

# Cortical mechanisms of binocular stereoscopic vision

A.J. Parker\* and B.G. Cumming

*University Laboratory of Physiology, Parks Road, Oxford, OX1 3PT, UK*

**Abstract:** The early neurophysiology of binocular vision is largely dominated by measurements of disparity selectivity in cortical neurons in various visual areas. Incisive progress has been made by the intensive study of the mechanism of disparity selectivity of V1 in cortical neurons and the development of a number of tests for the involvement of single neurons in the perception of stereoscopic depth. The picture that now emerges is that cortical area V1 must be a preliminary processing stage for the analysis of stereoscopic depth, whereas some of the extrastriate areas may actually be responsible for the generation of neuronal signals that underlie the perception of binocular depth.

## Introduction

This article is concerned with the way in which the visual areas of the cerebral cortex process information arising from the two eyes. It has long been established by psychophysical experiments that small differences in the images on the left and right retinæ are sufficient to generate a sensation of depth. The basic geometric relationships are outlined in Fig. 1. When the eyes are binocularly fixating a single target in the three-dimensional world, any object that is closer to or further from the depth plane of the fixation target falls on non-corresponding retinal points. Wheatstone (1838, 1852) demonstrated that binocular stimulation of this form is sufficient to provide observers with the impression that objects are at different distances. Julesz (1964, 1971) showed that this process is driven automatically by luminance edges or other low-level image features, in that no explicit recognition of form is required at the monocular level in order for the observer to experience binocular depth.

The intention here is to review recent progress towards understanding the neuronal basis of binocular depth perception and related phenomena. Much of the relevant evidence comes from studies of V1, the primary visual cortex. However, these experiments highlight some significant limitations of stereoscopic processing in area V1, a conclusion that points the way to a fresh look at the contributions of extrastriate cortex. Although many studies have measured the tuning characteristics for binocular disparity in extrastriate visual areas, it is argued here that a new range of tests needs to be applied to reveal more precisely the role of neurons in the various extrastriate areas.

Before embarking on these issues, we should address briefly the most significant issue of all, namely the functional role of stereoscopic vision and how it fits together with other important sensory and motor functions.

## Functions of stereoscopic vision

There can be no doubt that the fact that our species possesses two eyes bears no simple relationship to a single visual function. The bilaterality of eyes and ears clearly owes more to the midline symmetry of vertebrate development, than any specific selective advantage reinforced by evolutionary mechanisms.

---

\* Corresponding author: A.J. Parker, University Laboratory of Physiology, Parks Road, Oxford, OX1 3PT, UK. Tel.: +44-1865-272504; Fax: +44-1865-272469; E-mail: andrew.parker@physiol.ox.ac.uk

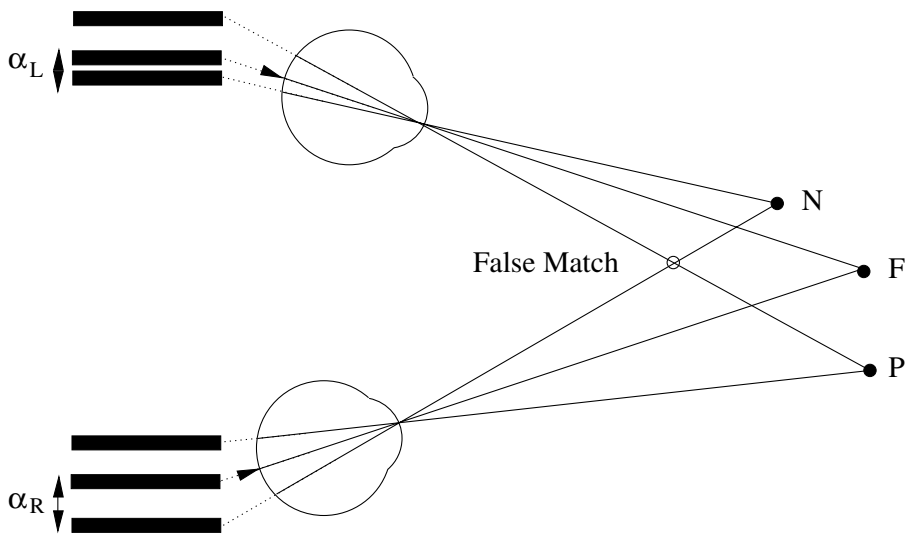


Fig. 1. Diagram to illustrate stereoscopic viewing geometry for horizontal disparities. The eyes are fixating point  $F$ . Point  $N$  has a non-zero disparity with respect to point  $F$  because its distance is less than  $F$ . Point  $P$  is at zero disparity because it is at the same distance as the fixation point. The angles subtended at the left and right eyes ( $\alpha_L$  and  $\alpha_R$ , respectively) by the gap between features  $N$  and  $F$  are different, whereas the angles between  $F$  and  $P$  are the same. The open circle marked 'False Match' shows that if  $N$  and  $P$  give rise to similar image features then there is the possibility of confusion about how to pair up the image features between the left and right eyes. The brain is remarkably adept at eliminating these potential confusions.

The different ways in which this symmetry has been exploited by lateral-eyed and frontal-eyed animals illustrates the opportunistic effect of biological selection. In lateral-eyed animals, such as the rabbit, the two eyes are used primarily to provide a panoramic view of the world with only a small region of binocular overlap. In frontal-eyed animals, such as primates, the basic sensory input from the two eyes is coordinated and exploited in a variety of ways.

### Bifoveation and vergence control

First, it is used for bifoveation, to ensure that the foveas of the two eyes remain locked on a single target, for example during a pursuit eye movement. Related to this function is the control of vergence position, which allows the system to acquire new targets at different depths or even to track the movement of a target in depth, albeit slowly compared with versional movements. This review will not examine the mechanisms responsible for generating these movements in any detail, but one aim is to identify candidate sensory control signals for the maintenance or adjustment of binocular vergence position.

