Available online at www.sciencedirect.com
SCIENCE DIRECT®
Vision Research 45 (2005) 3096–3108
Vision Research
www.elsevier.com/locate/visres

The structure of stereoscopic masking: Position, disparity, and size tuning

Christopher W. Tyler *, Leonid L. Kontsevich

Smith-Kettlewell Eye Research Institute, 2318 Fillmore Street, San Francisco, CA 94115, USA

Received 2 October 2004; received in revised form 28 June 2005

Abstract

The masking effect of a Gaussian blob on detection of a Gaussian target was measured as a function of the position, disparity, width and polarity of the mask. The data reveal a large degree of disparity-specific masking that cannot be explained by the masking of its monocular constituents. 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. The masking effects can be modeled as having three additive components, one that has a fixed disparity range and is polarity independent, one with a center/surround form keyed to both the disparity and the polarity of the mask, and one that derives from the monocular masking in each eye. Thus, the profound disparity interaction behavior is not limited to the simple monocular masking properties of the stimuli but reveals extensive connectivity across the disparity domain. Future models of disparity encoding will need to take these properties into account.

Keywords: Stereopsis; Depth perception; Disparity; Masking

1. Introduction

Masking is a 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. When masking occurs, either of two conceptual frameworks may be invoked to develop its neural interpretation. Under the assumption of univariance, the mask and test are assumed to impact the same coding channel in indistinguishable fashion, exploiting some processing nonlinearity to induce a variation in sensitivity to the test stimulus in the presence of the mask (Mansfield & 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 inhibitory masking does not imply univariance and does not measure the sensitivity of the processing channels for either stimulus, but the inhibitory relations between them. Distinguishing between univariant-channel masking and inhibitory interactions is not possible without additional information about the system characteristics. 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, it makes an unlikely channel structure and the more plausible interpretation is in terms of inhibitory interactions. It is for this reason that the inhibitory interpretation is preferred in framing the present masking results.

* Corresponding author. Tel.: +1 415 345 2105; fax: +1 415 345 8455.
E-mail address: cwt@ski.org (C.W. Tyler).

0042-6989/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2005.07.034

1.1. Disparity-selective local mechanisms

In detail, the conceptual framework of this investigation is the assumption of a parallel array of cortical disparity "registers," each selective for the presence of a stimulus with a particular location in x, y, z coordinate space. The layout of space before the two eyes is represented in Fig. 1 (with the y axis omitted for clarity) where the curved retinas are replaced by planar projection regions. The array of intersecting sight lines is known as the Keplerian array, since Kepler (1611) was the first to analyze the binocular geometry of local disparities. It is assumed that the same structure is mirrored in the visual cortex, where the initial stage of binocular combination is a *disparity registration* process indicated by the black circle at the intersection of two lines of sight. The disparity registration might be implemented by the various types of neuron with facilitatory responses to particular disparities, as recorded in the visual cortex of cat and monkey (Barlow, Blakemore, & Pettigrew, 1967; Poggio, Motter, Squatrito, & Trotter, 1985; Cumming, 2002).

This 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, Cormack, Schor, & Tyler, 1992). The best match or correlation in each local region of the visual field specifies when the local images are in register (Fig. 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). In practice there are, of course, both vertical and horizontal dimensions of field location, and the physiological array may not be as regular as depicted here.

This initial disparity registration process should not be envisaged as a defined perceptual output with disparity detection for an element at a particular depth, but as an array of disparity-coded signals subject to interactive reorganization by subsequent processing stages before consolidation of the resulting depth percept (Tyler, 1983, 1991). The conceptual elements of the input stage for cyclopean processing consist of the monocular receptive fields feeding the disparity registration stage.

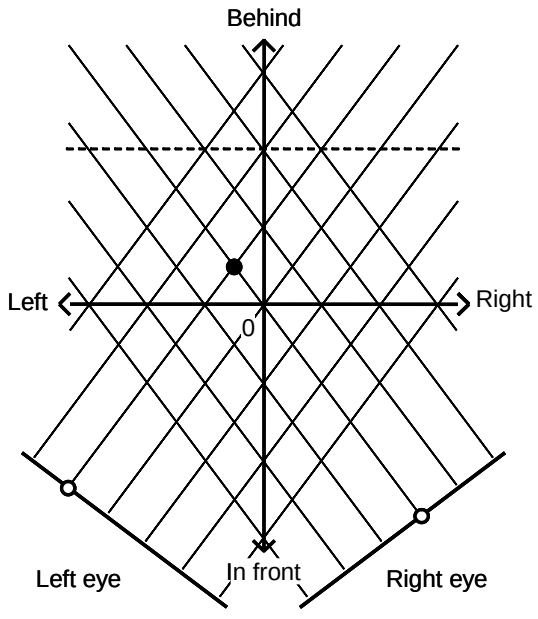

Fig. 1. Keplerian array of disparity detectors depicted by the intersections of the (oblique) lines of sight of the two eyes. The retinas of the left and right eyes are schematized as linear arrays, with one point in each array (open circles) indicated as the recipient of the image of the stimulus point 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.

