Vision: Seeing through the Gaps in the Crowd

Much of our visual experience is an 'illusion of completeness', without true conscious access to the details of objects in a typically cluttered natural scene. Recent results on this 'crowding' phenomenon are starting to bridge the gap between the striking limits of perception and the subjective impression of a rich visual world.

David Whitney

Conscious visual perception is more limited than we think. The limits are not simply imposed by acuity (the resolution of vision) but also by crowding — an impairment in the recognition of individual objects located in cluttered scenes. Figure 1A, for example, shows a natural scene filled with clutter, typical of our moment-to-moment visual experience. When fixating on the white cross, some characteristics of the scene are readily apparent — the visual textures and gist of the scene, for example. When probed more deeply, however, it is clear that we have very little access to the details or particulars; identifying the types of dish, or the number of pieces of silverware in the center of the image is difficult or impossible. This difficulty is the phenomenon of crowding. We can easily confirm that this is not due to limited acuity, because the individual objects are recognizable in the periphery when isolated from the clutter (Figure 1A,B). Because of crowding, we must make eye movements to individuate and scrutinize objects in naturally cluttered scenes — a costly and time consuming consequence of limited resources.

Crowding happens in the periphery of almost every natural scene that we look at, though it is usually studied in simplified contexts, using basic visual features such as lines, edges, gratings and letters (Figure 1B–1D). Experiments using these sorts of stimuli have revealed several defining characteristics of crowding, including density and eccentricity dependence — crowding usually happens when the spacing between features or objects is less than half the eccentricity of the object (Bouma's law [1–3]). Crowding is also stronger when the flanking object is on the far (more eccentric) side of the target than on the near (foveal) side of the target [1,4], and

is stronger in the upper visual field [5], but does not occur at the fovea [6].

There is no shortage of models for crowding, ranging from over-integration or pooling of low level features, to attentional resolution (for reviews, see [2,3,5]). The abundance of hypotheses reveals the continuing struggle to identify a single neural mechanism and computational model that can account for this fundamental limit on object recognition. A single neural mechanism, however, may be a misguided search. Among many recent results, three recent ones, in particular, suggest that our previous understanding of the locus and homogeneity of the phenomenon of crowding requires revision.

More Than Just Blind Over-Integration
In this issue of *Current Biology*, Levi and Carney [7] report that, paradoxically, increasing the size or number of flankers can, under certain circumstances, decrease crowding (Figure 1C). The authors found that crowding varied systematically as a function of the center-to-center distance between objects, independent of the size of the flankers or their edge location. Increasing the number of flankers beyond roughly four to eight also decreased crowding [7]. In their simplest form, models that invoke spatial integration, pooling or attentional resolution would generally predict that increasing the size or number of flanking (distractor) elements should increase crowding [2,3,5]. Levi and Carney [7] suggest, in contrast, that crowding occurs because the visual system independently samples a limited number of distinct features (or objects or positions; cf [8]) before integrating them.

A Face in the Crowd
Two additional results challenge our current notion of what perceptual crowding really is. First, crowding can be reduced when flanking objects are

perceived as a spatial group [9,10] or a temporal group [7], suggesting that grouping happens before crowding
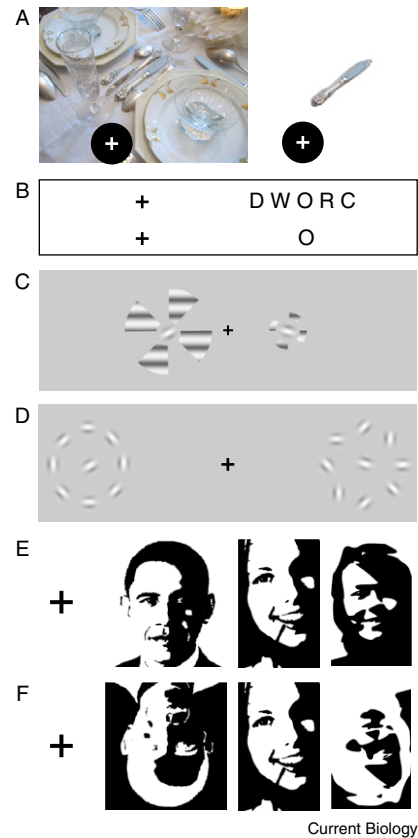

Current Biology

Figure 1. Visual crowding.
(A) While fixating on the white cross, it is extremely difficult to count or identify the number and type of silverware in the middle of the image. This difficulty is due to crowding — the deleterious effect of nearby clutter on object recognition. It is not due to poor acuity or visual resolution, as it is easy to identify the same object in the periphery when nearby clutter is removed (right side of panel A). Crowding happens for many features and objects, including letters (B): while fixating on the cross, identifying the middle letter in a string of peripheral letters is difficult, even though the letter is easily identified when presented in isolation (bottom of panel B). (C) Counterintuitively, increasing the size of flankers can reduce crowding. While fixating on the black cross, notice that the orientation of the central target Gabor patch is easier to recognize on the left than it is on the right. (D) Grouping precedes crowding. The central Gabor patch is easier to discriminate on the left than on the right, while fixating centrally, simply because the flanking Gabor patches group together on the left. (E,F) Holistic crowding. While fixating on the black cross, the central face is harder to recognize in (E) than it is in (F). This is due to crowding selectively between upright, holistic face representations, independent of the low-level features in the images.