### Figure-ground segregation

A second important use of binocular input is also unrelated to the perception of depth. This is the use of binocular vision for figure-ground segregation. A particular form of random dot stereogram illustrates this point (see Fig. 2). The stimuli to the left and right eyes are uniform in appearance but the fusion of the two images shows hidden structure. This arises because the central circular region is binocularly correlated between the left and right eyes' images, whereas the surround region is binocularly uncorrelated. In the real 3-D world, these differences of correlation are often associated with differences of depth. In the figure, there is no consistent depth difference between center and surround because the center is arranged to have a disparity of zero. The segregation is brought about purely by the differences in binocular correlation. This is not a new observation psychophysically but it illustrates a point that is important to remember in interpreting neurophysiological experiments: binocular information may contribute to the segmentation of images without the need for an explicit representation of depth.

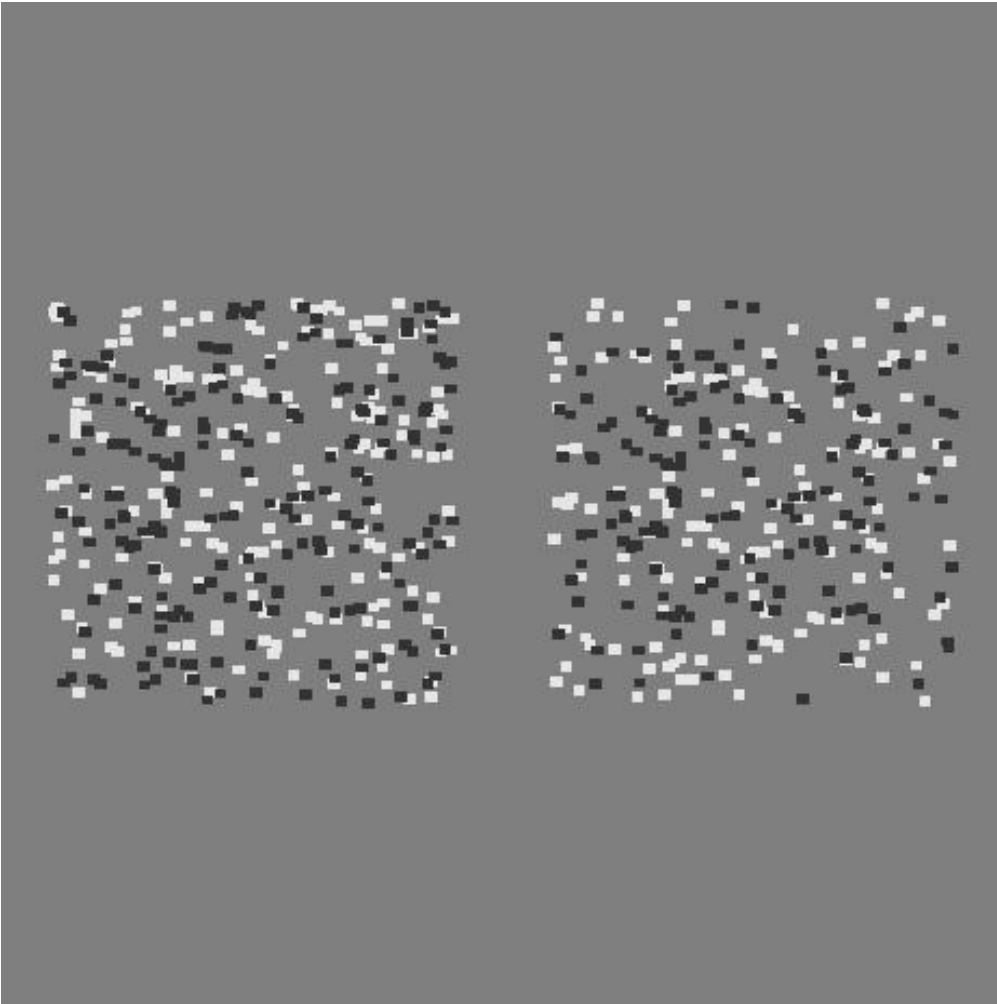


Fig. 2. A binocular pair of random dot pictures. Fusion of the two images results in the segregation of a central region where binocular correlation between right and left images is 100% from a surround region where the eyes' images are uncorrelated, 0%. Note that it does not matter whether these images are viewed with divergent or convergent fusion because the disparity is zero in the centre region and undefined in the surround.

### Depth perception

The more immediate reason why the visual system of frontal-eyed animals has special binocular properties is their ability to extract depth from binocular disparities. It is often mistakenly argued that this is only useful in the near workspace (within 0.5–2 m from the viewer). A simple calculation shows that the best human stereo thresholds imply an ability to distinguish between objects that are truly at infinity and those that are over 400 m away. The general principle can be readily verified for yourself by look-

ing out of a window in a building with a good view into the distance. Place a spot or other marker on the window, fixate the distance and open and close the left and right eyes alternately. The binocular parallax of the spot will be readily visible even when you are several meters from the window. Thus binocular vision is capable of contributing significantly to the perception of scene layout as well as discriminating depth in the near workspace.

The conversion of disparity into depth is also a significant issue. When the movement of the eyes changes the point of binocular regard from one depth

plane to another, the depth signaled by binocular disparity must be re-evaluated. The most obvious point is that there is a change in the distance from the observer at which objects fall on corresponding points on the left and right retinæ. This is implied by binoculation since the foveae themselves are, of course, corresponding points between the left and right retinæ. The necessary readjustment of vergence has long been acknowledged as a potential source of information about distance (Descartes, 1664). Also of significance is the fact that the depth relationship between different disparities must be recalibrated. The same depth difference far away from the observer creates a smaller disparity than the same difference in depth in an object closer to the observer. The source of signals providing this calibration has been studied intensively in recent years (see Mayhew and Longuet-Higgins, 1982; Cumming et al., 1991; Johnston, 1991; Sobel and Collett, 1991; Rogers and Bradshaw, 1993). An elegant recent synthesis is provided by Backus (Backus and Banks, 1999; Backus et al., 1999).