1.2. Global interactions

Beyond the disparity registration stage 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. 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: local cross-disparity inhibition (along the vertical line in Fig. 1) and disparity-specific pooling or facilitation (along the horizontal line in Fig. 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 processes as the disparity gradient limitation on the upper limit for depth reconstruction (Tyler, 1973), coarse-to-fine matching processes for building up the depth image from the monocular information (Marr, 1982), and so on. 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 goal of the present study is to provide a survey of the scope of such disparity-domain interactions.

Another kind of disparity-specific interaction has been reported by McKee, Bravo, Taylor, and Legge

(1994), in which the introduction of a highly shifted monocular counterpart to the stimulus in one eye eliminates the masking effect of a dichoptic mask on a corresponding monocular test target. The direct interpretation of this release from masking is that the monocular counterpart has shifted the Keplerian location activated by the disparate pair, moving it away from the inhibitory field of the monocular target within which it was originally in correspondence. However, our stimuli were all binocular, with the goal of measuring disparity interactions as a function of testing disparity, spatial frequency, and luminance polarity.

2. Experimental rationale

2.1. Masking paradigm

To take a maximally model-free approach to the question of local channel interactions through the Keplerian array, we adopted the masking sensitivity paradigm (introduced by Stiles, 1939, as the field sensitivity paradigm). In our version of this approach, the test stimulus is set at a fixed contrast level above its contrast detection threshold and the mask contrast required to return the test to threshold is determined as a function of the tuning parameter (position, disparity, or scale). The result is a tuning curve through some aspect of the Keplerian array that has three advantages over threshold elevation curves:

(a) 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 detection threshold.
(b) The test effect on neighboring channels and their potential interactions is minimized because the test contrast is near threshold.
(c) 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.

2.2. Local target structure

To maximize the local specificity of the probe, we needed stimuli that were simultaneously local with respect to four variables: eccentricity, extent, disparity, and spatial frequency. Previous studies have confounded these variables so that it is not possible to disambiguate the effects of retinal inhomogeneity from eccentricity, relative peak positions of test and mask, and spatial frequency effects. Classic sinusoidal grating targets maximize spatial frequency specificity but do so at the expense of both position and eccentricity definitions. More recent compromises have been adopted to constrict the stimulus to more homogeneous regions of retina, such as Gabor patches and sixth-derivative-of-Gaussian bars (Blake & Wilson, 1991; Halpern, Wilson, & Blake, 1996; Rohaly & Wilson, 1993, 1994, 1998; Wilson, Blake, & Halpern, 1991). The problem with such stimuli is that they still have multiple peaks that stimulate multiple disparity nodes in the Keplerian array, and they also allow unspecified phase interactions in between the peaks as the masking frequency is varied relative to test frequency (the beats of Graham & Nachmias, 1971).

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 & King-Smith, 1973). Such Gaussian bars allow measurement by contrast masking of our three requisite variables, position sensitivity, disparity sensitivity, and spatial frequency (scale) selectivity (Kontsevich & Tyler, 2004). Gaussians have only one peak, so that the position of the masking bar can be varied cleanly relative to the peak of the test bar. The lack of side-lobes in the 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. The Gaussian bars also provide a substantial degree of tuning in spatial frequency (Kontsevich & Tyler, 2004). If the Gaussian is smaller than the receptive field center, it provides less than optimal activation; if it is 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.

3. Methods

We used 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. A bright fixation square 10 arc min on a side was provided at the primary position of each eye, and observation of the observer's eyes during testing verified that fixation was stable in these practiced observers. Both test and mask blobs had a Gaussian horizontal profile of width 25 arc min at half height (or variable in one experiment) and a four-times wider Gaussian profile vertically. Background luminance was 25 cd/m^2. The

test stimuli were presented in a temporal raised-cosine bell of 2 s duration while the masks were static throughout the trial. For the test bar of a particular width the contrast was set at a level providing more that 95% correct. The contrast of the masking bar at a range of set positions, disparities or widths was varied to measure the mask contrast required to return the test to its criterion level of detectability.

