

Suppressive and facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking

Dennis M. Levi

College of Optometry, University of Houston,
Houston, TX, USA



Srividhya Hariharan

College of Optometry, University of Houston,
Houston, TX, USA



Stanley A. Klein

School of Optometry, University of California,
Berkeley, CA, USA



Peripheral vision is characterized by reduced spatial resolution and inhibitory spatial interactions that extend over long distances. This work had three goals. (1) We considered whether the extensive crowding in peripheral vision is a consequence of a shift in the spatial scale of analysis. To test this, using a large range of target sizes and spatial frequencies, we measured the extent of crowding for targets that were limited in their spatial frequency content. (2) We considered whether crowding in peripheral vision can be explained on the basis of contrast masking by remote flanks. To test this hypothesis, we measured and compared crowding in a direction-identification experiment with masking by remote flanks in a detection experiment. In each of the experiments, our targets and flanks were composed of Gabor features, thus allowing us to control the feature contrast, spatial frequency, and orientation. (3) We examined the relationship between the suppressive and facilitatory interactions in peripheral contrast detection and crowding. Our results show that unlike the normal fovea (Levi, Klein, & Hariharan, 2002), peripheral crowding is not scale invariant nor is it attributable to simple contrast masking. Rather, our results suggest that inhibitory spatial interactions in peripheral crowding extend over larger distances than in the fovea for targets of the same size. In peripheral vision, the critical distance for crowding is approximately 0.1 times the target eccentricity. Observers can easily detect the features that compose our targets (Gabor patches) under conditions where crowding is strong. Thus, we speculate that peripheral crowding occurs because the target and flanks are combined or pooled at a second stage, following the stage of feature extraction. In peripheral vision, this pooling takes place over a large distance.

Keywords: contrast sensitivity, spatial vision, masking, crowding

Introduction

Spatial interactions have been proposed to serve two different purposes in spatial vision. One is to sharpen the perception of individual features, the other to enable features to be grouped into forms. There is an essential tension between these two purposes: strong grouping is not compatible with individuating features, and vice versa. Peripheral vision is characterized by poor spatial resolution. However, peripheral spatial vision is especially compromised when features are presented in clutter. For example, the ability to recognize a small letter (which can be easily recognized in isolation) in peripheral vision is compromised by flanking letters as far away as 0.5 times the eccentricity of the target letter (e.g., Bouma, 1970; Saarinen, Rovamo, & Virsu, 1989; Toet & Levi, 1992; Kooi, Toet, Tripathy, & Levi, 1994; Tripathy & Levi, 1994; Palomares, LaPutt, & Pelli, 1999; Pelli & Palomares, 2000; Chung, Levi, & Legge, 2001). Furthermore, the ability to individuate features is much

coarser than visual resolution, but similar to the extent of crowding in peripheral vision (Intriligator & Cavanagh, 2001). Thus, peripheral vision may be characterized by excessive grouping (Kooi et al., 1994; Orbach & Wilson, 1999).

In their classical study, Flom, Weymouth, and Kahneman (1963) estimated the extent of spatial interactions by having observers judge the orientation of a near acuity threshold Landolt C while varying the distance of surrounding flanks from the C. Flom et al. suggested that the distance over which spatial interaction occurs is related to the size of the receptive fields that are most sensitive to the target. Because peripheral vision is characterized by reduced visual acuity, larger receptive fields will be engaged (because small receptive fields have low sensitivity), and this scale shift will result in proportionally larger crowding distances. This scale shift hypothesis makes several predictions. First, it predicts that in peripheral vision, the spatial extent of crowding (with broadband stimuli) will scale with (be proportional to) the uncrowded acuity. Indeed, for letter acuity (Flom,

Weymouth, & Kahneman, 1963) and Vernier acuity (Levi & Klein, 1985; Levi, Klein, & Aitsebaomo, 1985), the spatial extent of crowding appears to scale with the unflanked letter or Vernier acuity in amblyopia and peripheral vision, respectively. Second, the scale shift hypothesis predicts that the spatial extent of crowding will depend on the size or spatial frequency of the target. We have shown that in the normal fovea, the extent of crowding depends on target size over a 50-fold range of target sizes (Levi, Klein, & Hariharan, 2002). Thus, the large extent of peripheral crowding obtained with near acuity threshold targets might simply be a consequence of the large target sizes used in testing peripheral crowding (due to poor peripheral acuity). According to the scale shift hypothesis, when tested with the same physical size (and spatial frequency) targets, foveal and peripheral crowding should be similar.

One purpose of this work was to test the scale-shift hypothesis for crowding in peripheral vision. Specifically, we were interested in knowing whether the extent of crowding in fovea and periphery is the same when tested with identical stimuli (a test that had not been performed). To test this hypothesis, using a large range of target sizes and spatial frequencies, we measured the extent of crowding for targets that were limited in their spatial frequency content. Our bandlimited (0.825 octaves) stimuli ensured that the initial (linear) filters selected in peripheral vision would be similar in scale to those selected for foveal viewing.

Previously, we developed a “test-pedestal” model for our task, and showed that threshold elevation for crowding is similar to threshold elevation for masking as predicted by the model (Levi et al., 2002). Thus, in foveal vision, crowding may be explained as a form of contrast masking by flanks that overlapped the target. A second purpose of this work was to ask whether crowding in peripheral vision can be explained in the same way. A recent study suggests that in peripheral (but not central) vision, crowding in an orientation-discrimination task is distinct from masking (Parkes, Lund, Angelucci, Solomon, & Morgan, 2001).

In normal foveal vision, spatial interactions may be suppressive (as in crowding) or facilitative (enhancing sensitivity). Thus, in the fovea, high-contrast flankers can reduce the contrast needed to detect a small Gabor patch (Polat & Sagi, 1993, 1994; Levi et al., 2002). Polat and Sagi (1993, 1994) attributed this facilitation to long-range excitatory connections that are known to exist in cortical area V1 (e.g., Gilbert, 1998; Fitzpatrick, 2000). Interestingly, surround facilitation has been reported to be absent in the normal periphery (Williams & Hess, 1998; Xing & Heeger, 2000, 2001). This raises the question of whether the extensive crowding in peripheral vision reflects an imbalance in the strength of suppressive and facilitative interactions in peripheral vision? Thus, the third purpose of this study was to examine the relationship between the suppressive and facilitatory

interactions in peripheral contrast detection and crowding.

Methods

Our targets, flanks, and all experimental details were identical to those described earlier (Levi et al., 2002), and examples of the stimuli can be seen in their Figures 1, 2, 9, and 14. Three normal observers (two of the authors and M.F., a practiced observer who was naïve to the purpose of the experiments) were tested in peripheral vision (5 and 10 degrees, lower visual field). Viewing was monocular with the untested eye occluded with a black patch.

Experiment 1

We measured crowding in peripheral vision using the four-alternative forced-choice (4AFC) method of constant stimuli, as described previously (Levi et al., 2002; Experiment 1).