### Cortical area V1

The simple picture of the operation of binocular disparity selective neurons in V1 to be advanced here is that they respond primarily to the local disparity of binocular features presented within their receptive fields. The evidence for this is wide-ranging and considerable, but nonetheless there are views incompatible with this simple picture. Here, we initially concentrate on the positive evidence that has led to our conclusions and then attempt to interpret conflicting views in the light of this evidence.

### Absolute disparity

Before proceeding any further, it is necessary to make more precise our concept of disparity. Until now, we have referred to binocular stimuli as falling on corresponding or non-corresponding retinal points. Whilst a single point in the dark that falls on non-corresponding retinal points is sufficient to allow a response to the depth of the target, it has been clear since the work of Westheimer in the 1970s that the best binocular performance is achieved when there is more than one point visible

within the binocular visual field (Westheimer, 1979). Indeed it appears that the finest stereo judgments are supported by stimulus configurations in which the depth of one feature is judged relative to another.

The diagram in Fig. 3 indicates the distinction that is being made here. In the left-hand diagram, point  $d$  alone is at a non-zero disparity but the threshold for detecting that  $d$  is not at the fixation plane is made considerably lower by the presence of point  $f$ , the visible fixation point (Westheimer, 1979). It is therefore possible to conceive of the binocular disparity of point  $d$  in two different ways. With respect to the point of binocular regard of the eyes, it has a measurable disparity, which we refer to as its 'absolute disparity'. With respect to point  $f$ , it also has a measurable disparity, which we refer to as its 'relative disparity'. In the left hand figure, the value of absolute and relative disparity is the same:  $\alpha + \beta$ . Absolute and relative disparity can be distinguished by considering what happens when the eyes move, leaving  $d$  and  $f$  fixed at their original locations in 3-D space. This is illustrated in the right-hand version of the figure, where it can be appreciated that movement of the eyes changes the absolute disparity of all points in the binocular field, but the relative disparities between pairs of visible points are unaffected. Thus, the relative disparity between  $d$  and  $f$  remains  $\alpha + \beta$ , but the absolute disparity of  $d$  has decreased, due to the eye movement.

In Cumming and Parker (1999), we investigated how neurons in V1 respond to absolute and relative disparity. The strategy was to add a controlled extra amount of absolute disparity to the receptive field. This simulates the sensory consequences of a change in the vergence state. The extra absolute disparity was held at its target value by means of a feedback loop so that further fluctuations of vergence would have no effect on the added absolute disparity. The result was clear. The co-ordinate system for representing disparity in V1 is linked to absolute disparity. Changes of absolute disparity in the stimulus over the receptive field are the dominant factor in generating the disparity tuning curves of these neurons.

The stimuli used in Cumming and Parker (1999) consisted of a patch of dynamic random dots, which covered the binocular receptive field of the V1 neuron at all disparities tested in the experiment. Surrounding this central patch was an annulus of binoc-

?#1

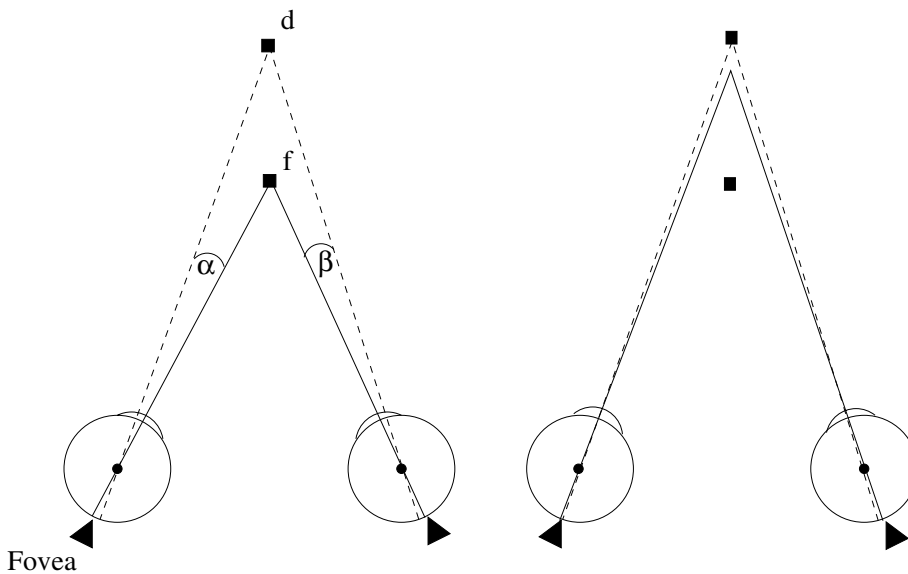


Fig. 3. The distinction between absolute and relative disparities. See text for details.

ular random dots always at zero disparity (same as the fixation point). Thus there is always a relative disparity signal present within the display. It is clear that this relative disparity is perceptually effective because removal of the surround raises the animals' psychophysical thresholds by a factor of 10 (Prince et al., 2000). Nonetheless, the firing of V1 neurons is completely dominated by their response to the absolute disparity of the central region of the display. This suggests that it is necessary to look outside V1, in other cortical visual areas, for signals that are more tightly linked to the perceptual responses of the animal.

We may tentatively assign two functions to the neuronal signals for binocular disparity in V1. First, it is possible to build receptive fields sensitive to relative disparity by combining the signals from several V1 receptive fields. Second, neurons that are sensitive to absolute disparity are potentially useful for controlling changes in the vergence state of the eyes. Therefore the signal from V1 neurons that respond to a target at a non-zero disparity provides a way of estimating the absolute disparity of the target. A signal of this kind would correspond to the sensory control signal that Rashbass and Westheimer (1961) deduced should underlie the generation of vergence eye movements.

