# 5

# The role of disparity interactions in perception of the 3D environment

CHRISTOPHER W. TYLER

## 5.1 Introduction

In understanding visual processing, it is important to establish not only the local response properties for elements in the visual field, but also the scope of neural interactions when two or more elements are present at different locations in the field. Since the original report by Polat and Sagi (1993), the presence of interactions in the two-dimensional (2D) field has become well established by threshold measures (Polat and Tyler, 1999; Chen and Tyler, 1999, 2001, 2008; Levi *et al.*, 2002). A large array of other studies have also looked at such interactions with suprathreshold paradigms (e.g., Field *et al.*, 1993; Hess *et al.*, 2003). The basic story from both kinds of studies is that there are facilitatory effects between oriented elements that are collinear with an oriented test target and inhibitory effects elsewhere in the 2D spatial domain of interaction (although the detectability of a contrast increment on a Gabor pedestal also reveals strong collinear masking effects).

The present work extends this question to the third dimension of visual space as specified by binocular disparity, asking both what interactions are present through the disparity dimension and how these interactions vary with the spatial location of the disparate targets. Answering these questions is basic to the understanding of the visual processing of the 3D environment in which we find ourselves. *Pace* Plato, the world is not a 2D screen onto which images are projected, but has the full extra dimensionality of a 3D space through which we have to navigate and manipulate the objects and substances of physical reality.

*Vision in 3D Environments*, ed. L. R. Harris and M. R. M. Jenkin. Published by Cambridge University Press.

It is worth emphasizing that the addition of the third, $z$, axis to the two axes of $x,y$ space does not just add another plane to the initial $x,y$ plane, but extends its dimensionality in a multiplicative fashion to vastly expand the scope of possible interactions. Thus, whereas there are two dimensions of pointwise interactions for a single dimension of $n$ points (i.e., the effect on the detectability of each $x$ point of stimuli at all other $x$ points), there are four dimensions of possible interactions in 2D space $(x,y \times x,y)$ and six dimensions of possible interactions in 3D space $(x,y,z \times x,y,z)$ that need to be evaluated in order to have a complete understanding of the visual processing of that space. Moreover, we cannot assume that visual space is homogeneous in its properties, as is classically assumed for physical space. The space of interactions is itself keyed to the absolute location of the targets in the space, adding an addition dimension of possible variations.

The question of disparity specificity was addressed in reverse with a suprathreshold paradigm by Hess *et al.* (1997), who measured the secondary effect of relative disparity on the spatial interactions among targets at different spatial locations. Here, the issue is the direct question of the disparity-domain interactions between elements at different disparities and the same spatial location. The issue then generalizes to the six-dimensional space of the disparity, the $x$, $y$ location, the orientation, the size, and the luminance polarity of a pair of targets. In each case, the interaction between two elements forms one relevant dimension, with the variation of this interaction over the range of absolute values as a second dimension. The total parameter space for pairwise interactions in 3D space is thus 12-dimensional. To keep the problem manageable, the enquiry is restricted to the horizontal ($x$) direction of space, focusing on the dimensions of absolute and relative disparity, absolute size, and relative $x$ location and relative polarity.

A note on the concept of absolute and relative disparity: one should distinguish between the concepts of "binocular disparity" and "convergence angle." Thus, zero disparity for the absolute-disparity ($z$) reference frame is assumed to be the angle between the fixation line when the eyes are converged on the fixation point. (Zero convergence angle is given by convergence on optical infinity, but its use as the zero for absolute disparity would require the awkward concept that targets with negative, or uncrossed, disparities relative to this convergence angle would have to be seen as more than an infinite distance away.) The test targets are set laterally from the fixation target either at the *same* (zero) disparity or at a range of different (absolute) disparities. The concept of *relative* disparity ($\triangle z$) comes into play when there are two targets in the visual field (other than the fixation point). The relative disparity is the difference in disparity between the absolute disparities of these targets. Also, for simplicity, it is assumed that

the horopter of zero absolute disparity through space is a horizontal line in the frontoparallel plane extending from the fixation point, which is a reasonable approximation within a few degrees of the fixation point (Ogle, 1950; Tyler, 1991).

A natural approach to the study of interactions is the masking paradigm, in which a salient masking stimulus is used to reduce the visibility of a second test stimulus. The presence of masking always implies a processing nonlinearity, because a linear system simply adds the mask to the test, without affecting the response to the test or its signal-to-noise ratio.

There are two conceptual frameworks in which masking may be interpreted. Under the assumption of univariance, the mask and test are assumed to impact the same coding channel in an indistinguishable fashion, exploiting some processing nonlinearity to induce a variation in sensitivity to the test stimulus in the presence of the mask (Mansfield and Parker, 1993). In a complex system, however, the mask and test may be processed by different channels with inhibitory interactions between them, as in the classic case of metacontrast masking (Foley, 1994). Such an inhibitory-masking interpretation does not imply univariance or that the data measure the sensitivity of the processing channels for either stimulus; instead, the data measure the inhibitory relations between them. Distinguishing between the univariant-channel masking and the inhibitory-interaction interpretation is not possible without additional information about the characteristics of the system. However, one strong criterion that can be employed is that univariant channel masking may generally be expected to decrease with distance between the test and mask along any stimulus dimension (such as position, disparity, or spatial frequency). If masking increases with distance, this implies an implausible channel structure, and the more likely interpretation is in terms of inhibitory interactions, as in the present treatment.

The primary stage of stereoscopic processing may be considered as a local cross-correlation process, performed by neurons tuned to different disparities, occurring at each location in the binocular visual field (Stevenson *et al.*, 1992). The best match or correlation in each local region of the visual field specifies when the local images are in register (Figure 5.1). As the eyes vary their vergence, the cortical projections of the visual scene slide over one another to vary their projected shifts (or disparities), which may be termed the Keplerian array of the binocular projection field. The 2D spatial structure of these local elements is elongated vertically for human stereopsis (Chen and Tyler, 2006). In practice there are, of course, both vertical and horizontal dimensions of field location, and the physiological array may not be as regular as in this idealized depiction.