Results

Crowding in peripheral vision extends over larger distances

In foveal vision, the extent of crowding depends on target size over a wide range of target sizes (an \approx 50-fold range of target sizes as seen earlier [Levi et al., 2002]). In peripheral vision, for a given target size (Figure 1A, solid symbols), crowding extends over considerably larger distances than in the fovea (Figure 1A, open symbols), even with our bandlimited stimuli.

Several points of interest are in Figure 1 (which shows some examples of our data). In the normal fovea, the distance over which flanks influence performance depends on the target size (compare the open symbols in Figure 1B) as noted in Levi et al. (2002). In peripheral vision, thresholds are elevated (more for small stimuli than for large—compare small and large solid symbols in Figure 1B), and the extent of crowding is increased. Note also that changing the carrier spatial frequency while fixing the target size (compare diamonds and circles in Figure 1A) has no effect on performance in fovea or periphery.

We quantified the extent of crowding by fitting Gaussian functions to the data (lines in Figure 1) and specifying the critical distance for crowding as the flank distance that causes the unflanked threshold to double (Equation 1, Levi et al., 2002). The critical distance (specified in arc min) is plotted as a function of target size for Gaussian Es (Figure 2) and for Gabor Es (Figure 3, inside abscissa).

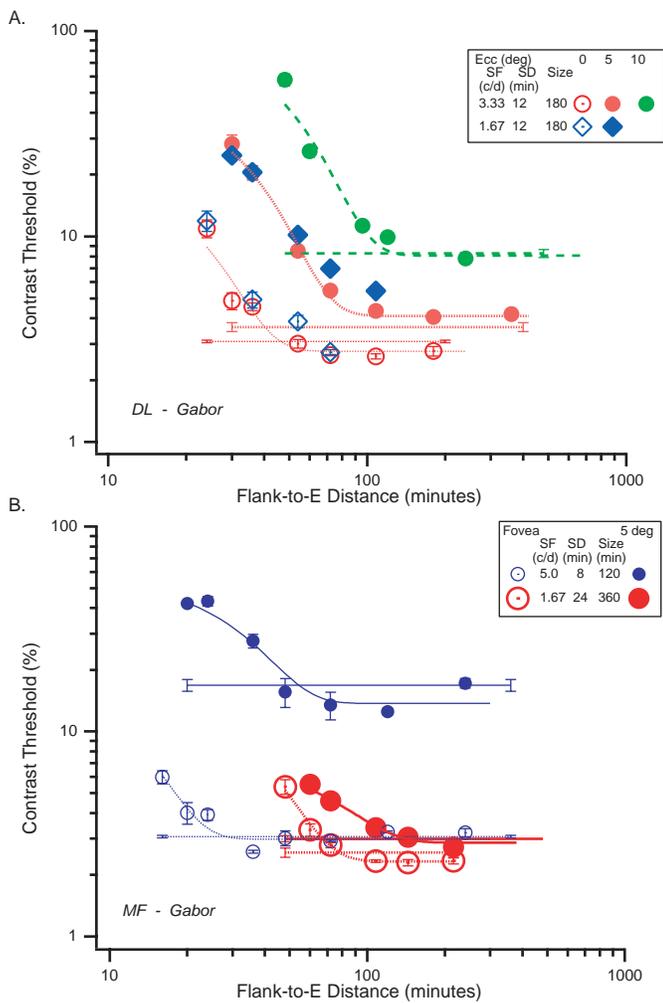


Figure 1. Contrast thresholds versus flank distance for Gabor Es. A. Examples of crowding functions for one target size with two different spatial frequencies at 0, 5, and 10 degrees for observer D.L. B. Examples of crowding functions for two target sizes in the fovea (open symbols) and 5 degrees (solid symbols) for observer M.F. The lines are Gaussian functions (see Levi et al., 2002) fit to the data.

In normal foveal vision, the critical distance (in minutes) is more or less proportional to target size, falling close to the line of unity slope for Gaussian Es (Figure 2) and for Gabor Es (Figure 3). For a fixed-size Gaussian E target (Figure 2), the critical distance is bigger in peripheral vision, and similar results are obtained when the flanks consist of 5 patches each, or just 2 appropriately placed patches. The results with Gaussian targets are similar to those of previous work: The stimuli are broadband; therefore, the increased extent of crowding may be a consequence of the peripheral visual system engaging large (low) spatial frequency filters. Note, however, that most previous studies used differently sized targets to test the fovea and periphery. Our results show that the extended crowding in peripheral vision is not simply a consequence of using larger peripheral targets.

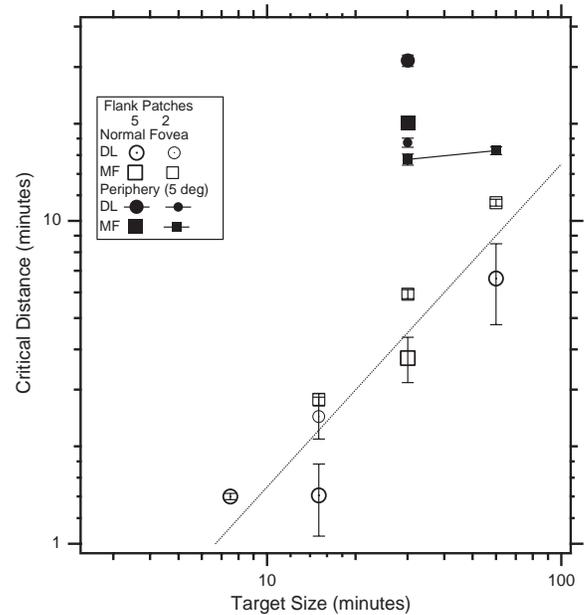


Figure 2. The critical distance (flank distance at which thresholds are elevated by a factor of 2, specified in arc min) versus total target size for Gaussian Es. The critical distance is bigger at 5 degrees (solid symbols) than at the fovea (open symbols), and similar results are obtained when the flanks consist of 5 patches each (larger symbols), or just 2 patches placed at the possible gap locations (see Levi et al., 2002, Figure 2, bottom right panel) (Data are shown here by the smaller symbols).