There are two recent pieces of evidence that support this proposal of at least a dual function for the disparity-selective neurons in V1. Their potential role in perceptual tasks is emphasized by the fact that they carry highly reliable signals about binocular disparity: the performance of the best neurons approaches that of the psychophysical observer (Prince et al., 2000). The potential role in the control of vergence is highlighted by the fact that there are parallel changes in neuronal signals and fast corrective vergence movements when the contrast of a stimulus is inverted in one eye to form a binocularly anti-correlated stimulus (Cumming and Parker, 1997; Masson et al., 1997).

This comparison between the different stimuli that control vergence and those that control psychophysical judgments of stereoscopic depth makes it clear that the output of V1 may be used in more than one way. In this context, it is significant that earlier investigations established unambiguously that changes of vergence position could be induced by changes in the absolute disparity of random-dot targets, even when the change in disparity did not give rise to a perceived change in stereoscopic depth (Erkelens and Collewijn, 1985). Despite these examples, in general there is relatively little data that directly compare perceptual judgments of binocular depth

and the control of vergence movements by assessing their respective sensitivities to different stimulus manipulations. More comparisons of this type will be valuable in delineating the contributions of neural signals to these distinctive binocular functions.

### **Binocular receptive fields and the energy model**

A highly successful account of the responses of binocular neurons in the primary visual cortex has been provided by the energy model (Ohzawa et al., 1990). The energy model describes the responses of complex cells and it is constructed from the outputs of binocular simple cells that themselves summate the input from left and right retina. The simple cells are generally assumed to have receptive fields that are accurately described by Gabor functions and in the simplest version of the model four such functions are used that differ from each other by  $\pi/2$  radians change of phase. The pooling of signals from simple cells to complex cells passes through a ‘half-squaring’ nonlinearity (a threshold below which the simple cell does not respond and a squaring relationship above the threshold). Since the signals from pairs of simple cells with a  $\pi$  radians difference of phase are added together, this effectively squares the output of this unit-pair. Essentially the same argument applies to the other unit-pair of simple cells, except they are phase-shifted by  $\pi/2$  radians. Therefore, the two unit-pairs form a phase-independent energy calculation.

This model had the primary aim of accounting for data derived from recording in the anesthetized, paralyzed cat, but it has proved to be successful in giving an account of the disparity-specific mechanisms in V1 of awake behaving primates. This progress has been fully reviewed by Cumming and DeAngelis (2001), who provide a detailed account of the nature of the disparity-detecting mechanisms in V1 and their reliance on phase and position disparities.

The fact that the energy model is generally applicable is highly informative about the nature of the stereoscopic computations carried out by cortical area V1. Although the model is backed up by a considerable body of evidence, there are some unresolved discrepancies with existing data from V1. A much more serious issue surrounding the energy model is that it fails to provide a complete account

of the processing of stereoscopic disparities. The ways in which the model fails are characteristic and arise from the fact that the energy computation is essentially local in nature. Thus it will fail to give an account of stereoscopic depth processing that involve global phenomena.

### **Limitations of the disparity energy model**

The limitations of the disparity energy model are of two forms. The first area of concern is whether all aspects of the neuronal response to disparity within the primary visual cortex are properly characterized by the disparity energy model. In these cases, it may be possible to modify the energy model to make it more faithful to the true picture of neuronal signals, without violating the essential characteristics of the way that the model computes stereo disparity. The second area of concern arises from cases where neither the energy model nor the responses of V1 neurons are sufficient to understand the stereoscopic perception of depth. Strictly speaking, examples of this kind do not really represent failures of the energy model since it was developed primarily to give an account of the physiological responses of V1 cortical neurons. Nonetheless, these limitations on the energy model are provocative and powerful, since they force us to look at the nature of stereo processing outside V1 and to consider how the signals from ‘energy-like’ disparity-detectors could be further processed to yield responses that better account for stereo depth perception.

### **Binocularly anti-correlated stimuli**

Binocularly anti-correlated stimuli are created by taking a conventional binocular stimulus and inverting the contrast of the stimulus in one eye with respect to the other. Thus in an anti-correlated random dot stereogram, each black dot in one eye’s image is partnered with a white dot in the other eye’s image and so on (Julesz, 1971). The response of the disparity energy model to this manipulation is an inversion of the disparity-tuning curve. Thus if a Gabor function is used to describe the disparity-tuning curve, then the energy model predicts that the disparity-tuning curve for binocularly correlated stimuli and anti-correlated stimuli should have the

same amplitude of response but a complete inversion of phase with respect to one another (Cumming and Parker, 1997).

Experimentally, for both dynamic random-dot stereograms in monkey V1 (Cumming and Parker, 1997) and for the responses of cat V1 neurons calculated by reverse correlation with rapidly presented bar stimuli (Ohzawa et al., 1997), the result is the same. The amplitude of the response to binocularly anti-correlated stimuli is weaker than expected and the range of phases is spread broadly, rather than tightly, around the expected phase inversion of  $\pi$  radians. This discrepancy cannot be repaired by something as simple as an output non-linearity in the relationship between neuronal firing rate and membrane potential within cortical cells. This would leave the range of phase relationships still tightly clustered around  $\pi$  radians. Moreover, an output non-linearity predicts that cortical neurons that show a tuned-inhibitory response to binocularly correlated stimuli should show amplitude ratios *greater* than one in response to binocularly anti-correlated stimuli. It also predicts that neurons with odd-symmetric disparity tuning profiles (so-called ‘near’ and ‘far’ cells) should always show amplitude ratios of exactly one. None of these various predictions is fulfilled.