## 5.2 Global interactions

Beyond the disparity registration stage described up to this point are the global interactions operating between the local-disparity nodes, which serve to refine the representation of the disparity image from its initial crude array of stimulated disparities to a coherent representation of the 3D surfaces present in the field of view (Tyler, 2005). A variety of such cooperative processes has been proposed by Julesz and others over the years, summarized in Julesz (1971, 1978) and Tyler (1983, 1991). There are two obvious types of mutual interaction between the disparity-selective signals for different regions of the stereo image: disparity-specific inhibition (along the vertical line in Figure 5.1) and disparity-specific pooling or facilitation (along the horizontal line in Figure 5.1). Each type may, in principle, operate anisotropically to different extents in the three-dimensional space of relationships through the Keplerian array: over the frontal plane, over disparities, or over some combination of the two. Such cooperativity among local-disparity mechanisms may be involved in solving the correspondence problem effectively (Tyler, 1975); it may also include such

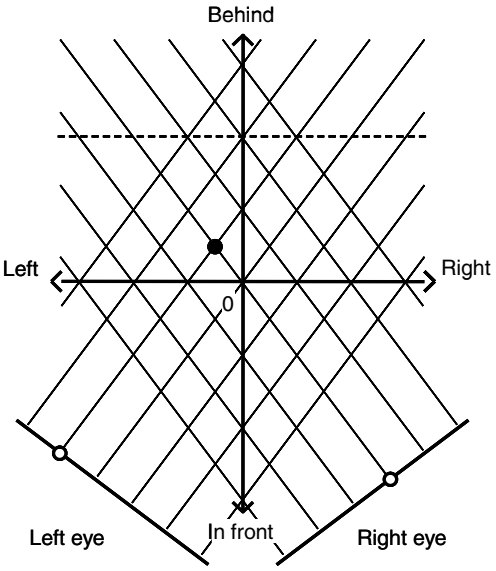


**Figure 5.1** A Keplerian array of disparity detectors depicted by the intersections of the (oblique) lines of sight of the two eyes (see Kepler, 1611). The retinas of the left and right eyes are schematized as linear arrays, with one point in each array (open circles) indicated as the recipient of the image of the stimulus point in space (filled circle). The solid horizontal line depicts the horopter ($x$ axis) and the solid vertical line the $z$ axis; the dashed line indicates a plane at constant uncrossed disparity.

processes as the disparity gradient limitation on the upper limit for depth reconstruction (Tyler, 1973), and coarse-to-fine matching processes for building up the depth image from the monocular information (Marr, 1982). These processes all may be conceived as taking place within the locus of global interactions following the interocular matching or disparity registration stage (but preceding the generation of a unified global depth image from the plethora of available disparity information (Tyler, 1983, 1991)). The present work was designed to provide an initial survey of the scope of such disparity-domain interactions as a function of location, testing disparity, spatial frequency, and luminance polarity.

### 5.3 Local target structure

To maximize the local specificity of the probe, the stimuli in the experiments described here were simultaneously local with respect to four variables: eccentricity, extent, disparity, and spatial frequency. Previous studies have confounded these variables so that it was not possible to disambiguate the effects of retinal inhomogeneity from eccentricity, relative peak positions of the test and mask, and spatial-frequency effects. To avoid such phase artifacts and to maximize the masking effects, we adopted a local-stimulus paradigm based simply on Gaussian bar test stimuli (cf. Kulikowski and King-Smith, 1973). Such Gaussian bars allow measurement by contrast masking of our three requisite variables, namely position sensitivity, disparity sensitivity, and spatial-frequency (scale) selectivity (Kontsevich and Tyler, 2004). Gaussians have only one peak, so that the position and disparity of the masking bar can be varied cleanly relative to the peak of the test bar. The lack of side lobes in Gaussian bars makes them particularly suitable for the study of stereoscopic disparity tuning by a masking paradigm because there is no aliasing of the disparity signal by spurious peak coincidences (as there would be with narrowband wavelet stimuli, for example). The use of such Gaussian bars in peripheral vision allows both disparity and position to be varied within a homogeneous retinal region. Gaussian bars also provide a substantial degree of tuning in spatial frequency (Kontsevich and Tyler, 2004, 2005). If the Gaussian is smaller than the receptive-field center, it provides less than optimal activation; if larger, it stimulates the inhibitory surround, tending to reduce the response. Gaussians thus have an optimal size tuning for center-surround receptive fields that translates into an effective peak spatial frequency.

### 5.4 Psychophysical procedure

The stimuli were Gaussian blobs to measure local position tuning, disparity tuning, spatial-frequency tuning, and polarity tuning at a location 5°

to the left of the fovea by the masking-threshold paradigm, which gives a direct measure of the channel tunings underlying the masking behavior. The methods we described in detail in Kontsevich and Tyler (2005). The psychophysical task was to set the test contrast at some multiple (2–4) of threshold and to vary the masking contrast by the Φ-adaptive staircase procedure (Kontsevich and Tyler, 1999) to the level that brought the test contrast back to its unmasked level of detectability. The advantage of this procedure is that it restricts the number of channels stimulated to those most sensitive to the test stimulus, ensuring that the results are focused on the most basic structure of the disparity-processing mechanisms. To take a maximally model-free approach to the question of local channel interactions through the Keplerian array, the masking-sensitivity paradigm, introduced by Stiles (1939) as the field sensitivity paradigm, was employed. The test stimulus was set at a fixed contrast level above its contrast detection threshold, and the mask contrast required to return the test to threshold was determined as a function of the tuning parameter (position, disparity, or scale). The result was a tuning curve through some aspect of the Keplerian array that had three advantages over threshold elevation curves:

(1) Output nonlinearities of the masking behavior do not affect the shape of the masking function, because the test contrast is always at the same level above the detection threshold.
(2) The effect of the test on neighboring channels and their potential interactions is minimized because the test contrast is near threshold.
(3) If the channel structure is sufficiently discrete that the near-threshold test is invisible to neighboring channels, the mask probes the shape of the inhibitory interactions all the way down its flanks to the maximum extent of masking.

