Current Biology 19, 1988–1993, December 15, 2009 ©2009 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2009.09.056

Report

Crowding in Peripheral Vision: Why Bigger Is Better

Dennis M. Levi[1,*] and Thom Carney[1]
[1]School of Optometry and Helen Wills Neuroscience Institute. University of California, Berkeley, Berkeley, CA 94720-2020, USA

Summary

We enjoy the illusion that visual resolution is high across the entire field of vision. However, this illusion can be easily dispelled by trying to identify objects in a cluttered environment out of the corner of your eye. This reflects, in part, the well-known decline in visual resolution in peripheral vision; however, the main bottleneck for reading or object recognition in peripheral vision is crowding. Objects that can be easily identified in isolation seem indistinct and jumbled in clutter. Crowding is thought to reflect inappropriate integration of the target and flankers in peripheral vision [1, 2]. Here, we uncover and explain a paradox in peripheral crowding: under certain conditions, increasing the size or number of flanking rings results in a paradoxical decrease in the magnitude of crowding—i.e., the bigger or more numerous the flanks, the smaller the crowding. These surprising results are predicted by a model in which crowding is determined by the centroids of ≈4–8 independent features within ≈0.5× the target eccentricity. These features are then integrated into a texture beyond the stage of feature analysis. We speculate that this process may contribute to the illusion of high resolution across the field of vision.

Results

The faulty-integration model for crowding is based on the notion that crowding occurs because the visual system erroneously combines signals from the target and the flankers and that this combination is compulsory (for recent reviews, [1] and [2]). One prediction of the faulty-integration model is that increasing the size of the flankers should result in increased crowding. We tested this hypothesis by varying the size of the flankers in a simple orientation-discrimination task. The target was a high-contrast Gabor patch presented at an eccentricity of 5° in the lower visual field, and the flankers were composed of four segments of an annular grating (Figure 1). In the first experiment, the position of the outer edge of the flanks was fixed, and the size of the flanks was increased by varying the inner radius. Increasing flank size resulted in monotonically worse performance (increased threshold elevation, Figure 1A), consistent with the faulty-integration hypothesis [1–3]. However, increasing the flank size also brought the inner edge of the flank closer to the target, and previous work has suggested that crowding is due to the inner edge of the flank [4].

In the second experiment, we fixed proximity by keeping the position of the inner edge of the flanks constant and increased the flank size by varying the outer radius. Increasing flank size resulted in a monotonic decrease in the magnitude of crowding—i.e., the bigger the flanks, the smaller the crowding (Figure 1B). The faulty-integration model predicts that the bigger (or more numerous) the flanks, the bigger the crowding. Our unexpected result suggests that the faulty-integration hypothesis may be faulty. To clarify this, we conducted a third experiment to test whether target-flank distance might determine the strength of crowding by fixing the flank size and varying the target-flank distance (Figure 1C, circles). Interestingly, all three data sets superimpose when plotted as a function of the center-to-center distance between target and flanks. Thus, it appears that it is the distance, not the size of the flankers, that determines the strength of crowding. When the flanks are centered less than ∼2.2° from the target (arrow in Figure 1C), crowding occurs, independent of flanker size. This target-flank distance represents the critical spacing for crowding at 5°. We have replicated these three experiments at other eccentricities (E) and find that the critical spacing is ∼0.5E (Figure 1D for E = 2.5°, 5.0°, and 10.0°), consistent with many previous studies [1, 2, 5, 6] and with the notion that the critical spacing represents a fixed distance on the cortex [7–10].

Control experiments (e.g., target detection [Figure 1C], radial/tangential anisotropies [see Figure 3] and size independence [see Figure S1B available online]) verify that the threshold elevation is due to crowding rather than masking or surround suppression (see [1] and [2]). Taken together, our results provide an explanation for the unexpected effect of flank size—increasing the flank size (with a fixed inner radius) increased the center-to-center distance between target and flankers. Crowding is determined by the centroids of features (both targets and flankers) within ∼0.5× the target eccentricity; however, this provides little insight into the integration process per se.