Recently in collaboration with Jenny Read, we have explored the properties of modified energy models that include a monocular non-linear stage prior to binocular combination (Read, Cumming and Parker, in preparation). The monocular stages feed into a binocular simple cell prior to the energy computation on the outputs of multiple simple cells to form a binocular complex neuron. Since the monocular inputs to the binocular simple cell are matched in their response to contrast polarity, the monocular non-linearity has the effect of reducing or eliminating the opportunity for binocular summation when the contrast is inverted in one eye. Models of this type can successfully predict the range of behaviors found in neuronal recordings. However, there is some penalty for these additional stages of processing. The energy model loses some of its elegance and simplicity and it is necessary to add more subunits that the minimum quadrature pairs to maintain the prediction of smooth receptive field profiles for cortical complex neurons. On the whole, the outcome is

a successful modification of the energy model that explains the response to binocular anti-correlation without discarding the local nature of the energy computation (see Ohzawa, 1998, for an alternative view of the effects of anti-correlation).

### Responses to local binocular stimulation

A more serious limitation of the energy model is illustrated by considering the response of the model to windowed sinusoidal grating stimuli (see Fig. 4). Stimuli of this general type provide an interesting test case for the relationship between the activity of single neurons and perception. As the disparity of the windowed sinusoid is altered, the perception of depth by psychophysical observers is controlled by the disparity of the entire display: thus perceived depth varies linearly with binocular disparity. The situation is rather different for a mechanism with a receptive field of limited spatial extent (as shown by the rectangle superimposed on the grating in Fig. 4). A mechanism of this kind receives a local signal about binocular disparity that is essentially identical for all added disparities that are related by a single period of the grating pattern. Note that this is true even when disparity of the grating pattern is exactly linked to that of the window.

The problem created by this stimulus for a mechanism with limited spatial extent is an example of the aperture problem, which is more typically identified with the processing of motion information (Marr and Ullman, 1981; Wallach: see Würger et al., 1996) but occurs for stereo processing as well (Morgan and Castet, 1997). In order to solve this problem, information about the depth signaled by the sharp boundary of the window needs to be put together with the local contour information given by the gratings. This allows the inherent ambiguity of the grating to be resolved. Perceptually, if there is no window boundary in place, then a spatially extended periodic stimulus is indeed ambiguous and creates the so-called ‘wallpaper’ effect, in which multiple depth matches are experienced at different times during the viewing of these stimuli (Smith, 1738).

Our recent recordings in cortical area V1 (Cumming and Parker, 2000) show that information from the window boundary fails to influence the firing patterns of V1 neurons. The responses of these neurons

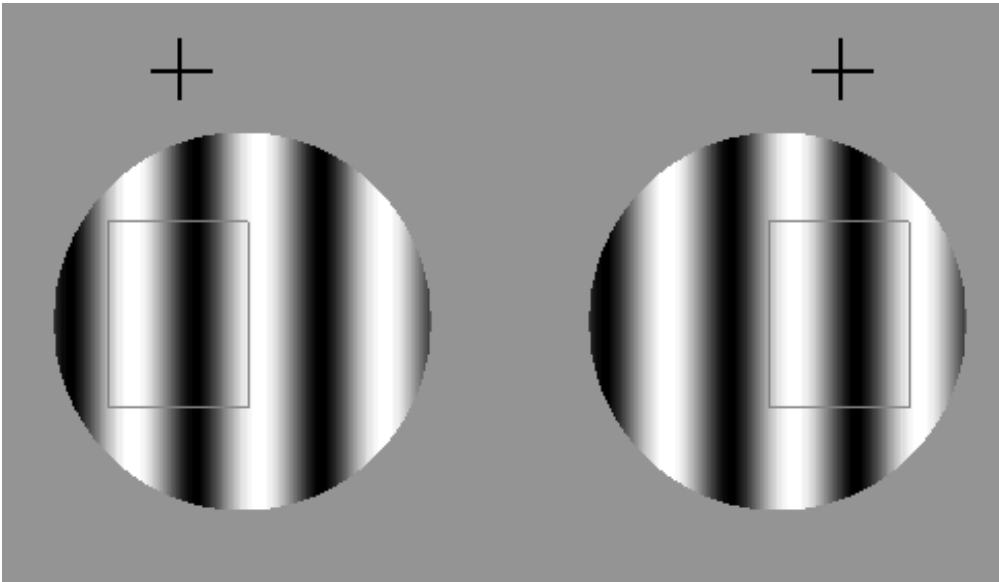


Fig. 4. Windowed sinusoidal grating stimuli used to test whether cortical neurons respond to local disparity or perceived depth. When binocularly fused (without the rectangular box present), the perception is of a single disk of grating at a different depth plane than the fixation point (crosses). A neuron with the receptive field shown by the rectangular box that responds only to local disparity continues to receive stimulation at zero disparity. (The rectangular box over the grating is present for illustrative purposes only and unsurprisingly has an additional effect on the binocular matching process, when attempting to fuse the images in this picture.)

are completely dominated by the visual stimulation created by the local contours within the receptive field. No consistent modulation of V1 responses were linked to the depth percept experienced by the psychophysical observers (either monkeys or humans) in response to these stimuli.

These observations for stereo bear a close analogy to earlier observations (Movshon et al., 1985) on the responses of V1 neurons to moving stimuli that are constructed to contain two component orientations (plaids). The main response of V1 neurons to plaids is dominated by the response of the orientation-selective classical receptive field, which is only sensitive to a component of the plaid that falls within the orientation bandwidth of the receptive field. Consequently, as the orientation and direction of the moving plaid is changed through  $360^\circ$ , the response of a strongly direction-sensitive V1 neuron is bimodal since it responds independently to each component orientation that forms the plaid. Neither of these responses occurs when the plaid is perceived to be moving in a direction similar to the direction-preference of the neuron. Thus the representation of the true direction of motion of the plaid stimulus in

the firing patterns of V1 neurons is ambiguous, in much the way that the binocular depth of the windowed grating patch is ambiguously related to the firing patterns of disparity-selective neurons in V1.