## 5.5 Position tuning

The first experiment was a control study to determine the position tuning of the local contrast-masking effect with respect to the mask phase. With both test and mask of width 25 arcmin, the mask contrast required to reduce the test back to threshold was maximal, implying that the masking *sensitivity* was minimal beyond about 30 arcmin separation from the test (Figures 5.2a–c). It is noteworthy that this masking function does not follow the form of the overlap between the test and mask stimuli (which would be a Gaussian of $\sqrt{2} \times 25 = 35$ arcmin width at half height, as shown by the dashed line in Figure 5.2d). Instead, it has an idiosyncratic form that evidently reflects the tuning of the underlying spatial channel (under the channel invariance assumption laid out in the

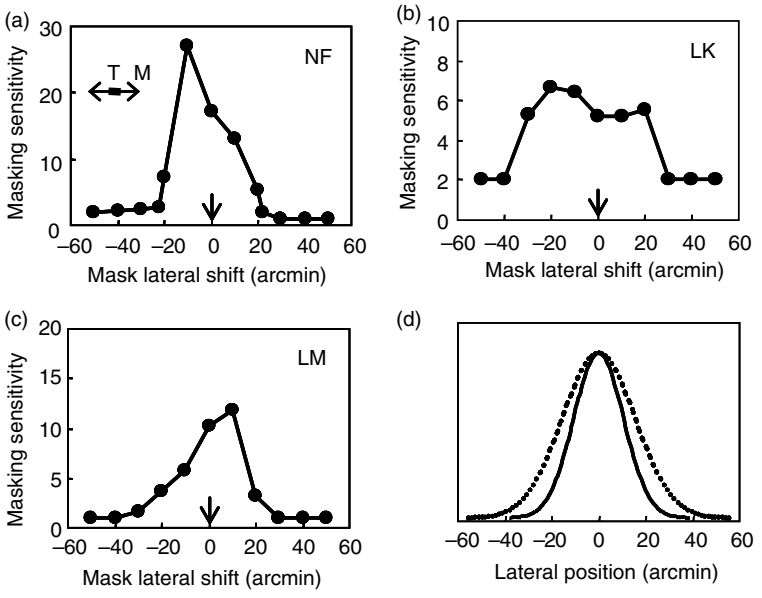


**Figure 5.2** Contrast masking as a function of lateral shift of the masking bar (M) from the test bar (T), both of which had profiles 25 arcmin wide at half height as shown by the solid line in (d). (a)–(c): Masking sensitivity with both mask and test at zero disparity (icon). Arrows at bottom depict test position. The mean standard errors were only ±0.04–0.09 log units, i.e., about the size of the symbols. (d) The Gaussian profile of the test and the mask (solid line) and the expected test/mask interaction profile (dashed line).

Section 5.1) rather than involuntary eye movements or retinal inhomogeneity, which should be consistent across observers.

The first disparity-related condition was used in an experiment to measure the degree of masking as a function of the lateral position of this disparate mask (dashed horizontal line in Figure 5.1). The mask disparity was set at 80 arcmin (40 arcmin displacement in each eye) for NF and 60 arcmin for LK. The masking effects (Figure 5.3) reveal the presence of long-range interactions across disparity, with their own tunings distinct from those found for monocular masking.

The data in Figure 5.3 show that there are substantial long-range interactions over disparity that are specific to the binocular locations of the stimuli (rather than to the locations of their monocular components). However, a role for the monocular components cannot be dismissed entirely. In fact, there is masking in the 20–60 arcmin range, where it would be predicted by the monocular-masking hypothesis, and the function tends to exhibit a separate peak from the main binocular masking, implying that disparity masking is a combination of the displaced monocular masking and a purely binocular component.

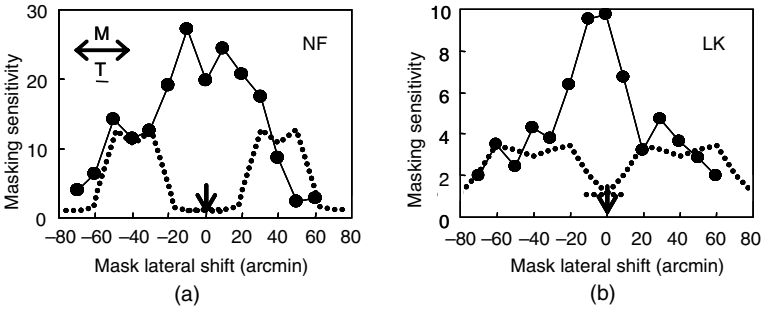


**Figure 5.3** Masking sensitivity (thick line) when the mask is behind the test (80 arcmin disparity for NF and 60 arcmin for LK). The dashed lines depict the disparity masking expected from the combined monocular components. The arrows at bottom depict the test disparity. The mean standard errors over the data set were ±0.04 log units for NF and ±0.04 for LK.

A quantitative, parameter-free prediction of the extent of monocular effect is given by adding the contrast of the monocular components to half that of the binocular test stimulus. The predicted masking is shown by the dashed lines in Figure 5.3, which provide a fair account of the outer flanks of the masking behavior, but the figure shows an additional binocular component centered on zero disparity of the disparity interactions that extends to about ±35 arcmin of positional shift. By this estimate, the purely binocular masking is comparable in frontal extent to the masking from the monocular components.

## 5.6 Disparity selectivity of contrast masking

The second experiment was to measure the masking sensitivity as a function of the stereoscopic disparity of the mask, which was varied in disparity along the same line of sight as the zero-disparity test stimulus. We know, from Figure 5.3, that there is substantial masking at large disparities, of the order of 1° (a half-disparity of ~ 30 arcmin in each eye). To facilitate comparison with the monocular conditions, the masking sensitivity in this experiment is depicted in Figure 5.4 as a function of "half-disparity," which can be conceptualized as the mask displacement from the standard test position in each eye. The data form neither a narrow peak like the monocular masking function nor a uniform mesa out to some disparity limit. Instead, the disparity masking exhibits an unexpected "batwing" shape without precedent in the stereoscopic literature. The masking was two or three times as strong at the peak half-disparities of ±40 arcmin for NF, ±30 arcmin for LK, and ±20 arcmin for LM as in the central region near zero disparity.