A second prediction of the faulty-integration hypothesis is that two or more rings of flanking segments should be more potent than one. To test this hypothesis, we fixed the size and position of one ring of flanks and added a second ring of flanks (Figure 2A). The rightmost point shows the threshold elevation with the inner flanker ring alone. The faulty-integration theory predicts that adding a second ring of flanks should increase crowding. However, adding the outer flanks reduces crowding. The closer the outer flank to the target, the smaller the crowding. This result is counterintuitive, and on first glance, it appears to represent an instance of disinhibition. In this view, crowding is due to flanks inhibiting the target location [11]. The outer flanks inhibit the inner flanks, thus reducing the inhibition. However, we think it can be understood in much the same way as the "surprising" effect of increasing flank size while fixing the outer edge of a single flanking ring. When the distance of the second flank is less than ∼2.5° (0.5E) from the target, it "blends" with the inner flank. Indeed, at the smallest distance, it is contiguous with the inner flank. Thus, the visual system treats the two flanks as one single large texture, effectively increasing the target-flank distance and reducing crowding because the centroid of the texture moves further away from the target. As the distance between the two flanking rings increases, the two flanks become more distinct as

*Correspondence: dlevi@berkeley.edu

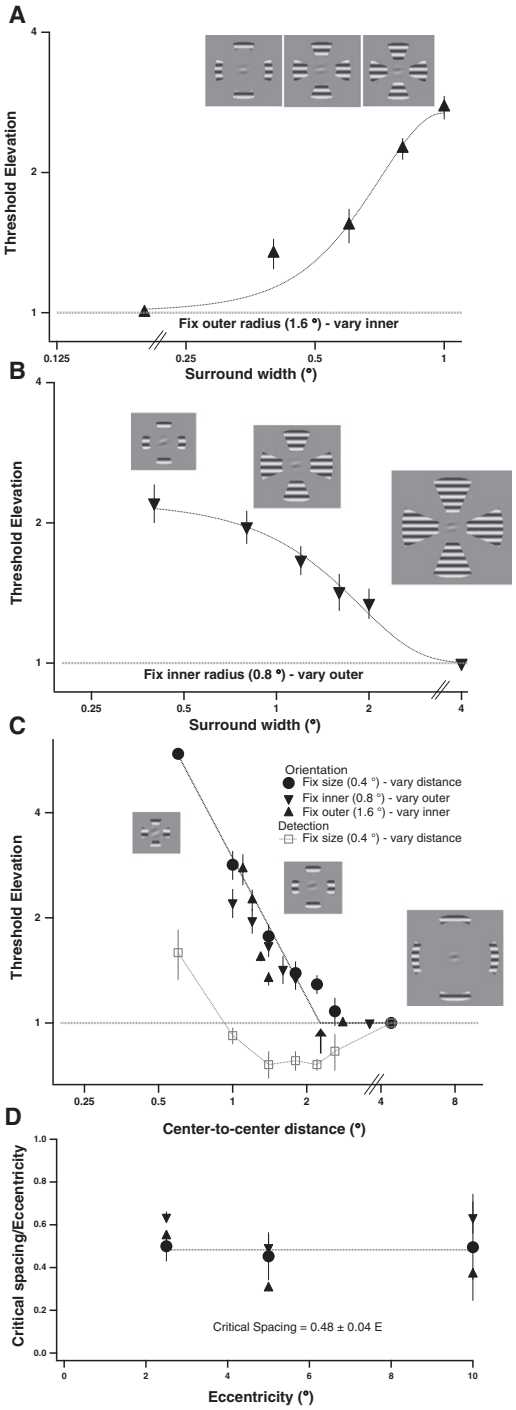

Figure 1. Testing the Faulty-Integration Model by Varying the Size and Location of the Flanks

(A) Fixing the position of the outer edge of the flanks (at 1.6°) and increasing the size of the flanks by varying the inner radius results in increased crowding (threshold elevation).

(B) Fixing the position of the inner edge of the flanks (at 0.8°),and increasing the size of the flanks by varying the outer radius results in decreased crowding.