Interestingly, similar results are obtained with Gabor patches (Figure 3). For very large target sizes (greater than about 150 minutes), the critical distance approaches (but is larger than) that of the normal fovea; however, for smaller targets, there appears to be a floor, so that the critical distance becomes a nearly constant (large) distance. The smallest critical distance depends on eccentricity (it is larger at 10 degrees than at 5 degrees). At both eccentricities, the smallest critical distance with Gabor patches is approximately 10% of the effective eccentricity (i.e., eccentricity, $E + E_2$, where E is the eccentricity and E_2 is the doubling eccentricity, is approximately 0.7 degrees [Levi et al., 1985Levi, Klein, & Aitsebaomo, 1985Levi, Klein, & Aitsebaomo, 1985Levi, Klein, & Aitsebaomo, 1985]; at an eccentricity of 5 degrees, it is about 0.5 degrees, and at an eccentricity of 10 degrees, it is about 1 degree. Note that the critical distance is roughly similar in size for the 4AFC task and the 2AFC task (which is discussed below). The gray symbols show the critical distances obtained with Gaussian Es (with 5 patches/flank, from Figure 2). For D.L., the Gaussian data are consistent with the 10% floor; for M.F., they are somewhat lower. Thus, in peripheral vision, the extent of crowding is not proportional to target size; however, the extent of interaction is stimulus dependent. We note that the critical distance is very much larger (more than 20

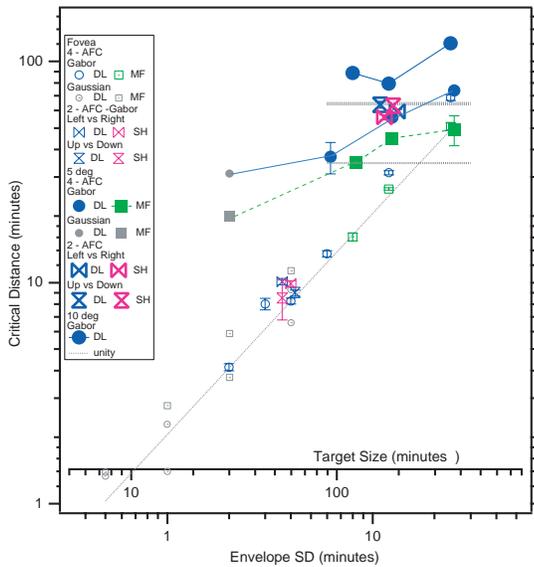


Figure 3. The critical distance versus patch standard deviation (bottom abscissa) or total target size (upper abscissa). The target size is 15 times larger than SD because target size is 5 times feature separation, which is 3 times larger than SD. In the periphery, there seems to be a floor, so that the critical distance cannot go below a nearly constant (large) distance. The horizontal dotted lines are shown at 10% of the effective eccentricity. The circles show data for the 4AFC task; bow ties and hourglasses for the 2AFC left/right and up/down, respectively. Gray symbols are Gaussian (5 patches/flank) from Figure 2.

times) than the resolution limit, but that it is similar to the critical distance for the crowding effect of flanks on a single Vernier target (Levi et al., 1985; Levi, Klein, & Aitsebaomo, 1985). In foveal vision, crowding is scale invariant and is primarily determined by target size (SD). When replotted as threshold elevation (i.e., flanked threshold/unflanked threshold) versus target-to-flank distance expressed in standard deviation units (SDU, i.e., target-to-flank distance [in arc min], divided by patch SD [in arc min]), foveal performance over a wide range of pattern sizes collapses into a more or less unitary function (see Levi et al., 2002, Figure 7). In peripheral vision (Figure 4, solid symbols), it is clear that crowding is not scale invariant. When plotted as threshold elevation versus target-to-flank distance (in SDU), it is clear that for small targets, the crowding does not scale to target size, but is disproportionately large—instead of the extent of crowding being ≈ 2.5 SDU as in the fovea, it may be much larger in the periphery (e.g., D.L. 5 degrees with Gaussian Es [not shown] is about 15 SDU). The extended crowding in peripheral vision is not a consequence of our choice of spatial frequency. The diamonds in Figure 4 show that in the periphery, as in the fovea, crowding is

similar for patches of the same size (SD) with spatial frequencies that are one octave apart. However, in the periphery, it is the eccentricity, rather than the target size, that determines the extent of crowding.

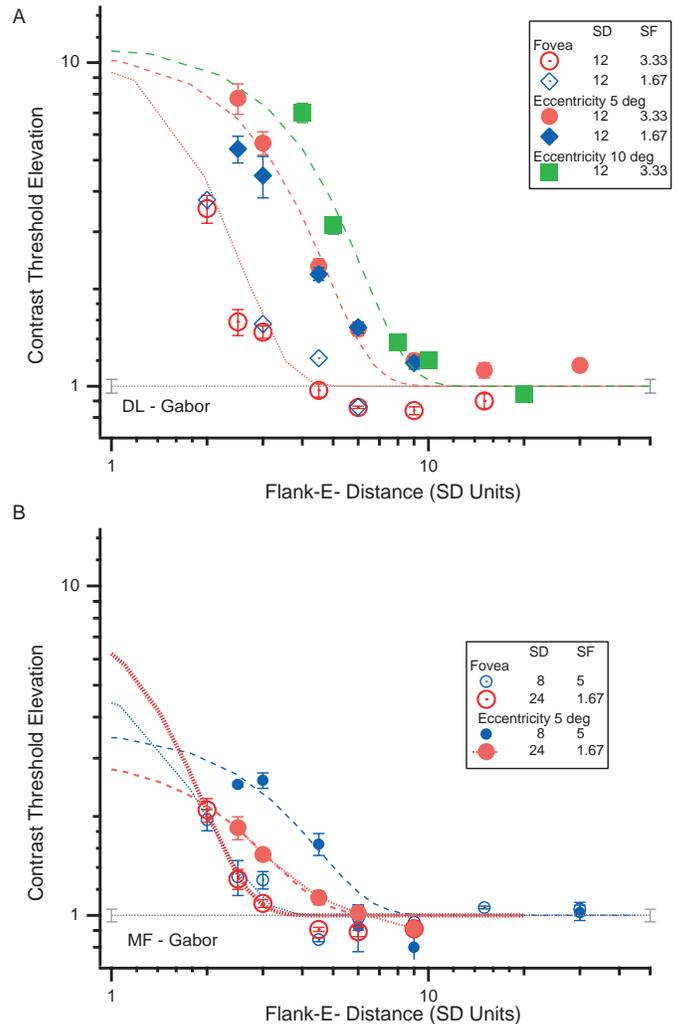


Figure 4. Foveal crowding is scale invariant; peripheral crowding is not. The data of Figure 1 are replotted as threshold elevation (i.e., flanked threshold/unflanked threshold) versus target-to-flank distance expressed in standard deviation units (SDU, i.e., target-to-flank distance [in arc min], divided by patch SD [in arc min]). When plotted in this way, foveal performance (open symbols) over a wide range of pattern sizes collapses into a more or less unitary function (see Levi et al., 2002, Figure 7); however, peripheral performance does not. In fovea and periphery, target spatial frequency makes little difference.

