Early visual cortex: **Smarter than you think**
David Bradley

A recent study has provided elegant evidence that the early visual area V2 plays an important role in image segmentation, piecing together parts of an object with the help of stereoscopic clues.

Address: Psychology Department, University of Chicago, 5848 S. University Avenue, Green 314, Chicago, Illinois 60637, USA.
E-mail: dave@ccp.uchicago.edu

Two people are stranded in the Tower of London. One is smart, the other not too smart. Both have a bed sheet. The idiot rips up the sheet and ties the ends together, hangs it out the window, and slithers halfway down the tower before reaching the end. The smart one makes a parachute and glides safely to earth.

The brain is not unlike the smarter of these two, in that it makes the most of what is available to it. Take, for instance, the use of stereo information by the visual system. Stereo cues, which derive from the slight displacement of one retinal image relative to the other, give direct information about the depth (distance) of objects. The brain, over the millenia, has figured this out, and the result is a profound sense of depth when we look at something with two eyes. Anyone who has seen a three-dimensional movie will attest to this; what one generally does not realise, however, is that three-dimensional movies, not the ordinary two-dimensional kind, mimic natural vision.

But the brain does not stop there. In fact, it uses stereo cues in ways that go beyond simply measuring the depth of things. A striking recent example comes from studies by Bakin, Nakayama and Gilbert [1], who have shown that neurons in cortical area V2 use a number of tricks that allow them to piece together image segments by making sensible interpretations of stereo cues. Sound easy? It certainly is not. The problem of grouping and separating image segments, known broadly as segmentation, is among the hardest in computational vision.

In the family of cortical visual areas, V2 is a child, sitting just above V1 in the known functional and anatomical hierarchy but beneath a sizeable list of specialized areas which are generally thought to carry out the fancier aspects of visual processing (Figure 1) [2]. But the new study by Bakin *et al.* [1], an electrophysiological experiment carried out with trained macaque monkeys, suggests that V2 is anything but naïve.

Figure 1
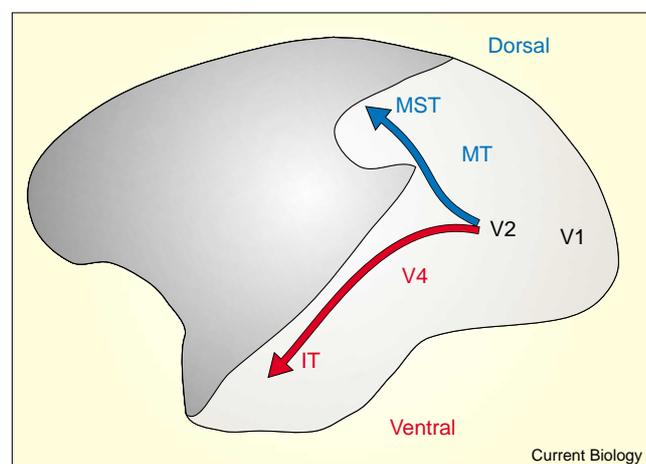
A schematic illustration of the primate brain, showing the dorsal (blue) and ventral (red) processing streams. The pathways bifurcate after V2. Area labels, such as V4, denote only approximate locations.

When humans are trying to see very dim lines, their task is easier if a colinear flanking line is included (Figure 2). The brain seems to understand that the two line segments are part of a larger line, and thus combines the signals from both. Putting a perpendicular line between the colinear segments tends to break the percept, and one loses the detection advantage from the flank. But if one then makes the perpendicular line appear closer than the colinear segments, using stereo cues, the disruptive effect goes away; presumably, the brain knows that continuous lines are bound to disappear occasionally behind occluders.

This knowledge appears to reside in V2. These neurons, like all visual neurons, see the world through windows known as receptive fields. An appropriately oriented bar in a V2 neuron's receptive field makes it respond (discharge action potentials), and a flanking bar makes the response bigger. This resembles the perceptual effect, where a target line with a flank is easier to see than the target line by itself. Bakin *et al.* [1] found that a perpendicular bar between the colinear segments decreases this response unless its stereo tag places it in front of the colinear segments (Figure 2, lower left). In this case, the neuron again seems to respond as if to the larger contour defined by the central and flanking colinear segments — that is, without the interrupting effect of the perpendicular segment. This, again, resembles the perceptual effect.