(C) Fixing the flank size and increasing the center-to-center distance results in decreased crowding (circles). The individual threshold data are shown in Figure S1. Replotting (A) and (B) in terms of the flank center-to-center distance superimposes the three data sets. The dotted line is a two-line fit to the data. We define the critical spacing as the point of intersection of the two lines (arrow). Open squares show that except at the closest distance at which the target and flanks overlap, flankers reduce detection thresholds (i.e., facilitate detection of the target).

(D) Critical spacing specified as a fraction of the target eccentricity for each of the three experiments at eccentricities of 2.5°, 5°, and 10° in the lower visual field.

separate objects, and the amount of crowding approaches that of a single flank. Interestingly, two flanking rings are never more potent than one at producing crowding.

Our results thus far provide little evidence that integration of flanking features increases crowding. Several previous studies in which either the size or the number of flankers increase have either failed to show significant increase in crowding [12] or shown a reduction of crowding. In each of these studies (including ours), varying the size or number of flankers has increased the eccentricity of the flanks and therefore increased their effective spacing. On the contrary, varying the number of flankers while maintaining a fixed spacing does have an influence on the strength of crowding [13–15]. Figure 2B shows the effect of varying the number (and location) of flanking segments while maintaining a fixed target-flank spacing. Increasing the number of flank segments from one to four has a substantial effect, following the prediction of linear integration of independent segments (dashed line). However, increasing the number of segments from 4 to 8 (a full annular surround) actually results in slightly reduced crowding, and substantially less than predicted by the integration theory. Figure 2B also illustrates the well-established anisotropy in crowding—not only is the critical spacing larger when two flanks are arranged radially (propellers) rather than tangentially (bowties) [2, 16] but the magnitude of crowding is also stronger too.

Why does doubling the number of segments from four to eight result in decreased rather than increased crowding? Increasing the number of segments to eight results in a single coherent annulus (a texture), rather eight distinct segments. Would making each segment distinct increase crowding? To test this notion, we created an annular surround consisting of either four or eight wedges of grating with variably sized gaps of uniform gray between the wedges (Figure 2C). Increasing the gap width results in increased crowding (Figure 2C, lower abscissa). This is paradoxical because the larger the gap, the smaller the amount of flanking grating (Figure 2C, top abscissa); however, it is consistent with the notion that integration requires independent flanking samples. With the optimal gaps, eight wedges results in stronger crowding than four wedges (≈17%; Figure 2C, compare black with gray symbols), less than the 41% ($\sqrt{2}$) predicted by integration theory. Thus, there may be a linear integration limit of between four and eight independent samples (see [1] and [2] for a detailed account of the integration model).

Discussion

There are many theories for crowding, from low-level lateral inhibition of the target by the flanking patterns [11] to high-level attentional bottlenecks [17]. However, there is a growing consensus for a two-stage model in which the first stage involves the detection of simple features (perhaps in V1) and a second stage is required for the combination or interpretation of the features as an object downstream from V1 [1–3]. Here, we investigated the spatial layout of this "faulty integration." Our key finding, that crowding depends on the distance between the centroids of objects, not the proximity of their edges (Figure 1C), is novel and provides a simple explanation