The experimental set-up for stereoscopic masking employed four screens, as depicted in Fig. 2. The observer viewed two monitor screens with each eye, combined in pairs through beam-splitter cubes; one screen was for the test and the other for the mask, to allow for independent control of test and mask contrast without re-computing the stimulus profile on each trial. The monitors were set at a distance of 1.28 m, so as to produce pixels of 1 arc min, and configured so that convergence was at the same distance. The test screens were viewed through additional front-surface mirrors, which were required for the spatial configuration of the four monitors. It was assumed that the horopter was frontoparallel, which is true to good approximation, since the empirical horopter inflects from convex to concave at about this viewing distance (Ogle, 1950). The four monitors controlled by four graphic cards in a Macintosh computer were accurately calibrated. The look-up tables for each monitor were updated asynchronously with the average delay between updates as small as 30 ms. Each look-up table was computed for its exact update time, so that there was no overall lag between gradual contrast transitions in different monitors, which determined our choice of the 2 s presentation epoch.

Since the test and mask were mixed with beam splitters, the Weber contrast of the mask did not exceed 0.5 and, therefore, the measured masking sensitivity could not have a value lower than 2 (a limitation that should be kept in mind when evaluating the data figures). Thresholds were measured in terms of masking sensitivity, the mask intensity required to return the test to the measured test threshold when initiated at a fixed multiple above threshold. The psychophysical task was to vary the mask contrast $C_{\mathrm{mask}}(x, z)$ to the level that brings C_{test} back to its unmasked level of detectability. The masking sensitivity parameter is defined as reciprocal to $C_{\mathrm{mask}}(x, z)$.

$$S_{\mathrm{mask}}(x, z) = 1/C_{\mathrm{mask}}(x, z).$$

Thresholds were determined in a two-interval forced-choice paradigm with the mask set to the same (variable) level in both intervals, while the fixed-contrast test was presented in one or other intervals at random. Maximum control was maintained over the observers' criterion effects by the Ψ adaptive staircase method that we have developed (Kontsevich & Tyler, 1999) based on minimum entropy criteria for setting successive test contrasts. Each staircase had 40 steps, and the threshold value was measured by at least four repeats of the staircase. If the results of the first four were inconsistent (i.e., had a range of more than a factor of 2), further staircases were run until four consistent values were obtained in sequence. This procedure resulted in a mean standard error of about ± 0.05 decimal log units in all experiments.

To survey the terrain of the multidimensional disparity masking space, we elected to concentrate on the performance of a single highly trained observer NF (56 years old, corrected to normal vision). Key results were replicated on two observers LK and LM (40 and 20 years old, both with normal vision). The masking sensitivity approach requires stable and consistent performance because it depends on a well-established value for threshold to set the value for twice the threshold level. Each session consisted of a full run through all masking values for a particular test condition. Where it became apparent that the value was unstable (by failure of the staircase to converge to criterion), the unmasked contrast detection threshold was re-measured to ensure that the test contrast was set at the specified multiple of the detection threshold contrast for the observer's current neural state. The test contrasts were set at 2× detection threshold for NF, 3× detection threshold for LM, and 4× detection threshold for LK to provide performance better than 95% correct for the test alone. These multipliers were the minimal integer values that met the performance criterion for each observer.

Statistical analysis of this kind of functional survey is problematic. We therefore took the approach of specifying the mean log error, which was typically of the order of 10–25%, and limiting discussion to effects that are at least a factor of 2, or 4–8 standard errors of the individual values, and therefore highly statistically significant. Moreover, all the effects are well-validated by neighboring points, and thus even more significant.

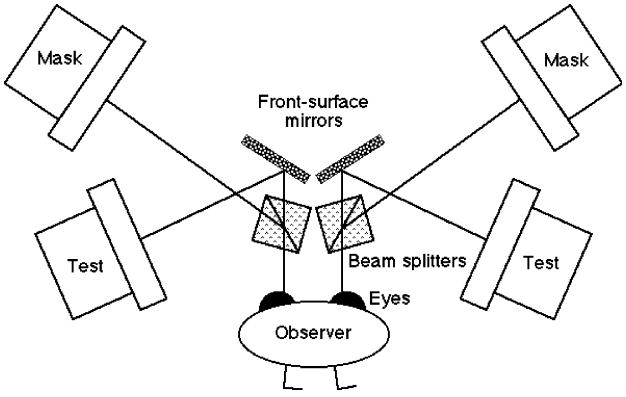

Fig. 2. Depiction of the binocular geometry of the monitors presenting the test and mask stimuli. (Horizontal distances to the monitors compressed.)

