s^2)}.$$  

Since $m > 1$, $\text{SNR}_2/\text{SNR}_1 > 1$ if $n > s$, and $\text{SNR}_2/\text{SNR}_1 < 1$ if $n < s$. Thus, increasing the element separation causes the signal-to-noise ratio to increase for distractor elements further away than the neighboring contour element and causes it to decrease for distractor elements closer than the neighboring contour element. The overall effect of increasing the separation on the signal-to-noise ratio will depend on the distribution of distractor positions, and the relative strengths of the inputs from the different elements due to other factors, such as co-circularity.

Acknowledgments

This work was supported by NSERC Grant no. RGPIN 46528-06 to Robert F. Hess. We thank Roger Watt for helpful comments on previous drafts of this paper.

Commercial relationships: none.
Corresponding author: Keith A. May.
Email: keith@keithmay.org.
Address: Vision Sciences Research Group, School of Life Sciences, University of Bradford, Richmond Road, Bradford, West Yorkshire, BD7 1DP, UK.

Footnotes

1 Note that, although the width of the bars within the Gabor elements was varied, the size of the Gaussian envelope was fixed across all conditions.

2 In Jones and Palmer's Gabor receptive field model (as in ours), $\sigma_x$ and $\sigma_y$ are the standard deviations of the envelope along its minor and major axes, respectively. A difference is that, in our model, the bars of the carrier were always parallel to the major axis of the envelope, whereas Jones and Palmer allowed the carrier to be oriented differently: For cell 0811, the carrier was oriented $5^\circ$ from the major axis of the envelope, so our filter kernels differed slightly from scaled versions of the Gabor function that Jones and Palmer fitted to this cell's receptive field.
References