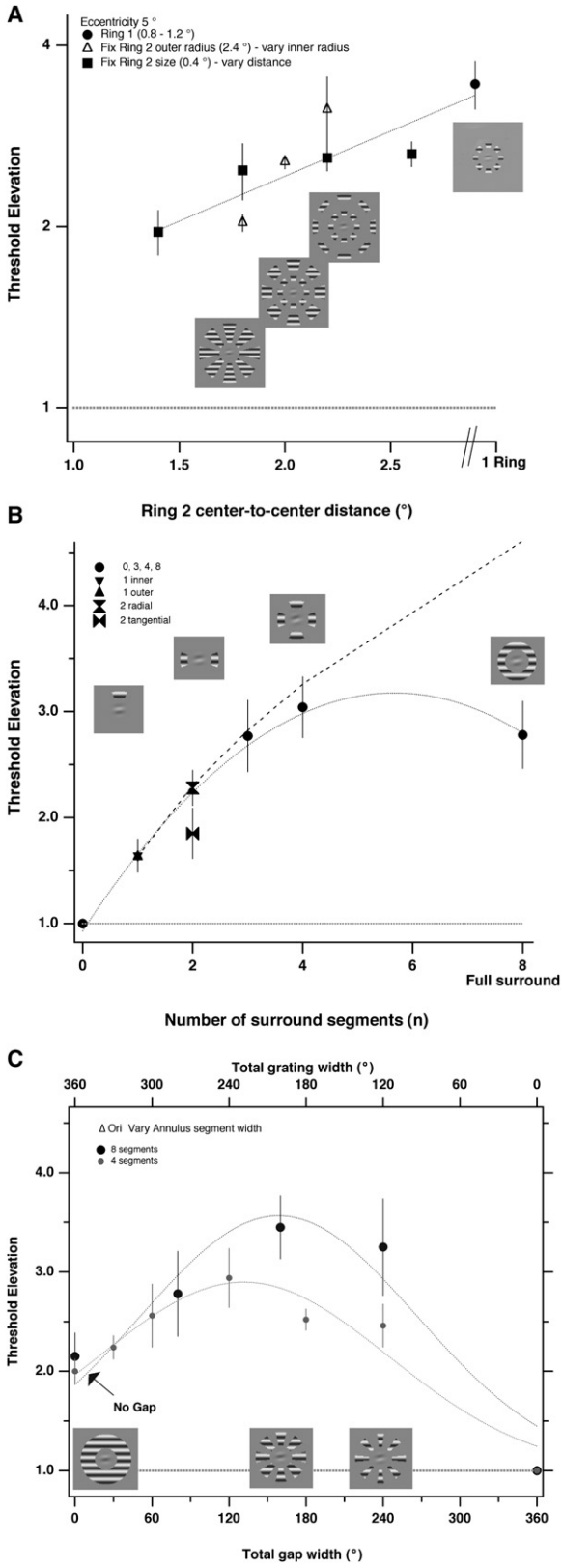

Figure 2. Testing the Faulty-Integration Model by Varying the Number of Flanking Rings, Segments, and Gaps

(A) Two-ring experiments. We fixed the size and position of the inner flanks and added a second set of flanks (i.e., two flanking rings). In one experiment, we fixed the size of the second ring and varied its distance (squares); in the other, we fixed the outer edge of the second ring and varied the flank size (triangles). The circle shows the threshold elevation with the inner flank alone.

(B) Varying the number (and location) of flanking segments. Increasing the number of flank segments (n) from one to four has a substantial effect, in keeping with the prediction of linear integration of independent segments (black dashed line). Increasing the number of segments from four to eight (a full annular surround) results in slightly reduced crowding. The dashed line shows the prediction of linear integration of independent segments, with threshold increasing in proportion to $\sqrt{n}$ for $n > 1$ ($n = 1$ is taken as the measured elevation with a single flank). The insets show examples of one "inner" flank (i.e., between the fixation point [which was 5° above the target]) and the target; two "tangential" flanks; and four and eight flanks.

(C) Gaps in the annulus. The flanks consisted of an annular surround consisting of either four (small circles) or eight (large circles) wedges of grating with variable sized gaps of uniform gray between the wedges. Increasing the gap width (lower abscissa) results in increased crowding. The larger the gap, the smaller the amount of flanking grating (top abscissa).

for the paradoxical finding that the bigger the flankers, the smaller the crowding and that two rings of flankers are less potent than one.

A Set-Size Effect in Crowding

Our results show a "set size" effect—increasing the number of flanks from 1 to ≈4 resulted in linear integration of independent features arranged in a ring around the target consistent with Põder and Wagemans [15]. Within ≈0.5× the target eccentricity (the "combination field" [9]), the visual system integrates ≈4–8 independent features into a crowded percept.