4. Results and discussion

4.1. Position tuning

The first experiment was a control study to determine the position tuning of the local contrast masking effect with respect to mask position (solid horizontal line in Fig. 1). With both test and mask of width 25 arc min at half height (Fig. 3D, solid line) corresponding to a peak spatial frequency of 0.97 cy/deg, the masker lateral position was varied in steps of 10 arc min, or finer if necessary to characterize the function adequately. Mask contrast required to reduce the test back to threshold was minimal beyond about ±20 arc min separation from the test for observer NF (±40 arc min for LK and ±30 arc min for LM) but highly effective inside these limits (Figs. 3A–C). The function was asymmetrical for observers NF and LM, with the peak displaced by 10 arc min toward fixation for NF (Fig. 3A) and 10 arc - min away from fixation for LM (Fig. 3C). 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 $\sqrt{2} * 25 \approx 35$ arc min width at half height, as shown by the dashed line in Fig. 3D). Instead the masking has an idiosyncratic form that evidently reflects the tuning of the underlying spatial channel (under the channel invariance assumption laid out in the first paragraph of the Introduction) rather than involuntary eye movements or retinal inhomogeneity, which should be consistent across observers.

The first disparity-related condition was to set the masking bar of 25 arc min at a non-zero disparity while the test remained in the horopter (zero disparity) and to measure the degree of masking as a function of lateral position of this disparate mask (dashed horizontal line in Fig. 1). The mask disparity was set at 80 arc min (40 arc min displacement in each eye) for NF and 60 arc min for LK. Three hypotheses come readily to mind for the outcome of this experiment:

H1. Masking is a simple combination of the monocular effects expected when each of the two bars moves into the masking range, depicted in Figs. 3A and B for the two observers. According to this hypothesis (dashed lines in Figs. 4A and B), the stimulus at about one degree of disparity should produce two masking peaks separated by a gap of reduced masking.

H2. Masking is narrowly disparity-specific (as in Stevenson et al., 1992), which would predict that no masking would occur between elements set at this large disparity from one another.

H3. Masking reveals the presence of long-range interactions across disparity with their own tunings distinct from those found for monocular masking. This hypothesis comes in two flavors; one in which the monocular masking combines with the binocular interactions, the other in which monocular interactions are suppressed in the presence of binocular stimuli (Tyler, 1971).

Clearly, the data in Figs. 4A and B eliminate the first two hypotheses since there is strong masking in the central region of the position range where both hypotheses predict an absence of masking. Thus, there must be substantial long-range interactions over disparity that are specific to the binocular locations of the stimuli, rather than the locations of their monocular components. On

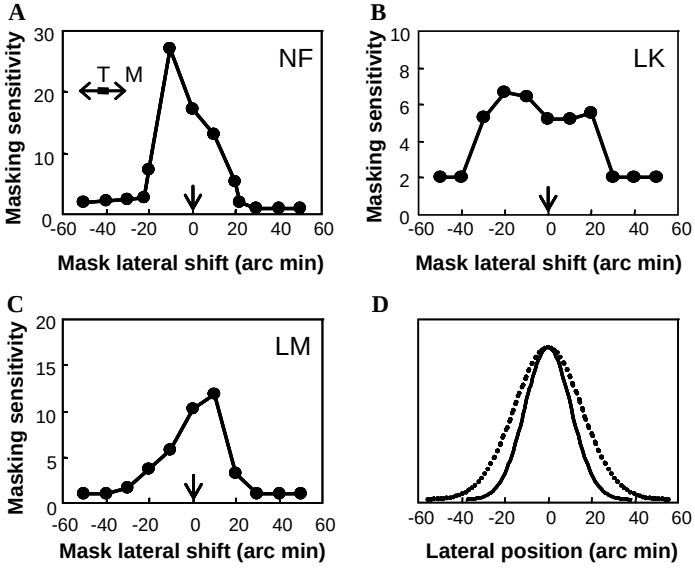

Fig. 3. Contrast masking as a function of lateral shift of the masking bar (M) from the test bar (T), both of which had profiles 25 arc min wide at half height shown as the solid line in (D). (A–C) Masking sensitivity with both mask and test at zero disparity (icon). Arrows at bottom indicate test position. Mean standard errors were ±0.04 log units for NF, ±0.09 for LK and ±0.05 for LM, i.e., about the size of the symbols. (D) The test and the mask Gaussian profile (solid line) and the expected test/mask interaction profile (dashed line).