This comparison of direction selectivity and disparity selectivity in V1 is made more compelling by the parallel between the responses of computational models of directional and disparity-sensitive detectors. A direction-sensitive motion detector based on the 'motion energy' principle (Adelson and Bergen, 1985) with orientation selective spatial filters has the same ambiguity in response to plaid stimuli as the bulk of V1 neurons. A disparity-selective detector based on the energy principle suffers the same ambiguity as V1 neurons in response to the windowed sinusoidal grating pattern. In the case of motion, there is strong evidence (Movshon et al., 1985) that the ambiguity of response to plaid stimuli is eliminated in the neuronal firing patterns within extrastriate visual cortex (area MT/V5). If this parallel between motion and disparity is correct, then it suggests that the key to disambiguating the windowed-grating stimulus may also be discovered in the extrastriate cortex.

## Fidelity of signals

Before turning to a discussion of the extrastriate visual areas, we should attempt to answer more precisely one obvious question: namely, given that the disparity-selective responses of V1 neurons exhibit several discrepancies with the perceptual response to binocular depth, what is the contribution of V1 neurons to the binocular perception of depth? The broad picture is a combination of information derived from studies on binocular responses in anesthetized, paralyzed preparations (where the opportunities for long-duration experiments on single neurons is maximized) and in awake, behaving monkeys (where the opportunity for parallel psychophysical experimentation arises and a natural binocular alignment on the stimuli can be achieved).

The recent years have seen an encouraging convergence of information from these two types of experiment. It now seems clear that V1 neurons respond primarily to absolute disparity of local contours within their receptive fields (Cumming and Parker, 1999, 2000) and it seems likely that some variant of the energy model (Ohzawa et al., 1990) will prove to be a sufficiently accurate description of most features of the neuronal response in V1 to binocular stimuli. The responsiveness to absolute disparity means that V1 neurons provide a sensory signal that is well suited for controlling the disparity-driven components of the vergence eye movement response. Indeed, a signal based on relative disparity would be more closely related to the perceptual response to binocular disparity but would be less suitable for controlling vergence.

In cortical area V1, Prince, Cumming and Parker (manuscript submitted for publication) report that some 40–55% of neurons are disparity selective, depending on the criteria for selectivity that are applied. It is hard to believe that this significant number of disparity-selective neurons is entirely devoted to controlling vergence position. It seems likely that these V1 neurons are also precursors to other binocular neurons whose responses are more closely linked to binocular depth perception. One new line of evidence that is consistent with this is the observation that individual V1 neurons are highly sensitive in signaling binocular disparity. The performance of the best V1 neurons is comparable with the psy-

chophysical performance of the monkeys themselves (Prince et al., 2000). This suggests that the V1 neurons are part of a cortical network that is devoted to the efficient and accurate registration of binocular image differences in preparation for further processing elsewhere, whether it is for vergence eye position or the perception of binocular depth.

## The role of extrastriate visual areas

Our knowledge of the responses of extrastriate cortex to binocular disparity is considerably less precise than the statements that we can make about V1. The relevant evidence is patchy and it is entirely conceivable that important and significant features have been completely overlooked. One issue is clear and consistently reported: the overwhelming proportion of visual neurons in extrastriate cortical areas are binocularly driven. Thus the potential for binocular disparity to influence the responses of these neurons is all-pervasive. Moreover, since the incidence of monocular neurons is low in extra-striate cortex, this suggests that the bulk of disparity selectivity observed in these areas derives from the disparity-selective neurons within the striate cortex.

This section will assess briefly some recent advances in three areas: the processing of depth and contour information in cortical area V2; the evidence for a clustered representation of disparities in cortical area V5/MT and the effects of electrical microstimulation within these clusters; the picture of cortical activity provided by recent functional magnetic resonance imaging (fMRI) studies of binocular depth perception.

## The diversity of V2

By the standards of visual neuroscience, the study of the processing of binocular disparity in V2 has a long history (Hubel and Wiesel, 1970). Neurons that change their responses systematically with the disparity of isolated bar contours have been associated particularly with the thick cytochrome oxidase stripes of V2 (Hubel and Livingstone, 1987; Peterhans and von der Heydt, 1993). If we consider the responses of V2 neurons in a wider context, it is clear that there are multiple opportunities for interactions between binocular disparity and other visual cues.

For example, there are numerous indications that the processing of texture boundaries, subjective contours, line terminators etc. within V2 is in some way more emphasized than within V1 (Peterhans and von der Heydt, 1993). The mechanisms by which sensitivity to these parameters is achieved have not been clearly elucidated.

It is important to consider the relationship of stereo disparity to these elements that seem to be primarily sensitive to perceptual segregation based on texture boundaries and line terminators. If these mechanisms are truly concerned with perceptual segregation, then one might expect that they will also be sensitive to regional segregation based on binocular disparity (Von der Heydt et al., 2000). This highlights the point that binocular correspondence is used in more than one way by the visual system: it is used for the recovery of depth and distance information, but it is also used for breaking camouflage and achieving segmentation of the visual scene (as illustrated by Fig. 2). Thus there are at least a priori reasons to expect a variety of ways in which disparity might influence the firing patterns of V2 neurons. Not only might there be neurons that change their response systematically with changes of disparity but there might also be neurons whose responses are modulated whenever there is a boundary defined by disparity. Bakin et al. (2000) have recently described modulatory effects of disparity on V2 neurons, using a variety of complex stimulus configurations.