Grouping and Configural Effects in Crowding

In peripheral vision, there is a predilection for objects to be "joined together by the visual system into a single Gestalt" [18–20]. Thus, crowding can be "relieved" with a "grouped" mask, i.e., multielement flankers that group together separately from the target, thus reducing crowding [18–25]. Here, we show that simply inserting gaps (blank wedges) in the flanking ring increases the amount of crowding (Figure 2C), because the gaps reduce the appearance of a coherent texture, and provide independent samples of the flanking features. We do not believe that it is the gaps (and associated broadband edges per se) that result in increased crowding. Filling the gaps with horizontally oriented grating has the same effect, consistent with Livne and Sagi [23], who found that crowding was weakest (or absent) when the flank orientations varied smoothly around a circle, creating the impression of a closed contour, and strongest when the smooth variation in flank orientation was interrupted by changing the orientations of four of the eight flanks. Moreover, we have confirmed the configural coherence effect with a different approach based on altering the stimulus temporal properties. Synchronously flickering wedges appear as a coherent flickering surround, whereas asynchronously flickering wedges do not (see Movie S1). When all eight wedges of a flanking ring flickered asynchronously, crowding was ≈40% stronger than when they flickered synchronously (threshold elevation = 2.98 ± 0.22 versus 2.15 ± 0.20 times threshold elevation, respectively; $p < .004$). These configuration effects modulate the strength of crowding at a fixed separation and may thus act independently on the basis of object information [23].

Similarity and Conspicuity in Crowding

When targets and flankers are similar they may group together, and it is well known that when the target and flankers are "ungrouped" from each other by making them dissimilar in

a fundamental property such as color, polarity or depth so that the target "pops out," crowding is greatly reduced [22, 25–28]. We tested how dissimilar our target and flankers had to be in orientation and spatial frequency in order to "break" the texture and allow individual features to "pop-out" completely free of crowding. We show here that the target orientation must differ from those of the flanks by ≈90° (Figure S2A) or their spatial frequencies must differ by about a factor of 4 (full width at half-height ≈2 octaves; Figure S2B), consistent with previous work [29].

The Locus of Crowding

One exception to the similarity rule is contrast. Crowding is strongest when the target contrast is lower (not equal to) the flank contrast [28, 29]. We measured the effects of flank contrast at different target contrasts (Figure 3A). Like Pelli et al. [14], we note that flanks become effective once they are visible. What is new here is that the effect of flank contrast is independent of target contrast. Thus, at low contrasts (below ≈30%) the effect of flank contrast is essentially identical at all target contrast levels. We quantified the critical contrast for crowding (the minimum contrast required to induce crowding) with a two-line fit (the gray line in Figure 3A). The critical contrast for each target contrast (obtained by separate fits to each data set; Figure 3B) is essentially independent of target contrast. The slope of the best fitting line is −0.04 ± 0.05—consistent with a slope of zero.

If the flankers acted at the level of processing of the target feature, we would expect the critical contrast to be proportional to the target contrast (the solid line in Figure 3B shows the predicted critical contrast based on Weber's law [slope = 1]; a Legge-Foley power law [30] would predict a slope of ≈0.6). On the contrary, if crowding takes place at a higher level (beyond that of individual features) at which textural combination occurs, the effect of flank contrast should be independent of the contrast of the target feature, and such an outcome is just what we found. Previous work has shown that the strength of crowding depends on flank contrast rather than similarity (i.e., pop-out does not occur when target and flanks differ in contrast) [28, 29]; here, we show that the critical contrast for crowding is independent of target contrast.