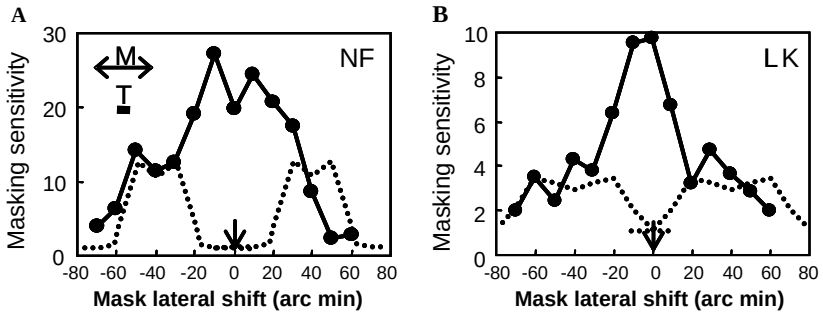

Fig. 4. Masking sensitivity (thick line) when mask is behind the test (80 arc min disparity for NF and 60 arc min for LK). Dashed lines depict disparity masking expected from the combined monocular components. Arrows at bottom indicate test disparity. Mean standard errors over the data set were ±0.04 log units for both NF and LK.

the other hand, a role for the monocular components cannot be dismissed entirely. In fact, there is masking in the 20–60 arc min 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. (Looking forward, this peak is replicated for the crossed disparity condition shown in a subsequent figure). This wide range of masking tends to support the first flavor of hypothesis 3, in which disparity masking is a combination of the displaced monocular masking and a purely binocular component.

We can make a quantitative, parameter-free prediction for the extent of the combined monocular effect by noting that the contrast of the monocular components added to the uniform field background is half that of the binocular test stimulus. On the assumption of linear superposition of the monocular masking effects, we would predict a masking effect of a half of the total measured in the monocular experiment, with the effect should be shifted by ±40 and ±30 arc min for the two observers to account for lateral shifts of the monocular stimuli to produce the mask disparity. The predicted masking is shown as the dashed lines in Figs. 4A and B, which provide a fair account of the outer flanks of the masking behavior. On this interpretation, the binocular component of the disparity interactions is that part of the function that exceeds the monocular predictions, and therefore extends to about ±35 arc min of positional shift (which is −70 arc min of disparity in this experiment). By this estimate, the purely binocular masking component is comparable in frontal extent with the masking from the monocular components.

4.2. 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 always held along the same line of sight as the zero-disparity test stimulus. We know, from Figs. 4A and B, that there is substantial masking at large disparities, of the order of 1° (half-disparity of ~30 arc min in each eye). Two hypotheses for the disparity range of this interaction arise.

H4. That the masking is maximal for small disparities and falls off with a single-humped function, such as a Gaussian, as disparity is increased.

H5. That the masking is carried by each of a series of disparity-selective channels (Stevenson et al., 1992) and therefore remains fairly uniform with disparity until it finally collapses when the extent of last channel is exceeded.

The data obtained do not support either of these hypotheses. To facilitate comparison with the monocular conditions, masking sensitivity in this experiment is depicted in Fig. 5 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, disparity masking exhibits an unexpected "batwing" shape without precedent in the stereoscopic literature. Masking was two or three times as strong at the peak half-disparities of ±40 arc min for NF, ±30 arc min for LK and ±20 arc min for LM as in the central region near zero disparity.

For this experiment, the role of monocular masking is more complex to evaluate than for Fig. 4, since it would require convolution with some assumed function of sensitivity to disparity. However, it seems worth noting that the masking in the central (±20 arc min) region matches that expected if this region were dominated by the monocular masking effect from Figs. 3A–C (shown as the thin lines in Figs. 5A–C). On this interpretation, the disparity-specific masking behavior is essentially restricted to two humps extending from about 15 to a maximum of 45–65 arc min in half-disparity in both near and far directions.

We were struck by the steepness of the fall-off in masking sensitivity at the outer edge of the disparity range for observer NF. To characterize this fall-off in

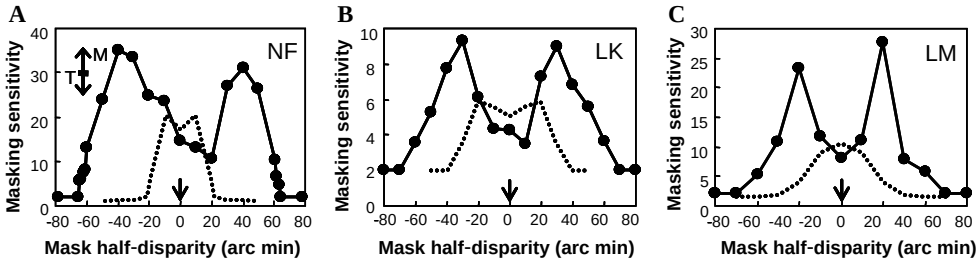

Fig. 5. Masking sensitivity as a function of disparity of a 25 arc min masking bar at the same mean visual direction (icon) as the 25 arc min test bar at zero disparity (arrow) for observer NF (A), observer LK (B), and observer LM (C). Solid lines depict the peculiar batwing form of the masking sensitivity, with a minimum near the test disparity and maxima at crossed and uncrossed disparities. Monocular masking sensitivity for the combined monocular positions implied by each mask disparity is plotted by dashed line for comparison. Arrows at bottom indicate test disparity. Mean standard error was ±0.04 log units for NF, ±0.09 log units for LK, and ±0.04 log units for LM.