Crowding causes 180-degree errors

Under conditions where crowding occurs, normal observers viewing foveally make a preponderance of 180-degree (mirror reversal) errors (Levi et al., 2002). Figure 5 shows the proportion of 90- and 180-degree errors under conditions where crowding occurs (small flank distance,

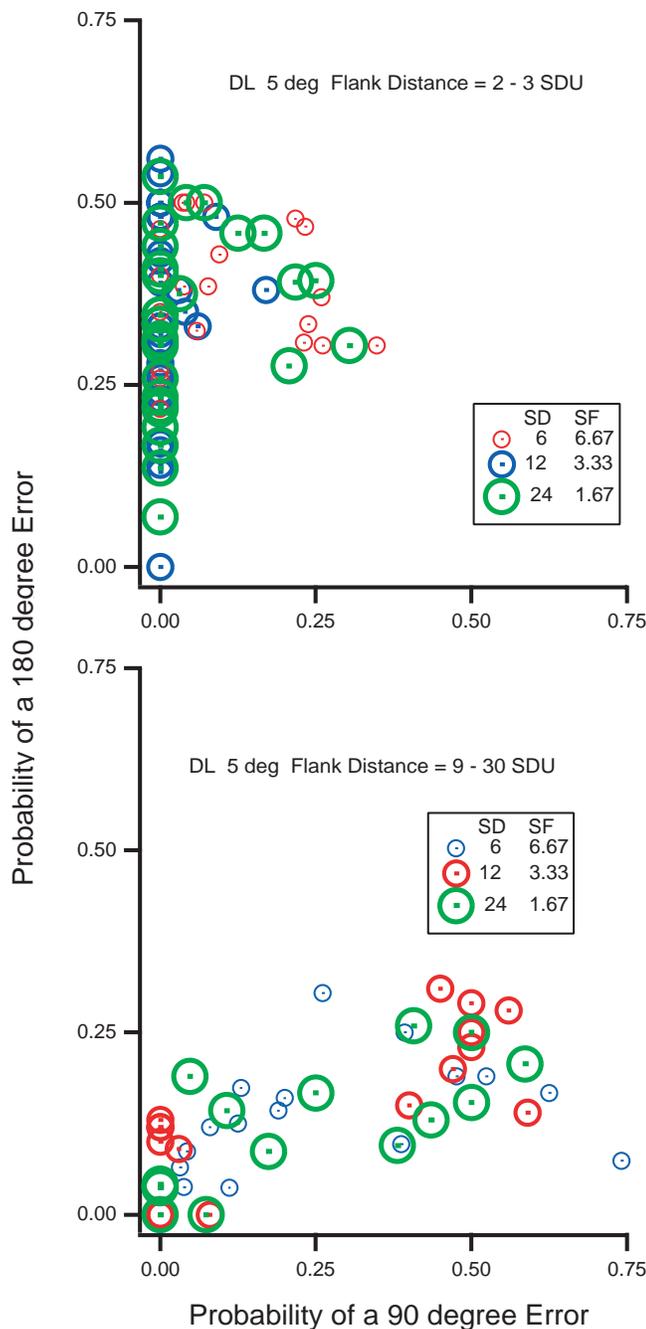


Figure 5. Confusion analysis. We classified the errors as either 180-degree errors (mirror image errors) or 90-degree errors (nonmirror image errors). This figure plots the proportion of 180-degree versus 90-degree errors under conditions where crowding occurs (small flank distances [2-3 SDU], top panel) and under conditions where there is little or no crowding (large flank distances [9-30 SDU], lower panel) for observer D.L. Data for different target sizes are coded by symbol size and color. Note that under conditions of crowding, there is a preponderance of 180-degree errors. Under conditions of crowding, the observer is able to correctly judge whether the legs of the E are oriented vertically or horizontally, but is unable to correctly identify the location of the gaps.

top panels) and under conditions where there is little or no crowding (large flank distances, lower panels) for peripheral viewing. In peripheral vision, as in the fovea, under conditions of crowding, there is a preponderance of 180-degree errors, suggesting that crowding is not simply a loss of visibility, or of orientation information (Hess, Dakin, & Kapoor, 2000), but is rather a specific loss of positional information. Under conditions of crowding, the observer is able to correctly judge whether the legs of the E are oriented vertically or horizontally, but is unable to correctly identify the location of the gaps

In the following experiment, we show that similar crowding occurs in the absence of an orientation cue, and that peripheral crowding is not simply contrast masking by remote flanks.

Experiment 2: Crowding Versus Masking in Peripheral Vision

In this experiment, we compared crowding in a 2AFC E direction-identification experiment (see Levi et al., 2002, Experiment 2) with masking by remote flanks in a single patch (solo) detection experiment (see Levi et al., 2002, Experiment 3).

Results

Crowding in the 2AFC experiment (with no orientation cue) is quite similar in strength and extent to that obtained in the 4AFC experiment (with the orientation cue, Figure 6; compare the critical distances in Figure 3). Interestingly, threshold elevation for solo detection (bow ties and hourglasses in Figure 6) is much weaker and less extensive than threshold elevation for the E direction-identification task. For example, for both observers, threshold elevation in the noncollinear configuration at the smallest flank distance is almost 8 times less than threshold elevation for the E at the comparable flank distance. In the E task, each flank consisted of 5 patches, whereas in the solo task, the flanks were single patches. However, increasing the number of flank patches (to 5) in the solo task did not materially influence the threshold elevation (small bow ties in Figure 6).

In foveal vision, we found strong facilitation by remote flanks in the solo detection task (Polat & Sagi, 1993, 1994). Interestingly, observer D.L. showed no facilitation in peripheral viewing, and S.H. showed some facilitation, but only with collinear flanks. Facilitation may depend critically on flank contrast; therefore, we measured threshold as a function of flank contrast at a fixed flank separation (6 SDU or 4.5 λ units), where S.H. showed the strongest facilitation in the periphery. Figure 7 shows that over the range that we measured, there is no evidence of facilitation for this observer.