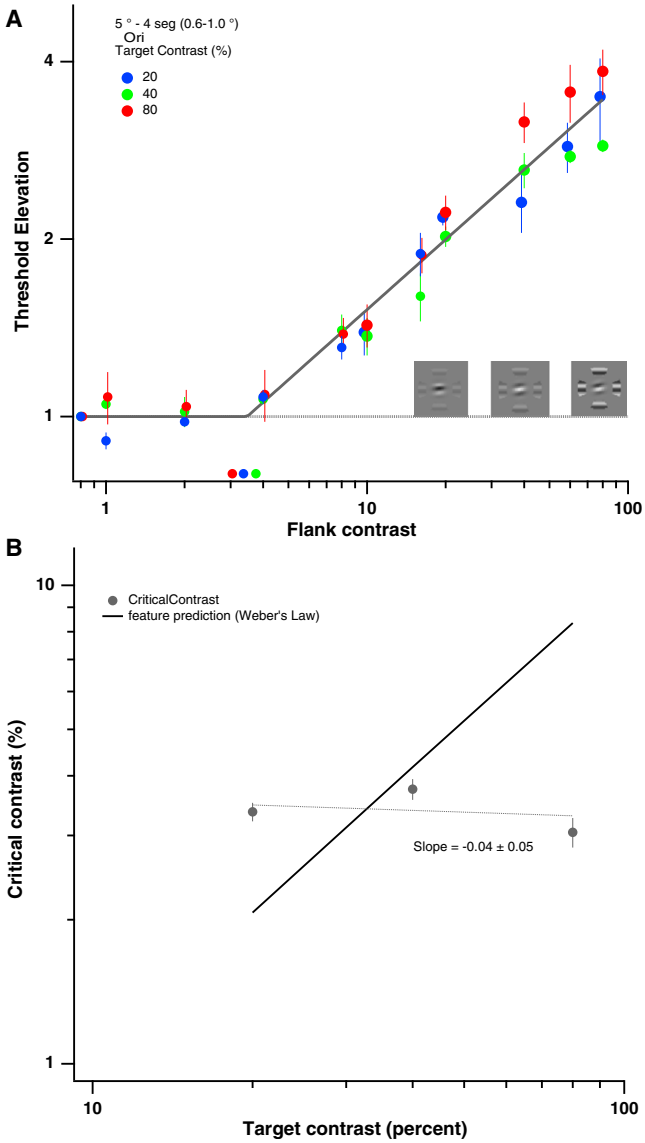

Figure 3. The Effect of Target and Flank Contrast

(A) The effect of target and flank contrast. The effect of flank contrast at three target contrasts—20% (blue), 40% (green), and 80% (red). The solid gray line is the best-fitting two-line fit to the entire data set. The fit is constrained to have a slope of 0 and a value of 1 at low contrast levels. The "critical contrast" is the contrast at which threshold elevation begins to increase. Critical contrast for each target contrast is shown by the colored symbols along the abscissa.

(B) The critical contrasts from separate two-line fits to each data set (target contrast) versus target contrast. The black line with unity slope shows the predicted values if the flankers acted at the level of the target feature. The dotted gray line is the slope of the best-fitting line.

The Illusion of High-Resolution Vision across the Field of Vision

With multiple saccades, we build up a representation of the visual field, much of it (in the periphery) at low resolution, but with knowledge of a world in high detail. At an eccentricity of only 5°, the ability to resolve an isolated letter or grating patch has diminished by about a factor of 3 or 4—from ≈1 arc minute in the fovea to 3 or 4 arc minutes [31]. However, it is not this reduced resolution that limits object perception in the periphery—it is crowding. Features up to several degrees from the target (≈2.5° for a target at 5°) further impair object recognition, rendering details of individual patterns unavailable for discrimination [12, 32]. Indeed, in a cluttered display, orientation signals are pooled rather than being lost through masking [13]. Although the simple integration model [13] is likely to be incomplete or incorrect in detail, models of crowding almost all invoke the extraction of independent features and some form of pooling or integration [1–3, 33, 34]. Our speculation, based on the centroid account of our data, is that the integration limit of ≈4–8 independent features may be closely related to the limited storage capacity of visual short-term memory in parietal cortex [35]. Beyond this limit, crowding "flattens" the percept of individually resolvable features into a coherent texture.

Experimental Procedures

Stimuli

The target was a briefly presented Gabor patch with horizontal carrier presented at an eccentricity of 2.5°, 5°, or 10° in the lower visual field. The carrier spatial frequency was 2.5 cycles/degree (c/°), and the Gaussian standard deviation of the patch was 0.4° when presented at 5° eccentricity and

was scaled according to eccentricity. Control experiments show that the critical spacing at a given eccentricity is independent of target size and spatial frequency. The target phase was randomized from trial to trial. The flanks were segments of an annular grating. In most experiments, the flank grating was horizontal and the spatial frequency was the same as the target (2.5 c/° at 5°). Target and flanks were presented simultaneously for 250 ms. For the orientation discrimination task, the target contrast was 40%, and unless otherwise specified, flank contrast was 80%. The stimuli were generated with the WinVis Psychophysical testing platform and were presented on a CRT screen (Sony Multiscan G400) with a mean luminance of 80 cd/m^2.