detail, we tested the masking effect in half-disparity increments of 2 arc min (Fig. 5A). It may be seen that the data remained highly consistent, despite the fact that this is in the range where the mask appears diplopic. In fact, the entire region of maximum masking is in the diplopic range, which begins at about 10 arc min. Nevertheless, the data fall from half their maximum sensitivity almost to baseline within about 6 arc min at both extremes. This rapid fall-off certainly belies the monocular structure of these stimuli, which was 25 arc min in width at half height (Fig. 3D). The data for observers LK and LM (Figs. 5B and C) do not show the same steep flanks, but otherwise replicate the same overall masking behavior.

Even disregarding this inconsistency in the tails, 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 half-disparities of ±40 arc min for NF, ±30 arc min for LK, and ±20 arc min for LM. Convolving these locations with the stimulus profile, and adding the unavoidable 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 scope of the masking effects reported so far, Fig. 6 plots the available data for two observers in 3D fashion on the plane of the horizontal Keplerian array depicted in Fig. 1. Here the vertical dimension (black curves) represents the masking sensitivity for the mechanism responding to a test stimulus at zero disparity and centered at the 5° peripheral location. The measured curves from Figs. 3–5, 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

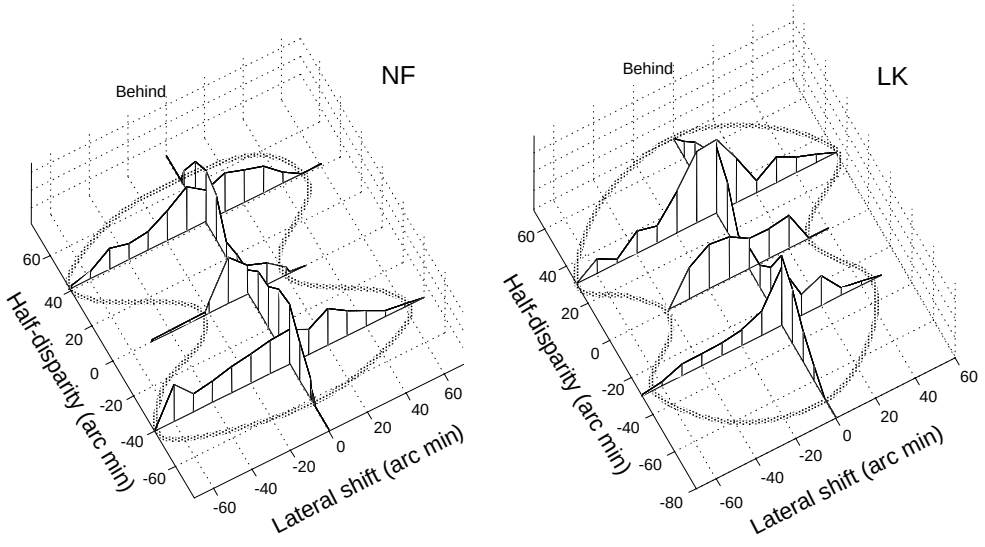

Fig. 6. Masking sensitivities replotted from Figs. 3–5, for the 25 arc min test bar at zero disparity as a function of the location of the 25 arc min masking bar, projected onto orthographic views of the Keplerian array of Fig. 1. The gray outline depicts the diaboloid form of the masking sensitivity limit in this disparity space.

meets the baseline and implicitly follows the masking sensitivity functions elsewhere. The increase in masking strength with disparity away from the horopter may imply transition from a fine to a coarse disparity system, with the inhibition implied by the masking behavior operating in the coarse-to-fine direction.

4.3. 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, we simply varied the width of the Gaussian stimuli constituting the test and mask stimuli. As developed in Section 1, the Gaussian test profile is a remarkably selective probe for particular spatial-frequency-selective mechanisms, below the peak of the sensitivity function (Kontsevich & Tyler, 2004; Kulikowski & King-Smith, 1973). We varied in tandem the width of the Gaussian test and mask stimuli in one-octave steps from 50 to 3 arc min (corresponding to peak spatial frequencies of 0.49–7.8 cycles per degree).