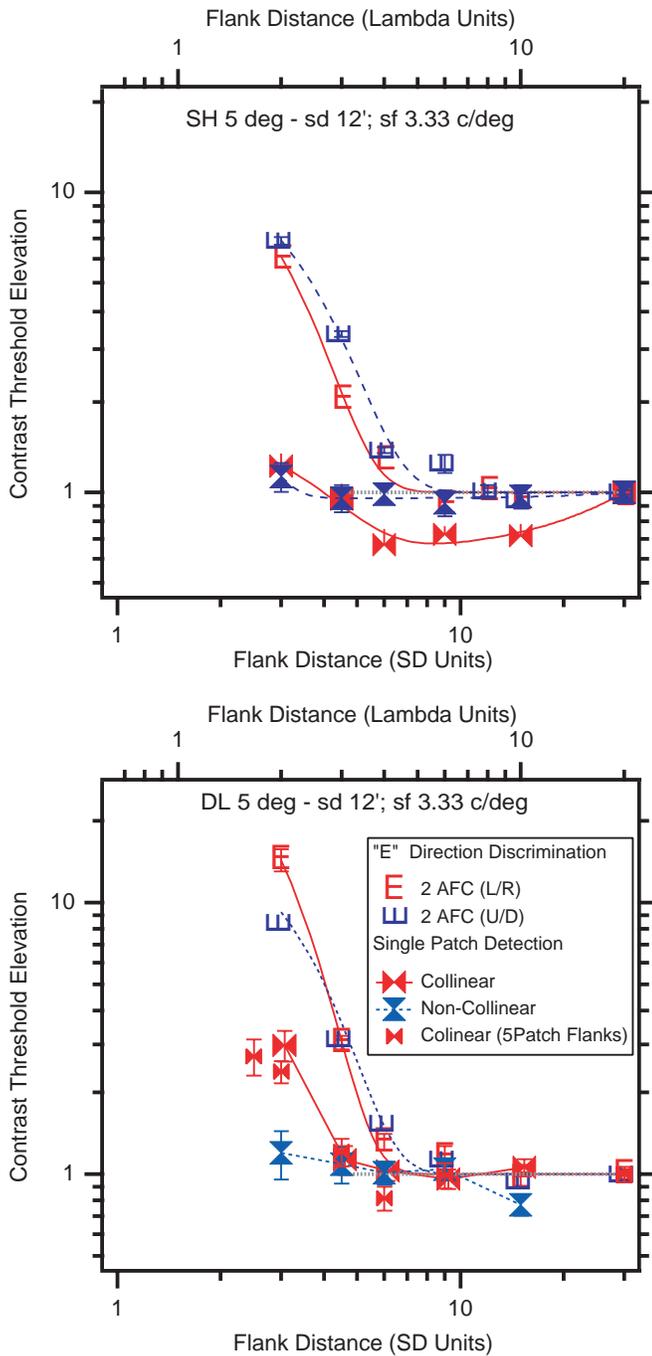


Figure 6. Crowding occurs for 180-degree direction discrimination for Es at 5 degrees in the lower field. For comparison, the bow ties and hourglasses show threshold elevation for detecting a single patch with a pair of flanks in a collinear and noncollinear arrangement, respectively. Data are for two observers (S.H., top; D.L., bottom). The small bow ties show single patch detection with each flank composed of 5 patches.

We have previously shown that in the fovea, the slope of the psychometric function was significantly flatter in the presence of flanks that produced facilitation than with no flanks (Levi et al., 2002). Interestingly, in peripheral vision, even in the absence of facilitation, we

found that flanks placed at 4.5 to 6 SDU from the target resulted in a significant flattening of the slope (relative to that with no flanks). Averaged across observers and conditions (i.e., collinear and noncollinear), the slope with flanks (at 4.5 and 6 SDU) was 1.2 ± 0.1 and without flanks was 1.8 ± 0.2 .

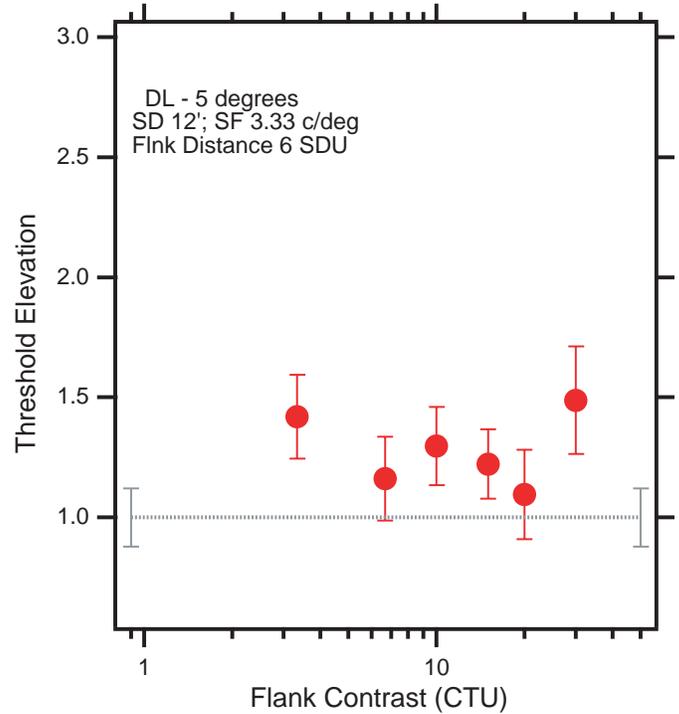


Figure 7. Threshold elevation as a function of flank contrast (in contrast threshold units) for observer D.L. at 5 degrees in the lower visual field.

To directly compare the effects of flanks in masking and crowding, we plot threshold elevation for crowding (E direction discrimination) against threshold elevation for masking (solo detection) for paired conditions (e.g., L/R E vs. collinear solo at the same flank distance, or U/D E vs. noncollinear at the same flank distance [Figure 8]). Each symbol in Figure 8 represents a paired measure. Data inside the red box show facilitation for solo detection and data inside the green box for E direction discrimination. The open symbols are foveal data from Levi et al., 2002, Figure 17. The solid symbols show that in the periphery, there is not much facilitation, but, more importantly, crowding is much stronger than contrast masking (i.e., threshold elevation for identification of the E direction is much stronger than threshold elevation for detecting a single patch). Threshold elevation for the single patch never exceeds a factor of 3, whereas threshold elevation for the E may be more than a factor of 10! It is also interesting to note a cluster of data where there is little or no masking (threshold elevation for the single patch is close to 1) but substantial crowding (threshold elevation for the E between two- and nine-fold).

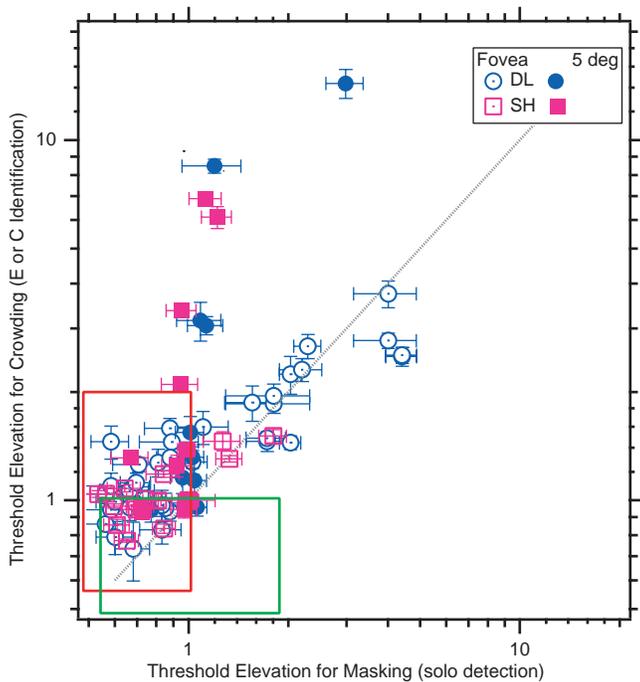


Figure 8. Threshold elevation for crowding (E or C direction discrimination) versus threshold elevation for masking (solo detection) for paired conditions (e.g., L/R E vs. collinear solo at the same flank distance; or U/D E vs. noncollinear at the same flank distance). Each symbol represents a paired measure in the fovea (open symbols) or at 5 degrees in the lower visual field (solid symbols). Results are shown for two observers. Data within the red box show facilitation for solo detection (abscissa values below 1). Data within the green box show facilitation for E direction discrimination (ordinate values below 1).

As in the fovea, crowding and masking are orientation specific in peripheral vision. Figure 9 shows significant threshold elevation for both tasks at a flank distance of 3 SDU when the target and flanks are iso-oriented (both have a horizontal carrier), and negligible threshold elevation when the flank orientation is ortho (i.e., vertical). However, unlike the fovea, threshold elevation for similar flank/target orientations is considerably stronger (about a factor of 5) for crowding (red bars) than for masking (blue bars).