In recent work with Owen Thomas, we have explored how the responses of V2 cortical neurons are modulated by the presence of disparities presented outside the classical receptive field. The stimuli were dynamic random-dot stereograms, arranged so that the depth of a circular central region could be altered independently of the depth of an annular surround. The central region extended just outside the classical receptive field as defined by plotting with bars and edges. In a number of neurons, alteration of the disparity of the surround profoundly affects the responses of the center to the disparity of the stimulus. In these neurons, changing the disparity of the surround alters the disparity preference of the central region. The change is such that the neuron tends to exhibit a consistent preference for the relative disparity between the center and the surround

(Thomas et al., 1999; Thomas, Cumming and Parker, in preparation).

In this regard at least, processing of binocular disparity in V2 appears to move closer to neuronal responses that match the perceptual characteristics of binocular depth judgments. In terms of areal dimensions, V2 is obviously a highly significant locus within the cortical visual pathways. For this reason, if no other, it is to be expected that other cortical areas of significant size (such as V4) will also have an important role in regard to binocular depth judgments. We will now turn to other sources of evidence about the role of other extrastriate cortical areas in the processing of binocular depth.

### **Microstimulation in V5(MT)**

Although cortical area V5(MT) is classically associated with the processing of motion information, it has been known for some time that the neurons there are sensitive to binocular disparity (Maunsell and van Essen, 1983). A recent series of studies has revealed a functional organization for disparity within V5(MT). The suggestion is that this has a columnar form, rather like the columnar arrangements in V5(MT) already identified for motion direction (DeAngelis and Newsome, 1999). There are distinct cortical regions within V5(MT) that show common disparity preferences (typically such that a neighborhood of cortex shows a consistent preference for either near or far disparities).

In recent experiments, DeAngelis et al. (1998) exploited this organization to set up a test of whether electrical microstimulation in area V5(MT) might influence judgments of binocular depth, in the same way that earlier studies in V5(MT) had established that electrical microstimulation influences judgments of motion direction (Salzman et al., 1990). The task of the monkey was to indicate whether a patch of dots was at a near or far distance with respect to the fixation plane. The proportion of dots signaling near or far was altered by diluting the dots present with some dots at randomly placed disparities. Thus the task is close in form to the original motion task in which the direction of motion is signaled by the fraction of dots that are moving consistently in one direction against a background of dots that are moving randomly in all directions. Electrical mi-

crostimulation in zones of V5(MT) whose neurons preferred near disparities tended to induce the monkey to report more often that the target was near with respect to the fixation plane, with the opposite effect for far disparities. Importantly, the effectiveness of microstimulation was greatest in zones of V5(MT) with a strong rather than weak selectivity for disparity.

### Functional magnetic resonance imaging

Recently, functional magnetic resonance imaging has been used in human observers to explore the visual cortical regions that are involved in stereoscopic depth perception. Unsurprisingly, these studies have revealed that there is a widespread activation of visual areas by stimuli whose disparity or degree of interocular correlation is changed systematically. Using a random-dot stereogram stimulus with a central square at a different depth plane, Mendola et al. (1999) reported activation in cortical area V3A, which could be due to activation of either disparity-specific mechanisms or contour-specific mechanisms responsive to the square. In experiments in which a binocular random element stimulus switches from depicting a single plane at the fixation depth to a pair of planes symmetrically disposed either side of the fixation plane, Backus, Fleet, Parker and Heeger (manuscript submitted for publication) demonstrate that cortical area V3A is especially sensitive to this manipulation, suggesting a disparity-specific response. It remains to be seen whether other stimulus paradigms will reveal the involvement of other visual areas: notably, in view of the well-characterized disparity sensitivity of V5(MT) neurons, the weakness of evidence for specific activations of this area presents a confusing picture at present.

### Conclusions

This is an interesting moment for the study of stereoscopic vision at the cortical level. A number of new paradigms have been added to the neurophysiologists' repertoire. Until recently, the emphasis has been on the pure measurement of disparity selectivity and the characterization of models of how binocular neurons are formed in V1. Such studies have generated fundamental information, but the emphasis is

now shifting towards the discovery of cortical sites whose neuronal properties could directly support some of the well-characterized properties of the perception of stereoscopic depth. We are as yet unclear which sites in the extra-striate cortex will eventually prove to be primarily responsible. Nonetheless it is clear that, at the neuronal level, there are interesting differences between cortical areas V2 and V5(MT) in comparison with the striate cortex V1 (see Cumming and DeAngelis, 2001, for a summary). The case of V3A is interesting, since although it has been known for some time (Poggio et al., 1988) that areas V3/V3A in the macaque contain disparity-specific neurons, it is recent human fMRI data that have emphasized its potential specialization. Perhaps this will be a clear case where fMRI data have been able to guide a neurophysiological investigation, rather than the other way round. The next 5–10 years should provide a wealth of new results that will address these questions.

### Acknowledgements

This work was supported by the Wellcome Trust. B.G.C. was supported by a Royal Society University Research Fellow.

### References

- Adelson, E.H. and Bergen, J.R. (1985) Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A Optics Image Sci. Vis.*, 2: 284–299.
- Backus, B.T. and Banks, M.S. (1999) Estimator reliability and distance scaling in stereoscopic slant perception. *Perception*, 28: 217–242.
- Backus, B.T., Banks, M.S., van Ee, R. and Crowell, J.A. (1999) Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision Res.*, 39: 1143–1170.
- Bakin, J.S., Nakayama, K. and Gilbert, C.D. (2000) Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *J. Neurosci.*, 20: 8188–8198.
- Cumming, B.G. and DeAngelis, G.C. (2001) The physiology of stereopsis. *Annu. Rev. Neuroscience*, 24: 203–238.
- Cumming, B.G. and Parker, A.J. (1997) Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature*, 389: 280–283.
- Cumming, B.G. and Parker, A.J. (1999) Binocular neurons in V1 of awake monkeys are selective for absolute, not relative, disparity. *J. Neurosci.*, 19: 5602–5618.
- Cumming, B.G. and Parker, A.J. (2000) Local disparity not