Figure 2
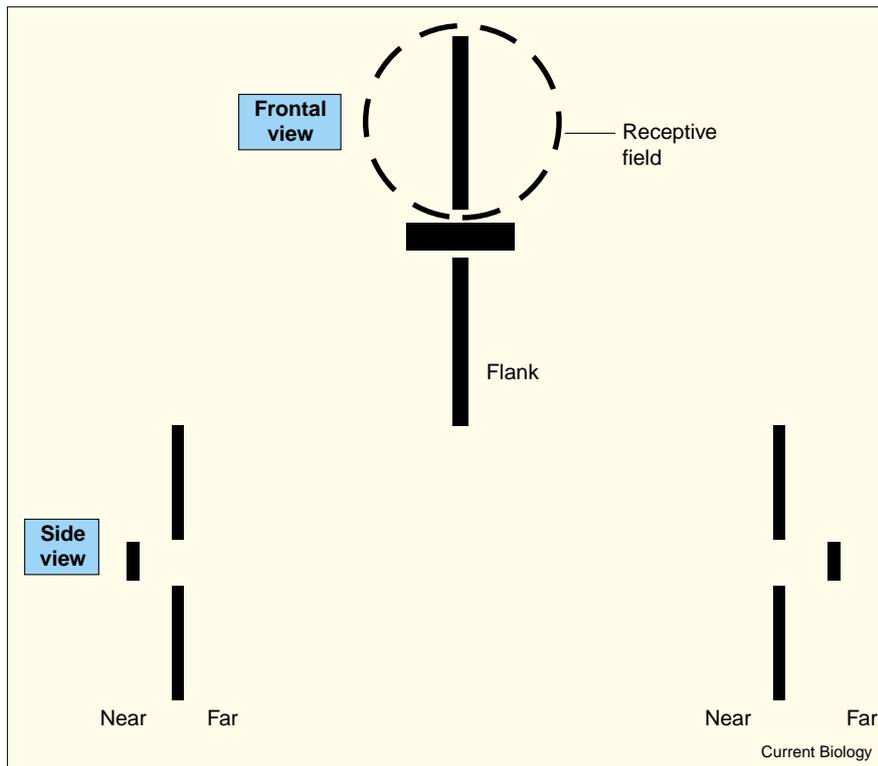
The colinear flank experiment carried out by Bakin *et al.* [1]. The perpendicular bar disrupts the facilitative effect of the flank on V2 firing rates, except when it appears closer (lower left), which is consistent with occlusion in nature.

There is another interesting perceptual effect that occurs when an object is partly obscured by its background. This happens when the object and its background have similar brightness, so that there is no contrast to define an edge between them. But we often perceive an edge anyway, and this edge is known as an illusory contour (Figure 3, left). This is a useful illusion to have — we cannot allow gaps in our perception every time part of an edge blends with its background.

Seminal experiments by Von der Heydt and colleagues [3] showed that V2 neurons respond when illusory contours appear in their receptive field, cleverly filling in what they think is really there. But if these neurons were really clever, they would avoid doing this in some cases. Imagine, for example, looking at a square patch with two bars protruding on opposite ends (Figure 3, right). If the protruding bars were perfectly aligned, one would have the perceptual option of connecting them with illusory contours. But if the visible bars were stereoscopically farther than the patch, it would be silly to do this; we should not entertain the illusion of a visible edge when that edge is in reality coursing behind an occluder. V2 neurons are apparently in on this secret. When Bakin *et al.* [1] showed a patch and protruding bars, V2 neurons responded as if to an illusory contour when the bars were stereoscopically in front of the patch, but not when they were stereoscopically at the back. Not surprisingly, one perceives an illusory contour in the former case, but not the latter.

These experiments suggest that perceptual contour integration, whether it involves assuming what goes on behind an occluder or constructing a perceptual edge where one visually does not exist, depends in no small part on V2 processing. Under the same conditions, Bakin *et al.* [1] found that V1 neurons, by contrast, generally do not complete contours very well. This is pleasing to the neurophysiologist because it suggests we understand the functional contribution of V2 — above and beyond what is already done in V1, that is. This is not to say that V1 neurons are not clever; indeed, they often modulate their responses depending on things outside their receptive field, and according to another study [4], may carry out contour integration in ways similar to V2. But the plain differences between V1 and V2 observed by Bakin *et al.* [1] suggest, given the constancy of conditions, that V2, at least, is the contour integration expert.

One of the striking things about the work of Bakin *et al.* [1] is that only parts of the stimuli were actually placed in the neurons' 'classical' receptive field, which is the part of visual space where localized stimuli evoke a response. For example, the colinear flank in the first experiment