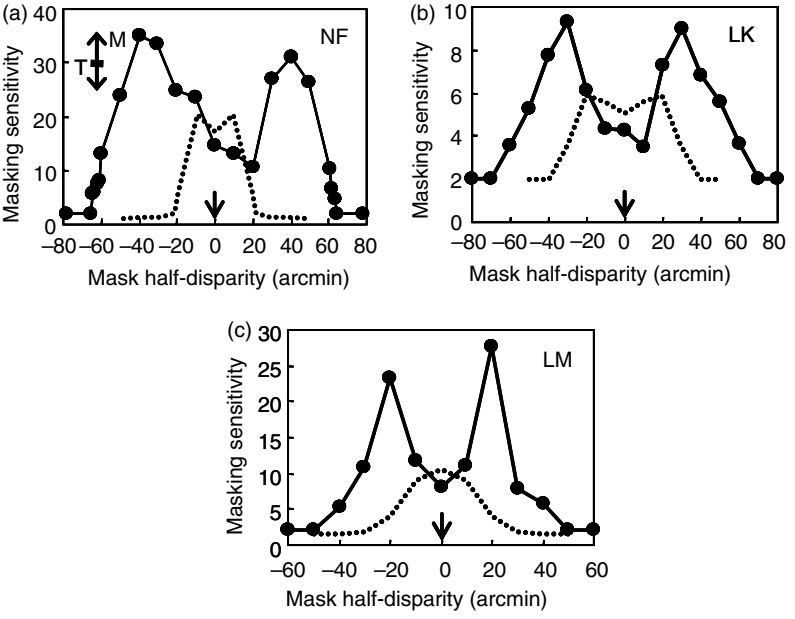


**Figure 5.4** Masking sensitivity as a function of the disparity of a 25 arcmin masking bar at the same mean visual direction (icon) as the 25 arcmin test bar at zero disparity (arrow) for three observers. The solid lines depict the peculiar batwing form of the masking sensitivity, with a minimum near the test disparity and maxima further away from it at crossed and uncrossed disparities. The monocular masking sensitivity for the combined monocular positions implied by each mask disparity is plotted as a dashed line for comparison. The arrows at bottom depict the test disparity. The mean standard errors were ±0.04–0.09 log units.

For this experiment, the role of monocular masking is more complex to evaluate than for Figure 5.3, since it would require convolution with some assumed function of sensitivity to disparity. However, it seems worth noting that the masking in the central region (±20 arcmin) matches the shape that would be expected if this region were dominated by the monocular masking effect from Figures 5.2a–c (shown as thin lines in those figures). On this interpretation, the disparity-specific masking behavior is essentially restricted to two humps extending from about 15 to a maximum of 45–65 arcmin in half-disparity in both the near and the far directions.

The width of the masking peaks is similar to that of the stimuli themselves, implying that the batwing form of the disparity masking could be attributable essentially to just two channels situated at the peak masking disparities. Convolving these locations with the stimulus profile, and adding the unavoidable

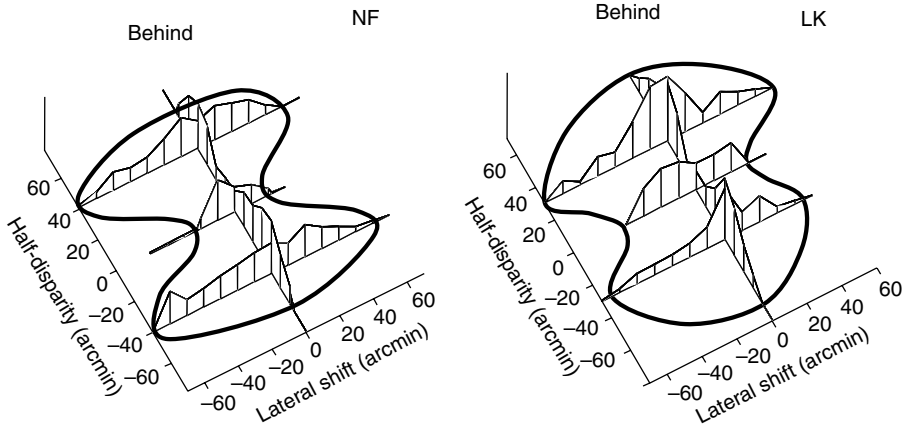


**Figure 5.5** Masking sensitivities replotted from Figures 5.2–5.4, for the 25 arc min test bar at zero disparity as a function of the location of the 25 arcmin masking bar, projected onto orthographic views of the Keplerian array of Figure 5.1. The thick outline depicts the diaboloid form of the masking-sensitivity limit in this disparity space.

monocular component for the zero-disparity mask would explain much of the masking behavior. Such an interpretation implies astonishing specificity in the cross-disparity interactions.

To provide a clearer representation of the significance of the masking effects reported so far, Figure 5.5 plots the available data for two observers in 3D fashion on the plane of the horizontal Keplerian array depicted in Figure 5.1. Here the vertical dimension (thick curves) represents the masking sensitivity for a mechanism responding to a test stimulus at zero disparity and centered at the 5° peripheral location. The measured curves from Figures 5.2–5.4, together with additional slices, combine to define a biconical ("diaboloid") form that expands along the monocular visual-direction lines as mask disparity increases. The outline (gray cross-section) is interpolated through the points where the masking-sensitivity curve meets the baseline and implicitly follows the masking-sensitivity functions elsewhere. The increase in masking strength with disparity away from the horopter may imply a transition from a fine to a coarse disparity system, with the inhibition implied by the masking behavior operating in the coarse-to-fine direction.

## 5.7 Size specificity of disparity masking

One may ask whether the range of disparity masking is (1) a unitary function or (2) specific for the different spatial-frequency channels. To access

a variety of spatial-frequency channels was achieved by varying the width of the Gaussian stimuli constituting the test and mask stimuli. As described in Section 5.3, the Gaussian test profile is a remarkably selective probe for particular spatial-frequency-selective mechanisms, below the peak of the sensitivity function (Kulikowski and King-Smith, 1973; Kontsevich and Tyler, 2004). The widths of the Gaussian test and mask stimuli were varied in tandem in one-octave steps from 50 to 3 arcmin (corresponding to peak spatial frequencies of 0.49 to 7.8 cycles per degree).