Discussion

Crowding in peripheral vision is not scale invariant nor is it attributable to simple contrast masking. Rather, our results suggest that inhibitory spatial interactions in peripheral crowding are larger than in the fovea for a comparable target size, and that there may be a floor at $\approx 0.1E$. This large extent of peripheral crowding is comparable with that of previous studies (for a summary, see Chung et al., 2001, Table 1), and we suspect that much of the variation in the reported extent is due to

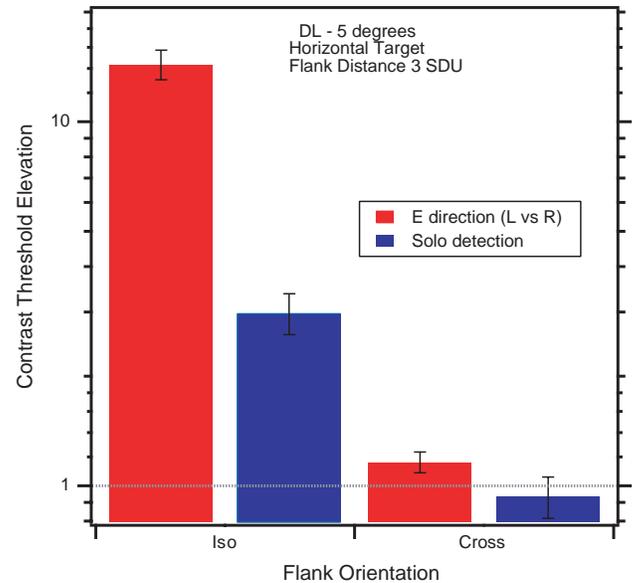


Figure 9. The effect of flank orientation in peripheral vision. Threshold elevation for crowding (red) and masking (blue) when the flank carrier orientation is Iso (i.e., horizontal like the target) or ortho (i.e., vertical or orthogonal to the target). The iso data are from Figure 6. Note that crowding and masking are orientation specific, and that crowding is much stronger than masking in the periphery.

the way in which the extent is specified. As we noted previously, for letter targets and letter flankers, the extent is often specified from target center to letter center (Levi et al., 2002). Our method of specifying the extent based on the separation between a flanking bar and an adjacent limb of the target (i.e., the gap) results in the masking extent being reduced five-fold (i.e., $\approx 0.1E$ rather than $\approx 0.5E$). The advantage of using the gap to specify the flank distance is that it facilitates comparison with masking of a single feature. We have previously shown that in foveal vision, masking of one of the features composing the target predicts threshold elevation for crowding (Levi et al., 2002). The novel result of this work is the demonstration that masking of a single feature does not predict peripheral crowding. What is the basis of the large extent of peripheral crowding? Our results rule out simple contrast masking as the sole factor. Figure 10 compares crowding in fovea (open symbols) and periphery (solid symbols) after discounting the effects of contrast masking. The data represent the linear differences between the ordinate and abscissa values of points in Figure 8, plotted here as a function of flank distance. Clearly, after subtracting out the effects of masking, there is essentially no residual threshold elevation in the fovea (the small elevation, ≈ 0.5 at around 4.5 SDU, is actually the effect of facilitation of detection for a single patch). In contrast, there is substantial and extensive threshold elevation in the periphery. We suggest that this represents genuine crowding, unconfounded by contrast masking. Note that

we linearly subtracted the effect of masking from that of crowding, rather than taking the ratio (crowding/masking); however, the conclusion is the same for the ratio. For the periphery, the ratio at the peak threshold elevation is between 5 and 7 (mean 5.75 ± 0.5), whereas in the fovea, it is 0.83 ± 0.09 . Below we consider several alternative explanations for the strong and extended inhibitory spatial interactions in peripheral vision.

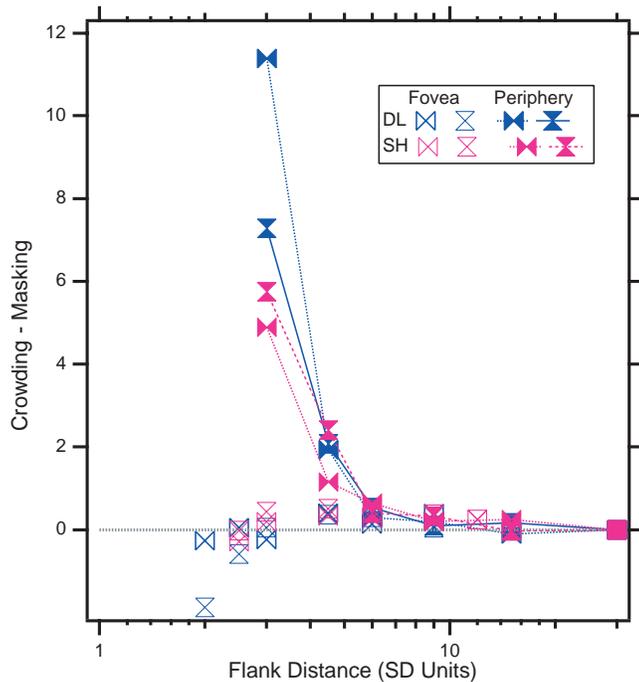


Figure 10. Crowding with masking discounted in fovea (open symbols) and periphery (solid symbols) for D.L. (blue) and S.H. (pink). The data represent threshold elevation for crowding (from the 2AFC E direction-identification experiment) after subtracting out threshold elevation for masking (from the corresponding single-patch detection experiment). Thus for the periphery, the data plotted here represent the difference in E threshold elevation and single patch threshold elevation from Figure 6. The foveal data are from Figures 11 and 12 of our earlier work (Levi et al., 2002).

Large Peripheral Receptive Fields

Although the optical quality of the eye changes little with eccentricity within the central 10 degrees or so (e.g., Jennings & Charman, 1978; Losada, Navarro, & Santamaria, 1993), there are marked neural changes in the retina and the visual cortex. In particular, the size and spacing of retinal and cortical receptive fields increase with eccentricity. We argue that crowding and masking must occur after information from the two eyes converges, because both are orientation specific (Figure 8) and because crowding occurs when the target is presented to one eye and flanks to the other (Flom, Heath, &

Takahashi, 1963; Kooi et al., 1994). Thus, it is reasonable to ask whether the large extent of peripheral crowding reflects pooling of target and flanks by the large peripheral receptive fields of cortical area V1. Specifically, does the extended peripheral crowding reflect a spatial scale shift to larger receptive fields (lower spatial frequencies) in the periphery? This is effectively the scale-shift hypothesis of Flom, Weymouth, and Kahneman, (1963). Most previous studies of peripheral crowding use broadband stimuli (e.g., letters) that are close to the acuity limit. Thus, crowding in peripheral vision is typically measured with larger (broadband) stimuli than in the fovea, and the resulting data may indeed reflect a shift in spatial scale. Recall that our Gabor E stimuli are composed of narrow-band features, identical to the features used in the masking experiments, and we show that peripheral crowding extends over a greater distance even when tested with the same size (and spatial frequency) stimuli as the foveal stimuli. Thus we argue that it is unlikely that the extended peripheral crowding reflects a simple shift in spatial scale of first stage filters toward larger (low spatial frequency) filters at an early stage of processing in V1.