The masking sensitivity functions for observer NF exhibited two peaks at all spatial frequencies (Fig. 7), but the peaks did not remain at the same disparities. Instead, the peak masking disparities shifted in rough correspondence with the change in stimulus width, peaking at approximately one stimulus width away from zero disparity all the way from 50 to 6 arc min stimuli. (Note that, as mentioned, all effects discussed are at least a factor of two and highly statistically significant relative to the errors of the order of 0.06 log units, or about 15%.)

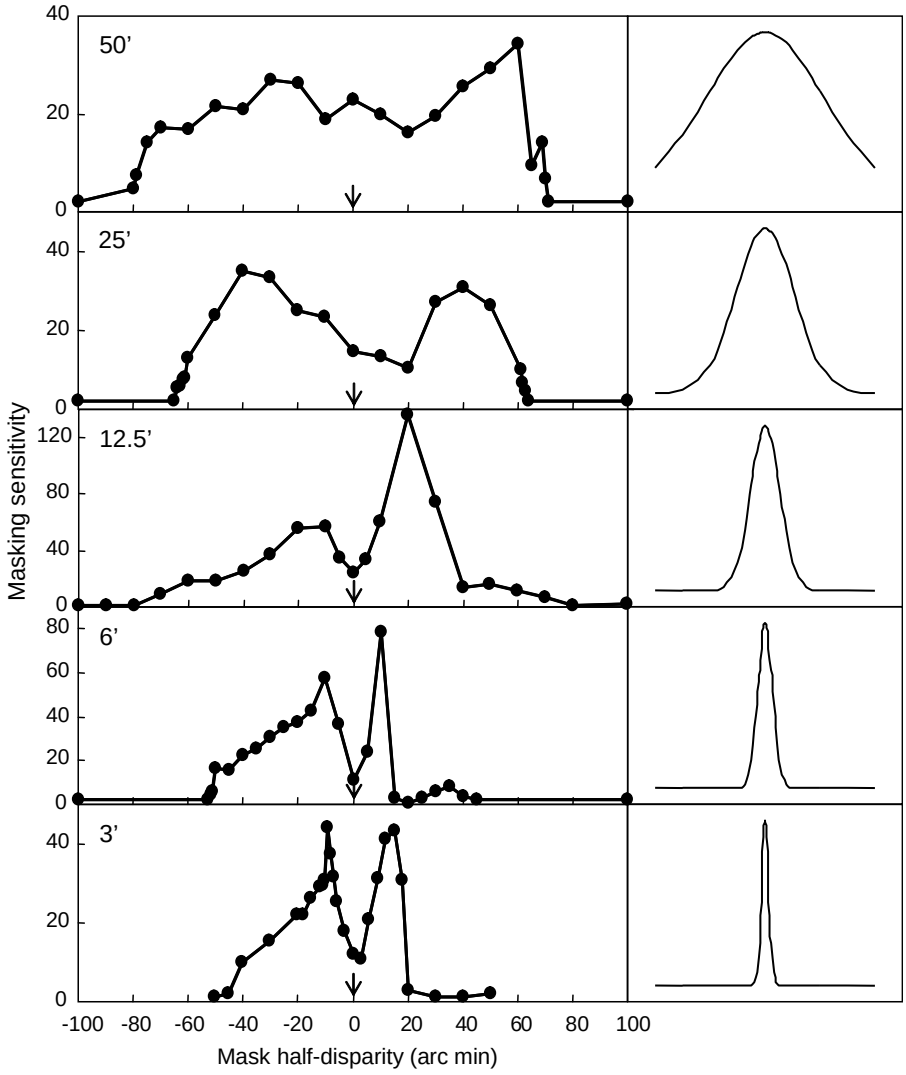

Fig. 7. Masking sensitivity as a function of masking disparity for the test bar at zero disparity, measured for a full range of bar widths as indicated (matched for test and mask). Thick lines on the left depict the peculiar batwing form of the masking sensitivities, with a minimum near the test disparity. The thin lines in the right panels show the monocular profiles of each stimulus, to scale. Arrows at bottom depict test disparity. Mean standard errors were ±0.06, ±0.04 ±0.12, ±0.08, and ±0.05 decimal log units, respectively, for the data sets from top to bottom.

The narrowest (3 arc min) stimulus failed to narrow the disparity range further, as should be expected when the Gaussian stimuli pass the peak of the contrast tuning function (Kontsevich & Tyler, 2004). These data indicate that the *size* selectivity of disparity masking has at least a 10-fold range.

On the other hand, the *disparity* range of disparity masking turns out to be asymmetric with spatial frequency over the 16-fold range that we were able to measure. If we focus attention on the upper limit of the masking range (the outer skirts of the functions in Fig. 7), it varies by only about a factor of 2 (from about −80 to −50 arc min) in the negative, or near, direction, compared with a factor of nearly 5 (from about +70 down to +15 arc min) 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 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.