The masking-sensitivity functions exhibited two peaks at all spatial frequencies (Figure 5.6), but the peak masking disparities did not remain at the same disparities. Instead, the peaks shifted in rough correspondence with the change in stimulus width, peaking at approximately one stimulus width away from zero disparity all the way from the 50 arcmin to the 6 arcmin stimuli. (Note that, as mentioned, all effects discussed are at least a factor of two, and are highly statistically significant relative to the errors, which are of the order of 0.06 log units, or about 15%.) The narrowest (3 arcmin) stimulus failed to narrow the disparity range further, as would be expected when the Gaussian stimuli pass the peak of the contrast tuning function (Kontsevich and Tyler, 2004). These data indicate that the size selectivity of disparity masking has at least a tenfold range from a width of $\sim 1^\circ$ down to $\sim 6^\circ$ of disparity at half height.

On the other hand, the disparity range of disparity masking turns out to be asymmetric with spatial frequency over the 16-fold range of the measurements. If we focus attention on the upper limit of the masking range (the outer skirts of the functions in Figure 5.6), this limit varies by only about a factor of two (from about $-80$ to $-50$ arcmin) in the negative, or near, direction, compared with a factor of nearly five (from about $+70$ down to $+15$ arcmin) in the positive, or far, direction. There is a corresponding peak asymmetry for this observer that is essentially replicated in all five measured functions. The near-disparity masks produce shallower flanks of masking than do the far disparities, especially for the narrower stimuli. Thus, the near/far disparity asymmetry is consistent across all measured Gaussian-stimulus widths. Note that this data set validates the batwing shape of the disparity-masking function, which becomes even more salient for the narrower stimuli.

## 5.8 Relationship of masking to test disparity: absolute or relative?

Having identified pronounced disparity selectivity in the effects of masking on a test at zero disparity, we may ask how the masking structure varies with the disparity of the test target. Two hypotheses arise.

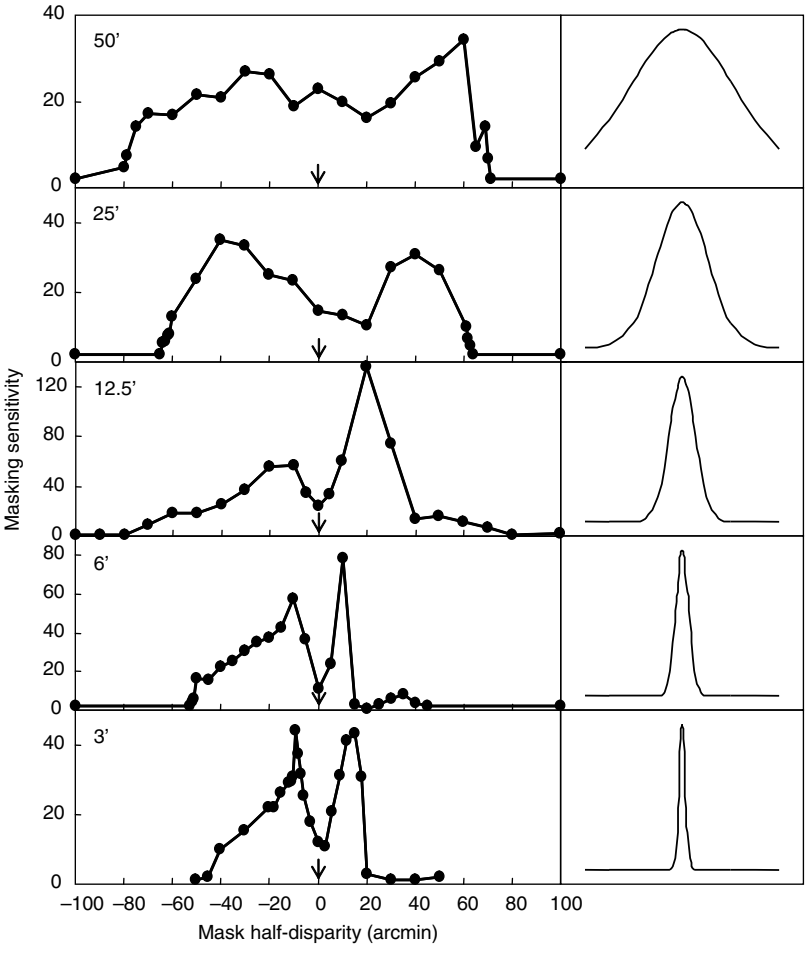


**Figure 5.6** Masking sensitivity as a function of masking disparity for the test bar at zero disparity, measured for a full range of bar widths as indicated (matched for test and mask). The thick lines on the left depict the consistent batwing form of the masking sensitivities, with a minimum near the test disparity. The thin lines in the right panels show the monocular profiles of each stimulus, to scale. The arrows depict the test disparity. From top to bottom, the mean standard errors were ±0.06, ±0.04, ±0.12, ±0.08, and ±0.05 decimal log units.

(1) The masking is a *unitary structure* that is unaffected by the location of the test probe. The masking should then remain at the same absolute disparity range regardless of the test disparity.

(2) The masking structure is *specific* to the disparity of the test probe. The masking function should then shift with disparity to remain locked to the range defined by the probe disparity.

Disparity-masking functions were measured for test disparities set from −80 to +80 arcmin in 40 arcmin increments. For each test disparity, the mask disparity was varied to generate a masking function similar to those in previous graphs. The masking functions conformed to neither prediction alone, but showed aspects of both kinds of hypothesized behavior (Figure 5.7). The outer limits of the masking did not seem to track the disparity of the test, but they were essentially symmetrical for all test disparities. This stability implies that there is a *generic* component of the masking that is relatively invariant with test disparity, providing a broad bluff of masking over the full range of visible disparities (see Figure 5.7 again). In fact, there was a slight tendency for the