Spatial Uncertainty

Peripheral vision is characterized by a high degree of spatial uncertainty (Pelli, 1985). Previously, we suggested that the facilitation, clearly evident in foveal solo detection with remote flanks, might be a consequence of the high-contrast flanks acting to reduce spatial uncertainty (and possibly uncertainty about the spatial frequency, orientation, etc.) of the near threshold target (Levi et al., 2002). In support of this argument, we showed that the slope of the psychometric function was significantly flatter in the presence of flanks that produced facilitation, than with no flanks. This result is a prediction based on the accelerated d' versus contrast function of the uncertainty model (Pelli, 1985). The present results, in agreement with previous work, suggest that facilitation is either absent or weak in peripheral vision (Williams & Hess, 1998; Xing & Heeger, 2000, 2001). In peripheral vision, peripheral uncertainty is exaggerated, so the finding that facilitation is reduced is counterintuitive. Based on the uncertainty model, one might predict even stronger facilitation in the periphery, unless peripheral uncertainty was so great that the location (and other details) of the high-contrast flanks is uncertain. This explanation seems paradoxical because we found that the slope of the peripheral psychometric function was also flatter in the presence of flanks than in their absence (mean exponents were ≈ 1.2 with flanks and 1.8 without). However, even if the flanks reduce uncertainty, in peripheral vision, the window of attention may be too large (so it includes the flanks) or mis-aimed (i.e., not directed to the appropriate location) to allow facilitation. In addition, although the uncertainty model

is widely recognized, other models (see [Hubner, 1993](#)) predict that the slope of the psychometric function may decrease (rather than increase) as a result of uncertainty. Although several studies (noted above) have failed to find facilitation in peripheral vision, we note that one of our observers (S.H.) did show some facilitation. Moreover, recent experiments suggest that under ideal conditions (lower spatial frequencies, lower flank contrast, and more distant and longer flanks), it is possible to obtain substantial facilitation in peripheral vision ([Yu, Klein, & Levi, 2001](#)). Thus, we conclude that peripheral facilitation is quite delicate, and a failure to find facilitation need not be due to a genuine loss of neural facilitation. Moreover, it seems unlikely that increased spatial uncertainty could explain the greater strength and extent of peripheral crowding.

Long-Range Inhibitory Connections

[Polat and Sagi \(1993, 1994\)](#) have argued that the facilitation by remote flankers observed in foveal detection is a consequence of excitatory long-range horizontal connections between neurons with like orientations in cortical area V1. These horizontal connections may be excitatory and inhibitory (e.g., [Fitzpatrick, 2000](#); [Gilbert, 1998](#)), and extend up to about 1 to 2 mm in primate area V1 ([Rockland & Lund, 1983](#); [Blasdel, Lund, & Fitzpatrick, 1985](#); [Fitzpatrick, Lund, & Blasdel, 1985](#); [Lund, Yoshioka, & Levitt, 1993](#); [Amir, Harel, & Malach, 1993](#)). Based on recent estimates of human cortical magnification, these cortical distances translate to approximately 0.1E in peripheral vision ([Beard, Levi, & Klein, 1997](#)). Thus, long-range inhibitory connections have approximately the requisite length to account for the floor distance ($\approx 0.1E$) of peripheral crowding. On the other hand, this close correspondence may be coincidental, because long-range connections appear to be too short to account for foveal crowding ([Levi et al., 2002](#)). Moreover, the fixed cortical distance of long-range connections predicts interactions over a fixed retinal distance, rather than interactions that are related to target size in the fovea. In addition, it is unclear why long-range interactions in masking should be different from long-range interactions in crowding.

Second-Stage Pooling

When several features are presented together, perception of the spatial details of an individual feature depends on (a) the ability of the visual system to resolve each feature (visual resolution), and (b) the ability of mechanisms at a subsequent stage to isolate each feature. [Intriligator and Cavanagh \(2001\)](#) refer to this as "attentional resolution," and they showed that in peripheral vision, the limits imposed by visual resolution and attentional resolution are quite different. [He, Cavanagh, and Intriligator \(1996\)](#) have argued that

peripheral crowding results from limitations set by attentional resolution. We prefer the more neutral notion that crowding reflects limited resolution at a stage beyond the initial filtering stage (see also [Chung et al., 2001](#)). Our crowding task requires that the observer not only detect the features but also isolate and localize the missing features (the gaps). Based on our masking experiments, observers can easily detect the features under conditions where crowding is strong. Thus, our speculation is that peripheral crowding occurs because the target and flanks are combined or pooled at a second stage that follows the stage of feature extraction. In peripheral vision, this pooling takes place over a long distance. Another way of saying this is that the second stage filter or template for identifying a target in peripheral vision is not well matched to the target ([Levi, McGraw, & Klein, 2000](#)). In several respects, this notion is similar to the proposition that patterns that generate considerable neural activity within a constrained region of the periphery are processed as textures, so that the details of individual patterns are not available for discrimination (e.g., [Wilkinson, Wilson, & Ellemberg, 1997](#); [Orbach & Wilson, 1999](#)). In these models, the texture processing reflects a second-stage analysis by complex cells. Our results are also consistent with the recent study of [Parkes et al. \(2001\)](#). They found that in peripheral (but not central) vision, crowding in an orientation-discrimination task was distinct from masking. Importantly, they were able to show that in a cluttered display, the orientation signals were pooled rather than being lost through masking. Their conclusion is that crowding reflects compulsory averaging of signals, the term we use to define texture perception "when we do not wish it to occur."

Is crowding in the periphery qualitatively different from that in the fovea? Crowding in foveal and peripheral vision appears to differ in two important ways. First, over the range of target sizes we measured, foveal crowding appears to be size invariant, whereas in peripheral vision, crowding occurs over a large spatial extent. Second, foveal crowding is reasonably well predicted from masking of detection of a feature, whereas peripheral crowding is much stronger than would be predicted by simple contrast masking. Our speculation is that target identification across the visual field involves the same two stages of processing: initial filtering and a second-stage template. The initial filtering stage (which limits detection) appears to be similar in foveal and peripheral vision. In foveal vision, however, the second stage filter (template) is exquisitely matched to the target (see also [Levi, Klein, & Carney, 2000](#)); whereas, in peripheral vision, because of limited resources, it is not ([Levi et al., 2000](#)). This view is consistent with the finding that crowding in foveal and peripheral vision is qualitatively different ([Hess, Dakin, Kapoor, & Tewfik, 2000](#)). However, our study also shows that this difference is not a consequence of a simple shift in the spatial scale of analysis.