Psychophysical Methods

The main experiments used a two-interval forced-choice (2IFC) orientation-discrimination task. On each trial, the reference (0°) and test (reference plus an orientation offset) were separately presented in the two brief stimulus intervals (250 ms each) in a random order separated by a 500 ms interstimulus interval. An observer's task was to judge which stimulus interval contained the more counterclockwise Gabor. Auditory feedback was given after each correct response. For reduction of uncertainty about the target location, each trial was preceded by two large 37° by 0.1° diagonal lines forming an "X" with a circular gap ($\approx$0.6 E in diameter) in the center, so that it did not extend into the area of the target and flankers. These lines disappeared during each stimulus interval. The target always appeared in the center of the gap. Detection thresholds were measured for the same stimuli and flankers with the same 2IFC staircase method.

Each staircase consisted of eight reversals with thresholds based on the last six reversals. The initial orientation difference between the test and the reference was sufficiently large, so that the observers could always make a correct discrimination. A classical three-down-one-up staircase rule was used, which resulted in a 79.4% convergence level. The geometric mean of the six reversals was taken as the threshold estimate for each staircase run. In a typical experiment, four to eight staircases were randomly interleaved, so that all flank conditions (including no flanks) were tested in the same run. Each experiment was replicated six to eight times, and the thresholds are the geometric means of the six to eight threshold estimates.

Our index of crowding is the "Threshold Elevation" (i.e., the flanked thresholds divided by the unflanked thresholds). A threshold elevation of 1 indicates no crowding.

Observers

Five adults with normal or corrected-to-normal vision served as observers (although not all observers performed every experiment). All testing was monocular with the untested eye patched (to allow comparison with future amblyopic observers). Observers were given substantial practice prior to data collection. The data reported are mean threshold elevation data averaged across observers. Individual data followed the same trends. Figure S1 shows individual data for the "fixed size" experiment (circles in Figure 1c). Figure S1A shows the "raw data" (with orientation thresholds in degrees), and Figure S1B shows the individual data normalized by the unflanked thresholds (i.e., plotted as the Threshold Elevation).

Supplemental Data

Supplemental Data include two figures and one movie and can be found with this article online at http://www.cell.com/current-biology/supplemental/S0960-9822(09)01769-2.

Acknowledgments

This work was supported by grants R01EY01728 and R01EY04776 from the National Eye Institute, NIH. We thank S. Chung, S. Song, E. Rislove, and S. Klein for their helpful comments on an earlier version of the paper.

Received: July 9, 2009
Revised: September 18, 2009
Accepted: September 21, 2009
Published online: October 22, 2009