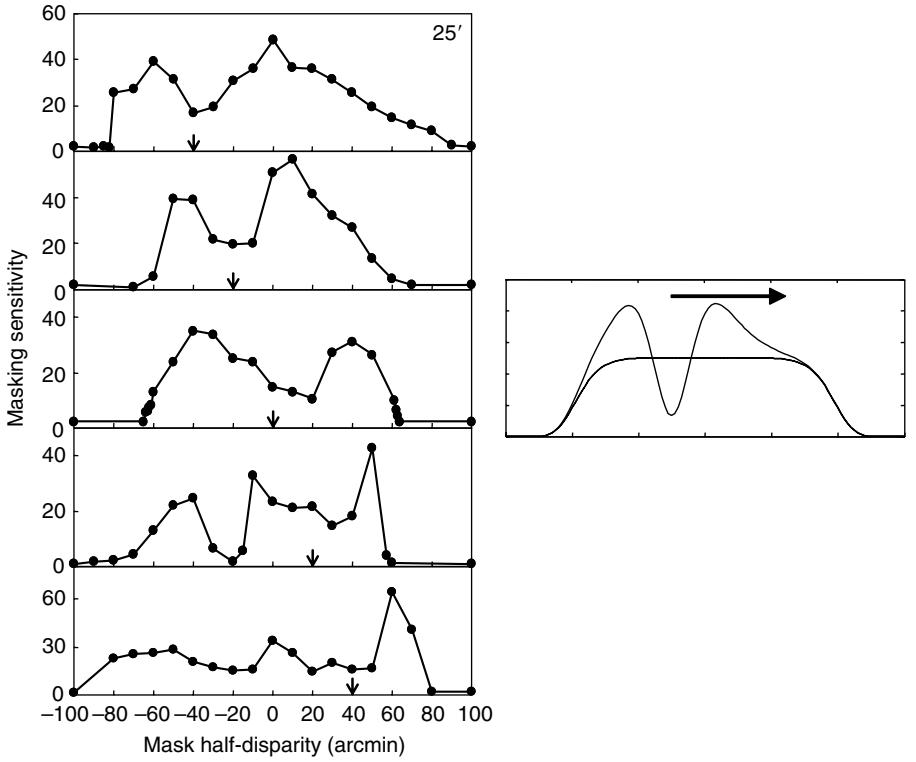


**Figure 5.7** Contrast-masking sensitivities as a function of the disparity of the masking bar for the 25 arcmin test bar at a range of different test disparities (arrows), each showing a minimum near the test disparity indicated by an arrow (together with other irregularities). The arrows at the bottom depict the test disparity, which varied from −80 to 80 arcmin disparity with 40 a arcmin step, with mean standard errors of ±0.01 to ±0.06 decimal log units (i.e., less than the size of the symbols). The inset at the right depicts the output of a conceptual model of the dual-process masking mechanism.

masking range to widen in the disparity range opposite to the test disparity, as though the presence of the test had a reciprocal inhibition on the range of the generic component.

However, there is also a local *disparity-specific* masking effect that appears to be overlaid on this generic masking range. This takes the form of a double peak of enhanced masking at disparities of about 15 arcmin on either side of the test disparity. There also seems to be a tendency for the masking at the test disparity to be lower than that at the extremes of the generic component, just before it starts to descend to zero masking. Taken together, these effects imply that the disparity-specific masking takes the form of a *wavelet* of central disinhibition flanked by peaks of enhanced masking.

## 5.9 Computational model

The behavior described up to this point may be captured in a computational model based on the following equations. The Keplerian space of Figure 5.1 may be expressed as a composite of position ($x$) and disparity ($z$) dimensions. When scaled in appropriate units, the left-eye parallels are given by $l = x + z$, and the right-eye parallels by $r = z - x$. The monocular masking behavior is modeled by a Gaussian along the $l$ and $r$ directions:

$$M_{\mathrm{m}} = e^{-(l/\sigma_{\mathrm{m}})^2} + e^{-(r/\sigma_{\mathrm{m}})^2}. \tag{5.1}$$

The generic component of the disparity-domain masking plateau $M_z$ is given by a variant of the Gaussian with a higher power:

$$M_z = e^{-(z/\sigma_z)^8}. \tag{5.2}$$

The disparity-specific component $M_{\mathrm{d}}$ is a function keyed relative to the disparity of the test target, (symbolized by $\triangle z$), rather than to the convergence angle (the absolute disparity, $z$). In these terms, the disparity-specific component is a wavelet given by a difference of Gaussians keyed to the relative disparity, scaled with the stimulus width $\sigma_{\mathrm{s}}$:

$$M_{\mathrm{d}} = e^{-(2\ \triangle z/\sigma_{\mathrm{s}})^2} - e^{-(\triangle z/\sigma_{\mathrm{s}})^2}. \tag{5.3}$$

This idea of a combined generic masking plateau $M_z$ and a disparity-specific wavelet $M_{\mathrm{d}}$ is depicted in the right panel of Figure 5.7, where the wavelet of inhibition is superimposed on a stable base function and travels with the disparity of the test stimulus as implied by the arrow. The data support the idea of two peaks and a dip between them (evident at all spatial frequencies in the data

of Figure 5.6, and all test disparities in Figure 5.7). This traveling dip represents a minimum in the masking function, in that there is always less masking when the mask is at the test disparity than when it is at adjacent disparities. The dip therefore implies a component of masking due to some facilitatory influence between mechanisms at neighboring disparities, rather than to local masking by two targets at the same disparity.

The conceptual structure of the full scope of disparity interactions is derived as a weighted sum of the three masking effects described by Eq. (5.1)–(5.3):

$$M = a_{\mathrm{m}} M_{\mathrm{m}} + a_z M_z + a_{\mathrm{d}} M_{\mathrm{d}}. \tag{5.4}$$

Thus, the full model consists of two local monocular masking effects, a broad generic disparity-masking effect, and a disparity-specific facilitatory wavelet effect. Equation (5.4) leaves open the issue of whether and how the scaling constants $a_{\mathrm{m},z,\mathrm{d}}$ depend on the contrast and polarity of the targets.