- perceived depth is signaled by binocular neurons in cortical area V1 of the macaque. *J. Neurosci.*, 20: 4758–4767.
- Cumming, B.G., Johnston, E.B. and Parker, A.J. (1991) Vertical disparities and perception of 3-dimensional shape. *Nature*, 349: 411–413.
- DeAngelis, G.C. and Newsome, W.T. (1999) Organization of disparity-selective neurons in macaque area MT. *J. Neurosci.*, 19: 1398–1415.
- DeAngelis, G.C., Cumming, B.G. and Newsome, W.T. (1998) Cortical area MT and the perception of stereoscopic depth. *Nature*, 394: 677–680.
- Descartes, R. (1664) *Tractatus de homine*. Paris.
- Erkelens, C.J. and Collewijn, H. (1985) Motion perception during dichoptic viewing of moving random-dot stereograms. *Vision Res.*, 25: 583–588.
- Hubel, D.H. and Livingstone, M.S. (1987) Segregation of form, color, and stereopsis in primate area 18. *J. Neurosci.*, 7: 3378–3415.
- Hubel, D.H. and Wiesel, T.N. (1970) Cells sensitive to binocular depth in area 18 of the macaque monkey cortex. *Nature*, 225: 41–42.
- Johnston, E.B. (1991) Systematic distortions of shape from stereopsis. *Vision Res.*, 31: 1351–1360.
- Julesz, B. (1964) Binocular depth perception without familiarity cues. *Science*, 145: 356–362.
- Julesz, B. (1971) *Foundations of Cyclopean Perception*. University of Chicago Press, Chicago.
- Marr, D. and Ullman, S. (1981) Directional selectivity and its use in early visual processing. *Proc. R. Soc. Lond. B*, 211: 151–180.
- Masson, G.S., Busetini, C. and Miles, F.A. (1997) Vergence eye movements in response to binocular disparity without depth perception. *Nature*, 389: 283–286.
- Maunsell, J.H.R. and van Essen, D.C. (1983) Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *J. Neurophysiol.*, 49: 1148–1167.
- Mayhew, J.E.W. and Longuet-Higgins, H.C. (1982) A computational model of binocular depth-perception. *Nature*, 297: 376–378.
- Morgan, M.J. and Castet, E. (1997) The aperture problem in stereopsis. *Vision Res.*, 37: 2737–2744.
- Movshon, J.A., Adelson, E.H., Gizzi, M.S. and Newsome, W.T. (1985) The analysis of moving visual patterns. *Pont. Acad. Sci. Scripta Varia*, 54: 118–151.
- Mendola, J.D., Dale, A.M., Fischl, B., Liu, A.K. and Tootell, R.B. (1999) The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *J. Neurosci.*, 19: 8560–8572.
- Ohzawa, I. (1998) Mechanisms of stereoscopic vision: the disparity energy model. *Curr. Opin. Neurobiol.*, 8: 509–515.
- Ohzawa, I., DeAngelis, G.C. and Freeman, R.D. (1990) Stereoscopic depth discrimination in the visual-cortex — neurons ideally suited as disparity detectors. *Science*, 249: 1037–1041.
- Ohzawa, I., DeAngelis, G.C. and Freeman, R.D. (1997) The neural coding of stereoscopic depth. *NeuroReport*, 8: R3–R12.
- Peterhans, E. and von der Heydt, R. (1993) Functional organization of area V2 in the alert macaque. *Eur. J. Neurosci.*, 5: 509–524.
- Poggio, G.F., Gonzales, F. and Krause, F. (1988) Stereoscopic mechanisms in monkey visual cortex: binocular correlation and disparity selectivity. *J. Neurosci.*, 8: 4351.
- Prince, S.J.D., Pointon, A.D., Cumming, B.G. and Parker, A.J. (2000) The precision of single neuron responses in cortical area V1 during stereoscopic depth judgements. *J. Neurosci.*, 20: 3387–3400.
- Rashbass, C. and Westheimer, G. (1961) Disjunctive eye movements. *J. Physiol.*, 159: 339–360.
- Rogers, B.J. and Bradshaw, M.F. (1993) Vertical disparities, differential perspective and binocular stereopsis. *Nature*, 361: 253–255.
- Salzman, C.D., Britten, K.H. and Newsome, W.T. (1990) Cortical microstimulation influences perceptual judgments of motion direction. *Nature*, 346: 174–177.
- Smith, R. (1738) *A Complete System of Optics in Four Books*. Cambridge.
- Sobel, E.C. and Collett, T.S. (1991) Does vertical disparity scale the perception of stereoscopic depth? *Proc. R. Soc. Lond. B*, 244: 87–90.
- Thomas, O.M., Cumming, B.G. and Parker, A.J. (1999) The responses of binocular V2 neurons to relative and absolute disparities. *Soc. Neurosci. Abstr.*, 25: 821.7.
- Von der Heydt, R., Zhou, H. and Friedman, H.S. (2000) Representation of stereoscopic edges in monkey visual cortex. *Vision Res.*, 40: 1955–1967.
- Westheimer, G. (1979) Cooperative neural processes involved in stereoscopic acuity. *Exp. Brain Res.*, 36: 585–597.
- Wheatstone, C. (1838) Contributions to the physiology of vision. I. On some remarkable and hitherto unobserved phenomena of binocular vision. *Phil. Trans. R. Soc.*, 128: 371–394.
- Wheatstone, C. (1852) Contributions to the physiology of vision. II. On some remarkable and hitherto unobserved phenomena of binocular vision. *Phil. Trans. R. Soc.*, 142: 1–17.
- Würger, S., Shapley, R. and Rubin, N. (1996) 'On the visually perceived direction of motion' by Hans Wallach: 60 years later. *Perception*, 25: 1317–1368.

QUERIES:

?#1: Please note that Fig. 4 has become Fig. 3 and vice versa (page 4)