References

1. Levi, D.M. (2008). Crowding–an essential bottleneck for object recognition: A mini-review. Vision Res. *48*, 635–654.
2. Pelli, D.G., and Tillman, K.A. (2008). The uncrowded window for object recognition. Nat. Neurosci. *11*, 1129–1135.
3. Levi, D.M., Hariharan, S., and Klein, S.A. (2002). Suppressive and facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. J. Vis. *2*, 167–177.
4. Takahashi, E. (1968). Effects of flanking contours on visual resolution at foveal and near-foveal loci. PhD Thesis. U.C. Berkeley.
5. Bouma, H. (1970). Interaction effects in parafoveal letter recognition. Nature *226*, 177–178.
6. Andriessen, J.J., and Bouma, H. (1976). Eccentric vision: Adverse interactions between line segments. Vision Res. *16*, 71–78.
7. Levi, D.M., Klein, S.A., and Aitsebaomo, A.P. (1985). Vernier acuity, crowding and cortical magnification. Vision Res. *25*, 963–977.
8. Motter, B.C., and Simoni, D.A. (2007). The roles of cortical image separation and size in active visual search performance. J. Vis. *7*, article 6, 1–15.
9. Pelli, D.G. (2008). Crowding: A cortical constraint on object recognition. Curr. Opin. Neurobiol. *18*, 445–451.
10. Liu, T., Jiang, Y., Sun, X., and He, S. (2009). Reduction of the crowding effect in spatially adjacent but cortically remote visual stimuli. Curr. Biol. *19*, 127–132.
11. Polat, U., and Sagi, D. (1993). Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. Vision Res. *33*, 993–999.
12. Wilkinson, F., Wilson, H.R., and Ellemberg, D. (1997). Lateral interactions in peripherally viewed texture arrays. J. Opt. Soc. Am. A Opt. Image Sci. Vis. *14*, 2057–2068.
13. Parkes, L., Lund, J., Angelucci, A., Solomon, J.A., and Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. Nat. Neurosci. *4*, 739–744.
14. Pelli, D.G., Palomares, M., and Majaj, N.J. (2004). Crowding is unlike ordinary masking: distinguishing feature integration from detection. J. Vis. *4*, 1136–1169.
15. Põder, E., and Wagemans, J. (2007). Crowding with conjunctions of simple features. J. Vis. *7*, article 23, 1–12.
16. Toet, A., and Levi, D.M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. Vision Res. *32*, 1349–1357.
17. He, S., Cavanagh, P., and Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. Nature *383*, 334–337.
18. Banks, W.P., Larson, D.W., and Prinzmetal, W. (1979). Asymmetry of visual interference. Percept. Psychophys. *25*, 447–456.
19. Banks, W.P., and White, H. (1984). Lateral interference and perceptual grouping in visual detection. Percept. Psychophys. *36*, 285–295.
20. Estes, W.K., Allmeyer, D.H., and Reder, S.M. (1976). Serial position functions for letter identification at brief and extended exposure durations. Percept. Psychophys. *19*, 1–15.
21. Põder, E. (2006). Crowding, feature integration, and two kinds of "attention". J. Vis. *6*, 163–169.
22. Saarela, T.P., Sayim, B., Westheimer, G., and Herzog, M.H. (2009). Global stimulus configuration modulates crowding. J. Vis. *9*, article 5, 1–11.
23. Livne, T., and Sagi, D. (2007). Configuration influence on crowding. J. Vis. *7*, article 4, 1–12.
24. Louie, E.G., Bressler, D.W., and Whitney, D. (2007). Holistic crowding: Selective interference between configural representations of faces in crowded scenes. J. Vis. *7*, article 24, 1–11.
25. Malania, M., Herzog, M.H., and Westheimer, G. (2007). Grouping of contextual elements that affect vernier thresholds. J. Vis. *7*, article 1, 1–7.
26. Nazir, T.A. (1992). Effects of lateral masking and spatial precueing on gap-resolution in central and peripheral vision. Vision Res. *32*, 771–777.
27. Põder, E. (2007). Effect of colour pop-out on the recognition of letters in crowding conditions. Psychol. Res. *71*, 615–715.
28. Kooi, F.L., Toet, A., Tripathy, S.P., and Levi, D.M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. Spat. Vis. *8*, 255–279.
29. Chung, S.T., Levi, D.M., and Legge, G.E. (2001). Spatial-frequency and contrast properties of crowding. Vision Res. *41*, 1833–1850.
30. Legge, G.E., and Foley, J.M. (1980). Contrast masking in human vision. J. Opt. Soc. Am. *70*, 1458–1471.
31. Levi, D.M., Klein, S.A., and Aitsebaomo, A.P. (1984). Detection and discrimination of the direction of motion in central and peripheral vision of normal and amblyopic observers. Vision Res. *24*, 789–800.

32. Orbach, H.S., and Wilson, H.R. (1999). Factors limiting peripheral pattern discrimination. Spat. Vis. *12*, 83–106.
33. Baldassi, S., Megna, N., and Burr, D.C. (2006). Visual clutter causes high-magnitude errors. PLoS Biol. *4*, e56.
34. Palmer, J., and Moore, C.M. (2009). Using a filtering task to measure the spatial extent of selective attention. Vision Res. *49*, 1045–1064.
35. Todd, J.J., and Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. Nature *428*, 751–753.