Figure 3

Illustration of illusory contours. On the left is shown a Kanisza triangle, which is perceived as having illusory edges in the regions between the 'pacman' figures. On the right is shown the simplified stimulus used by Bakin *et al.* [1]: if the vertical bar is placed stereoscopically closer than the square patch, it forms an illusory contour over the patch.
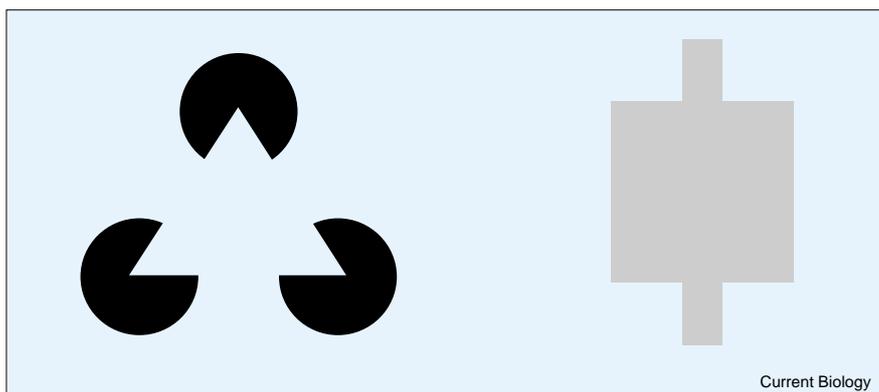
Current Biology

(Figure 2), when presented alone, caused no response, but it enhanced the response when presented in combination with a central line segment. This is called a surround effect. Similarly, in the illusory contour tests, the visible bars were completely outside the classical field. Surround effects have been appreciated for some time; the importance here is to emphasize that V2 neurons are not just interested in localized events, but rather in contour relationships that span their classical field and beyond.

The primate brain is generally thought of as having two functionally different streams, or pathways. The 'dorsal' stream climbs up and forward and includes, among others, areas V1, V2, MT and MST (Figure 1). All of these regions contain neurons that respond especially well to moving targets; in fact MT and MST are rich with them. For this reason, the dorsal stream is sometimes called the motion pathway. The 'ventral' stream descends away from the dorsal one, going forward and down into the inferior temporal lobe. It contains, among others, areas V1, V2, V4 and IT. These areas, in general, contain neurons sensitive to static object qualities, such as texture, shape and color. In fact the highest area, IT, has neurons that appear to recognize certain types of object. For these reasons, the ventral stream is sometimes called the object pathway.

Both pathways have neurons that are selective for — respond differentially to — stereo cues. In the case of the dorsal pathway, this has been something of a curiosity. In earlier thinking, stereo information was thought to be useful for describing depth, and thus three-dimensional form. What does that have to do with motion processing? One might think stereo information could be used to compute motion through depth — objects coming closer to, or moving away from, the observer. But our motion-sensitive neurons, as far as we know, sense only the left–right, up–down aspect of motion, and do not seem to use stereo information to compute motion through depth [5].

Perhaps one finds stereo signals in the dorsal pathway because — not excluding other functions — they facilitate image segmentation. After all, what good are motion signals to the brain if they cannot be sorted and grouped? The combined motion of several objects means nothing, so it is critical that motion cues be grouped and sorted according to which objects they come from. A similar argument holds, of course, for static objects; imagine perceiving someone's face as the disjointed appearance of various facial features. Obviously, it is the grouping of features that tells us about the face.

And so it is perhaps not coincidental that V2 sits at the branchpoint of the dorsal and ventral pathways, feeding both streams with the benefit of its computation. That benefit seems to involve basic grouping operations that link same-object contours with each other, critical information for defining the boundaries of image segments. This is not to say that segmentation is all done in V2. Area MT, for example, has segmenting functions that also rely on stereo cues [6,7], and various dorsal and ventral areas probably carry out different types of segmentation. But V2 may be at a crucial stage where basic linking operations precede more specialized computations that occur in higher dorsal and ventral areas.

Neurophysiologists often speak of linearity. In our tower example, the idiot uses a linear operation to make use of the sheet, in that two sheets would have got him twice as far. The smarter one uses a nonlinear operation in the sense that he got results that were more than the sum of what he started with. With two sheets, he might have flown to France. Area V2 is nonlinear like this — given a metric about luminance (edge contrast) and a metric about depth (stereo), V2 neurons condition one with the other, and in so doing are able to discover one of the most precious clues in vision: connection.

Now that is smart.

References

1. Bakin JS, Nakayama K, Gilbert CD: **Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations.** *J Neurosci* 2000, **20**:8188-8198.

2. Felleman DJ, Van Essen DC: **Distributed hierarchical processing in the primate cerebral cortex.** *Cereb Cortex* 1991, **1**:1-47.

3. von der Heydt R, Peterhans E, Baumgartner G: **Illusory contours and cortical neuron responses.** *Science* 1984, **224**:1260-1262.

4. Sugita Y: **Grouping of image fragments in primary visual cortex.** *Nature* 1999, **401**:269-272.

5. Maunsell JH, Van Essen DC: **Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity.** *J Neurophysiol* 1983, **49**:1148-1167.

6. Bradley DC, Andersen RA: **Center-surround antagonism based on disparity in primate area MT.** *J Neurosci* 1998, **18**:7552-7565.

7. Bradley DC, Qian N, Andersen RA: **Integration of motion and stereopsis in middle temporal cortical area of macaques.** *Nature* 1995, **373**:609-611.