The three components and the resultant masking pattern are illustrated in Figure 5.8 for the case of a same-polarity mask varying in disparity around the test location. The panels depict a Keplerian array in the format of Figure 5.1, with the visual axes for the two eyes schematized as lying along the two diagonals of the panels. The monocular masking effect takes place near the monocular retinal locations of the binocular mask and consequently produces a cross-like inhibition pattern in the Keplerian array along the lines of projection of the monocular projections for the two eyes (Figure 5.8, first panel). The wavelet component produces two inhibitory areas, in front of and behind the mask (second panel). The generic binocular component spreads broadly across both depth and lateral positions (third panel). These three components combined

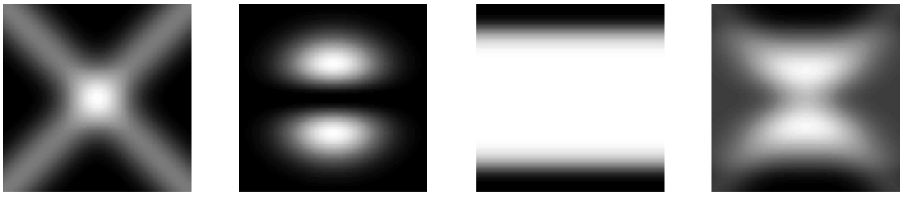


**Figure 5.8** Masking components in the Keplerian-array coordinates of Figure 5.1 (lateral position on the horizontal axis and near/far disparity on the vertical axis). The summed monocular, binocular polarity-specific, and generic binocular components of the masking are depicted separately in the respective panels. They combine to produce the signature "diaboloid" shape of the masking in the "combined" panel, which matches the empirical masking pattern (see Figure 5.5).

produce an inhibitory pattern of disparity interactions (fourth panel) similar to the empirical behavior shown in Figure 5.5.

In addition to these two structural regularities, there seem to be some idiosyncrasies in the masking behavior for the uncrossed (positive) test disparities in Figure 5.7. At a test disparity of 10 arcmin, an additional minimum appears at −10 arcmin mask disparity. This feature is confirmed more weakly in the −20 arcmin curve, but a symmetric dip does not appear in the cases of crossed (negative) test disparity.

## 5.10 Polarity specificity of disparity masking

The scaling parameters in Eq. (5.4) may be regarded as abstract scalars, but there is an alternative option, to treat them as predictions derived from the (signed) Weber contrast which would make the predicted function depend on the sign of the luminance contrast polarity. To test one aspect of this prediction, the contrast polarity of the mask target was inverted to be negative (dark) Gaussians in both eyes relative to the gray background, while the test target was retained as positive. The prediction is that any component of the masking function that is polarity-specific should invert. Conversely, if the masking is determined by the contrast energy at a particular location in space, no change should be expected in the masking function.

The data clearly demonstrate that the use of a dark instead of a light mask does radically alter the masking function, although it does not generate facilitatory behavior in test detectability, as shown in Figure 5.9a. To a first approximation, the batwing form of the test-disparity-selective component seems to be inverted, as depicted in the conceptual model of Figure 5.9b, in which the inversion matches the qualitative features of the dark-bar masking, whose maximum occurs near the (zero) test disparity. The regions of maximum same-sign masking, at about ±40 arcmin disparity, now exhibit local minima on either side of the peak in the opposite-sign masking effect, sharing many of the features of the conceptual model of a disparity-specific wavelet added to a polarity-invariant masking plateau (as in the inset to Figure 5.7). However, the plateau component of the masking model in Figure 5.9b is unchanged, implying that both the monocular and the generic components are proportional to the absolute rather than the signed Weber contrast ($a_m \approx |c|$, $a_z \approx |c|$). This initial assay reveals that a polarity-specificity-masking mechanism exists and may be employed in evaluating the multicomponent hypothesis of disparity masking. For a complete understanding of the contrast relationships, however, one would have to measure the same-sign masking for both light and dark stimuli, and for

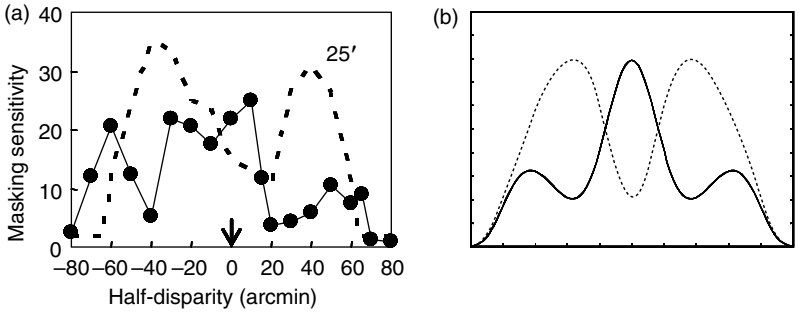


**Figure 5.9** Comparison of masking by a dark 25 arcmin bar (full curve) with that by a light 25 arcmin bar (dashed curve, reproduced from Figure 5.4 for test and mask with positive polarity). Note the inversion of the two main peaks and the central trough around a moderate masking level to generate two troughs flanking the central peak. The arrows at the bottom depict the test disparity. The mean standard error was ±0.03 log units over the three to five measurements per point. (b) Conceptual model showing the effect of inverting the polarity of the disparity-specific wavelet added to a polarity-invariant masking plateau.

the opposite pair of opposite-sign stimuli, and do so as a function of both spatial frequency and contrast.

## 5.11 The nature of disparity masking

The original motivation for this masking study, as implied in the introduction (Section 5.1), was as a technique for evaluating the local channel structure of stereoscopic processing along the lines of the paradigm of Stromeyer and Julesz (1972) for luminance contrast or Stevenson *et al.* (1992) for dynamic noise disparity planes. However, masking in general is composed of at least two distinct processes: self-masking within a channel as the presence of the mask activates the channel and degrades its ability to respond to an additional stimulus, and inhibitory responses between channels as activation of one channel reduces the response in a neighboring channel by reciprocal inhibition. In spatial vision, the masking paradigm has generally revealed a simple structure that is readily interpretable in terms of self-masking within channels. The lateral masking effects in Figure 5.2a conform to the self-masking model, for example.