(Figure 1D). Models of crowding that invoke low-level integration of features at an early stage of processing have difficulty accounting for this finding. Second, crowding can happen between features or objects at different levels of visual processing. For example, crowding happens selectively between holistic face representations, independent of the low-level feature-based crowding within those faces (Figure 1E,F). Both grayscale [11] and two-tone Mooney faces [12] are more crowded by upright flanker faces than they are by inverted ones.

This inversion effect in crowding is not simply 'similarity' in any low-level featural sense (for example [13,14]) — the similarity is in the 'faceness', the holistic nature of the face. Holistic crowding does not happen for cars or other non-face objects and it is not simply grouping (as in Figure 1D), because the flankers do not perceptually 'group' in a Gestalt sense any more with each other than they do with the target. Crowding cannot, therefore, be due to a single bottleneck; it cannot even be a single high-level bottleneck. There are layers of crowding: whatever mechanism or set of mechanisms contributes to crowding (be that integration, attentional resolution, positional averaging, and so on), the process must occur redundantly at multiple stages of visual processing.

The Crowded Explanatory Gap
A model of crowding that can account for the diversity of empirical findings will be a milestone, to be sure, but it may leave potentially unanswered the broader question of how we get from rarified percepts of objects in crowds to phenomenologically rich percepts of the world; despite crowding, we nonetheless have an 'illusion of completeness', we feel that the visual periphery is meaningful.

At the very end of their article, Levi and Carney [7] leave us with the intriguing speculation that crowding results in a 'flattened' percept. This idea is worth exploring. Perhaps there is a sensory threshold for what counts as 'rich' or 'meaningful' — within some limit, the default percept is that the peripheral visual field is organized and detailed (like the default percept is of a stationary world, even though the image of the world is constantly jittering on the retina because of small

eye movements and tremor [15]). The nature of the information on which this sensory threshold operates is unclear, but one possibility is summary statistics. Natural scenes are filled with similar objects, textures and features, resulting in the perception of ensembles (groups of trees, bricks, faces, flocks of soaring birds), and these may contribute to our rich perceptual experience of the world [16–18]. These ensembles are perceived whether or not crowding happens [17–20].

Thus, crowding may not be necessary for the illusion of completeness, but it may force the issue — obligating the visual system to efficiently compress (as opposed to filter or dismantle) the crowded information into a summary statistical representation. This happens at a number of independent levels ranging from low-level features and textures to high-level objects. Developing a model that accounts for the diverse effects of crowding while simultaneously bridging the explanatory gap between a 'flattened' percept and a rich visual impression of the world remains an important goal for vision science in the near future.

References
1. Bouma, H. (1970). Interaction effects in parafoveal letter recognition. Nature *226*, 177–178.
2. Levi, D.M. (2008). Crowding – an essential bottleneck for object recognition: a mini-review. Vis. Res. *48*, 635–654.
3. Pelli, D.G., and Tillman, K.A. (2008). The uncrowded window of object recognition. Nat. Neurosci. *11*, 1129–1135.
4. Petrov, Y., Popple, A., and McKee, S. (2007). Crowding and surround suppression: Not to be confused. J. Vis. *7*, 1–9.
5. He, S., Cavanagh, P., and Intiligator, J. (1997). Attentional resolution. Trends Cogn. Sci. *1*, 115–121.
6. 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.
7. Levi, D.M., and Carney, T. (2009). Crowding in peripheral vision: why bigger is better. Curr. Biol. *19*, 1988–1993.
8. Greenwood, J.A., Bex, P.J., and Dakin, S.C. (2009). Positional averaging explains crowding with letter-like stimuli. Proc. Natl. Acad. Sci. USA *106*, 13130–13135.
9. Banks, W.P., Larson, D.W., and Prinzmetal, W. (1979). Asymmetry of visual interference. Percept. Psychophys. *25*, 447–456.
10. Livne, T., and Sagi, D. (2007). Configuration influence on crowding. J. Vis. *7*, 1–12.
11. 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*, 21–11.
12. Farzin, F., Rivera, S.M., and Whitney, D. (2009). Holistic crowding of Mooney faces. J. Vis. *9*, 11–15.
13. Chung, S.T., Levi, D.M., and Legge, G.E. (2001). Spatial-frequency and contrast properties of crowding. Vis. Res. *41*, 1833–1850.
14. 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.
15. Murakami, I., and Cavanagh, P. (1998). A jitter after-effect reveals motion-based stabilization of vision. Nature *395*, 798–801.
16. Alvarez, G.A., and Oliva, A. (2009). Spatial ensemble statistics are efficient codes that can be represented with reduced attention. Proc. Natl. Acad. Sci. USA *106*, 7345–7350.
17. Chong, S.C., and Treisman, A. (2003). Representation of statistical properties. Vis. Res. *43*, 393–404.
18. Haberman, J., and Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. Curr. Biol. *17*, R751–R753.
19. Ariely, D. (2001). Seeing sets: representation by statistical properties. Psychol. Sci. *12*, 157–162.
20. 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.

The Department of Psychology, The University of California, Berkeley, CA, USA.
E-mail: dwhitney@ucberkeley.edu

Motors: Unleashing Mitochondria

Mitochondria can move along and interact with actin, yet the identity of the protein(s) mediating the interactions in metazoans is unknown. A new study reveals that a novel unconventional myosin, Myo19, is a mitochondria-associated motor that may play a role in either the transport or tethering of this organelle.

Margaret A. Titus

The eye-catching acrobatics of cytoplasmic organelles and their robust trafficking up and down axons provided strong impetus for cell

biologists to determine what drives these movements. The now classical images of organelles, such as mitochondria, trapped in a microtubule meshwork in the axon [1] led to intense studies on the role of microtubule