Conclusions

1. Peripheral crowding is not scale invariant nor is it attributable to simple contrast masking.
2. Inhibitory spatial interactions in peripheral crowding extend over larger distances than in the fovea for targets of the same size, and there may be a floor at $\approx 0.1E$.
3. Observers can easily detect the features that compose our targets (Gabor patches) under conditions where crowding is strong.
4. We speculate that peripheral crowding occurs because the target and flanks are combined or pooled at a second stage that follows the stage of feature extraction. In peripheral vision, this pooling takes place over a large distance.

Acknowledgments

We are grateful to Hope Queener for programming these experiments. This research was supported by National Eye Institute Research Grants R01EY01728 and RO1 EY04776, Core Center Grant P30EY07551, and short-term training grant T35EY07088. Commercial Relationships: None.

References

- Amir, Y., Harel, M., & Malach, R. (1993). Cortical hierarchy reflected in the organization of intrinsic connections in macaque monkey visual cortex. *Journal of Comparative Neurology*, *334*, 19–46. [PubMed]
- Beard, B. L., Levi, D. M., & Klein, S. A. (1997). Vernier acuity with non-simultaneous targets: The cortical magnification factor estimated by psychophysics. *Vision Research*, *37*, 325–346. [PubMed]
- Blasdel, G. G., Lund, J. S., & Fitzpatrick, D. (1985). Intrinsic connections of macaque striate cortex: axonal projections of cells outside lamina 4C. *Journal of Neuroscience*, *5*, 3350–3369. [PubMed]
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*, 177–178. [PubMed]
- Chung, S. T. L., Levi, D. M., & Legge, G. E. (2001). Crowding: A classical spatial-frequency masking effect? *Vision Research*, *41*, 1833–1850. [PubMed]
- Fitzpatrick, D., Lund, J. S., & Blasdel, G. G. (1985). Intrinsic connections of macaque striate cortex: Afferent and efferent connections of lamina 4C. *Journal of Neuroscience*, *5*, 3329–3349. [PubMed]
- Fitzpatrick, D. (2000). Seeing beyond the receptive field in primary visual cortex. *Current Opinion in Neurobiology*, *10*, 438–443. [PubMed]
- Flom, M. C., Weymouth, F. W., & Kahneman, D. (1963). Visual resolution and spatial interaction. *Journal of the Optical Society of America A*, *53*, 1026–1032.
- Flom, M. C., Heath, G. G., & Takahashi, E. (1963). Contour interaction and visual resolution: Contralateral effects. *Science*, *142*, 979–980.
- Gilbert, C. D. (1998). Adult cortical dynamics. *Physiological Reviews*, *78*, 467–485. [PubMed]
- He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, *383*, 334–337. [PubMed]
- Hess, R. F., Dakin, S. C., & Kapoor, N. (2000). The foveal ‘crowding’ effect: Physics or physiology? *Vision Research*, *40*, 365–370. [PubMed]
- Hess, R. F., Dakin, S. C., Kapoor, N., & Tewfik, M. (2000). Contour interaction in fovea and periphery. *Journal of the Optical Society of America A*, *17*, 1516–1524. [PubMed]
- Hubner, R. (1993). Different ways of modeling spatial-frequency uncertainty in visual signal detection. *Biological Cybernetics*, *9*, 457–462.
- Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*. In Press.
- Jennings, J., & Charman, W. N. (1978). Optical image quality in the peripheral retina. *American Journal of Optometry and Physiological Optics*, *55*, 582–590. [PubMed]
- Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision*, *8*, 255–279. [PubMed]
- Levi, D. M., & Klein, S. A. (1985). Vernier acuity, crowding and amblyopia. *Vision Research*, *25*, 979–991. [PubMed]
- Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research*, *25*, 963–977. [PubMed]

- Levi, D. M., Klein, S. A., & Carney, T. (2000). Unmasking the mechanisms for Vernier acuity: Evidence for a template model of Vernier acuity. *Vision Research*, 40, 951-972. [PubMed]
- Levi, D. M., Klein, S. A., & Hariharan, S. (2002). Suppressive and facilitatory spatial interactions in foveal vision: Foveal crowding is simple contrast masking. *Journal of Vision*, 2(2), 140-166, <http://journalofvision.org/2/2/2/>, DOI 10.1167/2.2.2 [Link].
- Levi, D. M., McGraw, P. V., & Klein, S. A. (2000). Vernier and contrast discrimination in central and peripheral vision. *Vision Research*, 40, 973-988. [PubMed]
- Losada, M. A., Navarro, R., & Santamaria, J. (1993). Relative contributions of optical and neural limitations to human contrast sensitivity at different luminance levels. *Vision Research*, 33, 2321-2336. [PubMed]
- Lund, J. S., Yoshioka, T., & Levitt, J. B. (1993). Comparison of intrinsic connectivity in different areas of macaque monkey cerebral cortex. *Cerebral Cortex*, 3, 148-162. [PubMed]
- Orbach, H. S., & Wilson, H. R. (1999). Factors limiting peripheral pattern discrimination. *Spatial Vision*, 12, 83-106. [PubMed]
- Palomares, M., LaPutt, M. C., & Pelli, D. (1999). Crowding is unlike ordinary masking [Abstract]. *Investigative Ophthalmology and Visual Science*(Suppl.), 40, S351.
- Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, 4, 739-744. [PubMed]
- Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America A*, 2, 1508-1532. [PubMed]
- Pelli, D. G., & Palomares, M. (2000). The role of feature detection in crowding [Abstract]. *Investigative Ophthalmology and Visual Science*(Suppl.), 41, S37.
- Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. *Vision Research*, 33, 993-999. [PubMed]
- Polat, U., & Sagi, D. (1994). The architecture of perceptual spatial interactions. *Vision Research*, 34, 73-78. [PubMed]
- Rockland, K. S., & Lund, J. S. (1983). Intrinsic laminar lattice connections in primate visual cortex. *Journal of Comparative Neurology*, 216, 303-318. [PubMed]
- Saarinen, J., Rovamo, J., & Virsu, V. (1989). Analysis of spatial structure in eccentric vision [Abstract]. *Investigative Ophthalmology and Visual Science*, 30, 293-296.
- Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, 32, 1349-1357. [PubMed]
- Tripathy, S. P., & Levi, D. M. (1994). Long-range dichoptic interactions in the human visual cortex in the region corresponding to the blind spot. *Vision Research*, 34, 1127-1138. [PubMed]
- Williams, C. B., & Hess, R. F. (1998). Relationship between facilitation at threshold and suprathreshold contour integration. *Journal of the Optical Society of America A*, 15, 2046-2051.
- Wilkinson, F., Wilson, H. R., & Ellemberg, D. (1997). Lateral interactions in peripherally viewed texture arrays. *Journal of the Optical Society of America A*, 14, 2057-2068. [PubMed]
- Xing, J., & Heeger, D. J. (2000). Center-surround interactions in foveal and peripheral vision. *Vision Research*, 40, 3065-3072. [PubMed]
- Xing, J., & Heeger, D. J. (2001). Measurement and modeling of center-surround suppression and enhancement. *Vision Research*, 41, 571-583. [PubMed]
- Yu, C, Klein, S.A, & Levi, D.M. (2002). Unpublished observations.