At 5° eccentricity, the masking range extends about 1° around the lines of sight of the two eyes and 1–3° in disparity, depending on the size of the test stimuli. It is evident from the model structure shown in Figure 5.8, however, that there is a large degree of disparity-specific masking that cannot be explained by the masking of its monocular constituents. The pattern of the results for

local masks in the disparity domain reveals a complex structure of inhibitory interrelations among channels at different disparities. In particular, some of the inhibitory pathways, such as the generic masking plateau, are polarity-independent, while others, such as the disparity-specific disinhibitory wavelet, are polarity-specific. The masking is, in turn, specific to the position, disparity, size (spatial frequency), and contrast polarity of the mask. The resultant structure of disparity interactions is captured qualitatively by a computational model of these three masking components. Both computational and neurophysiological analyses will benefit from building such masking behavior into future models of disparity encoding of depth information.

## 5.12 Relation to the 3D environment

In making fixations on objects in the 3D environment, it is evident that the binocular system will be faced by a large array of features at various spatial positions relative to the feature being fixated. The present results imply that any objects within the sector between the lines of sight of an object will degrade the ability to resolve its disparity. For solid objects, this issue does not usually arise, since they will be likely to occlude each other, so the relevant situation is when the objects are partially transparent, such as when one is looking in a glass-fronted shop window or looking through the surface of a lake. The present results are particularly relevant to situations where information is presented in this transparent form, such as in the "head-up" displays being developed for aircraft cockpits. The data imply that there are extended 3D masking effects in this situation, which are much stronger than those predicted from the local monocular masking effects. These interactions imply that, in complex 3D environments, we have to navigate a maze of 3D spatial interactions in the stereoscopic visual representation in order to make sense of the visual world.

### Acknowledgments

This work was supported by AFOSR grant number FA9550-09-1-0678.

### References

Chen, C. C. and Tyler, C. W. (1999). Spatial pattern summation is phase-insensitive in the fovea but not in the periphery. *Spat. Vis.*, 12: 267–285.

Chen, C. C. and Tyler, C. W. (2001). Lateral sensitivity modulation explains the flanker effect in contrast discrimination. *Proc. R. Soc. Biol. Sci.*, 268: 509–516.

Chen C. C. and Tyler, C. W. (2006). Evidence for elongated receptive field structure for mechanisms subserving stereopsis. *Vis. Res.*, 46: 2691–2702.

Chen, C. C. and Tyler, C. W. (2008) Excitatory and inhibitory interaction fields of flankers revealed by contrast-masking functions. *J. Vis.*, 8: 10.1–10.14.

Field, D. J., Hayes, A., and Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local "association field". *Vis. Res.*, 33: 173–193.

Foley J. M. (1994). Human luminance pattern-vision mechanisms: masking experiments require a new model. *J. Opt. Soc. Am. A*, 11: 1710–1719.

Hess, R. F., Hayes, A., and Kingdom, F. A. (1997). Integrating contours within and through depth. *Vis. Res.*, 37: 691–696.

Hess, R. F., Hayes, A., and Field, D. J. (2003). Contour integration and cortical processing. *J. Physiolol. (Paris)*, 97: 105–119.

Julesz, B. (1971). *Foundations of Cyclopean Perception*. Chicago, IL: University of Chicago Press.

Julesz, B. (1978). *Global Stereopsis: Cooperative phenomena in Stereoscopic Depth Perception*. Berlin: Springer.

Kepler, J. (1611). *Dioptrice*. Augsburg: Vindelicorum.

Kontsevich, L. L. and Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vis. Res.*, 39: 2729–2737.

Kontsevich, L. L. and Tyler, C. W. (2004). Local channel structure of sustained peripheral vision. In B. E. Rogowitz and T. N. Pappas (eds.), *Human Vision and Electronic Imaging IX*, pp. 26–33. SPIE, 5292, Bellingham, WA: SPIE.

Kontsevich, L. L. and Tyler, C. W. (2005). The structure of stereoscopic masking: position, disparity, and size tuning. *Vis. Res.*, 43: 3096–3108.

Kulikowski, J. J. and King-Smith, P. E. (1973). Spatial arrangement of line, edge and grating detectors revealed by subthreshold summation. *Vis. Res.*, 13: 1455–1478.

Levi, D. M., Klein, S. A., and Hariharan, S. (2002). Suppressive and facilitatory spatial interactions in foveal vision: foveal crowding is simple contrast masking. *J. Vis.*, 2: 140–166.

Mansfield, J. S. and Parker, A. J. (1993). An orientation-tuned component in the contrast masking of stereopsis. *Vis. Res.*, 33: 1535–1544.

Marr, D. (1982). *Vision*. San Francisco, CA: W. H. Freeman.

Ogle, K. N. (1950). *Researches in Binocular Vision*. Philadelphia, PA: Saunders.

Polat, U. and Sagi, D. (1993). Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments. *Vis. Res.*, 33: 993–999.

Polat, U. and Tyler, C. W. (1999). What pattern the eye sees best. *Vis. Res.*, 39: 887–895.

Stevenson, S. B., Cormack, L. K., Schor, C. M., and Tyler, C. W. (1992). Disparity tuning in mechanisms of human stereopsis. *Vis. Res.*, 32: 1685–1694.

Stiles, W. S. (1939). The directional sensitivity of the retina and the spectral sensitivities of the rods and cones. *Proc. R. Soc. (Lond.) B*, 127: 64–105.

Stromeyer, C. F. and Julesz, B. (1972). Spatial-frequency masking in vision: critical bands and spread of masking. *J. Opt. Soc. Am.*, 62: 1221–1232.

Tyler, C. W. (1973). Stereoscopic vision: cortical limitations and a disparity scaling effect. *Science*, 181: 276–278.

Tyler, C. W. (1975). Characteristics of stereomovement suppression. *Percept. Psychophys.*, 17: 225–230.

Tyler, C. W. (1983). Sensory processing of binocular disparity. In C. Schor and K. J. Ciuffreda (eds.), *Basic and Clinical Aspects of Binocular Vergence Eye Movements*, pp. 199–295. London: Butterworth.

Tyler, C. W. (1991). Cyclopean vision. In D. Regan (ed.), *Binocular Vision*. Vol. 9 of *Vision and Visual Disorders*, pp. 38–74. New York: Macmillan.

Tyler, C. W. (2005). Spatial form as inherently three-dimensional. In M. Jenkin and L. R. Harris (eds.), *Seeing Spatial Form*. Oxford: Oxford University Press.