4.4. Relationship of masking to test disparity: Absolute or relative?

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

H6. 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 test disparity.

H7. The masking structure is *specific* to the disparity tapped by the test probe. The masking function should then shift with disparity to remain locked to the relative range defined by the probe disparity.

Disparity masking functions were measured for test disparities set from −80 to + 80 arc min in 10 arc min increments. For each test disparity, mask disparity was varied to generate a masking function similar to those in previous graphs. The masking functions measured for observer NF conformed to neither prediction alone, but showed aspects of both kinds of hypothesized behavior

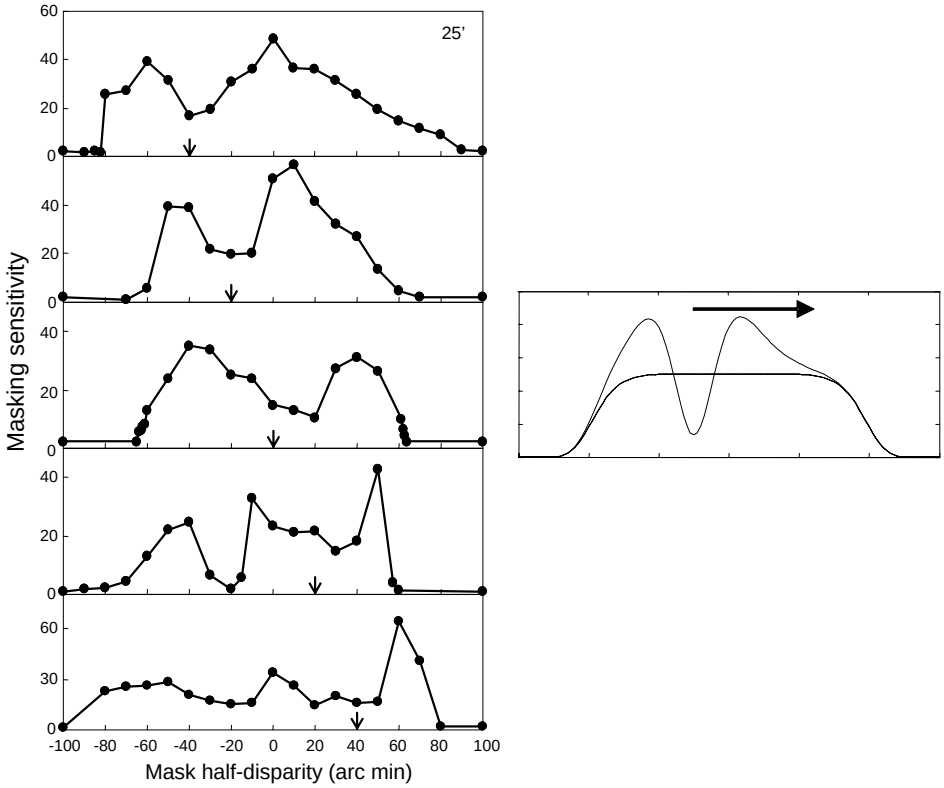

Fig. 8. Contrast masking sensitivities as a function of the disparity of the masking bar for the 25 arc min test bar at range of different test disparities (arrows), each showing a minimum near the test disparity indicated by an arrow (together with other irregularities). Arrows at bottom depict test disparity, which varied from −80 to 80 arc min disparity in 10 arc min steps. Mean standard errors were ±0.04, ±0.03, ±0.04, ±0.01, and ±0.06 decimal log units respectively for the data sets from top to bottom. Inset at right depicts the output of a conceptual model of the dual-process masking mechanism.

(Fig. 8). The full range of the masking function did not seem to track the disparity of the test. The range was essentially symmetrical for all test disparities, although it broadened from about 65 arc min for the zero-disparity test to about 85 arc min for the most extreme test disparities (both positive and negative). 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 (Fig. 8). In fact, there was a slight tendency for the range to widen in the disparity range opposite the test disparity, as though the presence of the test had a reciprocal inhibition on the range of the generic component. The masking seems to have a steep flank on the side nearest to the test disparity (and on both sides for the zero-disparity test). This steep flank is coupled with enhanced masking for masks at about ±20 arc min from the test disparity.

However, there is also a local *disparity-specific* masking effect appears to be overlaid on this generic masking range. This idea of the combined generic masking plateau and disparity-specific wavelet is depicted in the conceptual model of the right panel of Fig. 8, 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 the dip between them (evident at all spatial frequencies in the data of Fig. 8, left panel) that remained locked to the test disparity (arrows) in the sense that there was always a dip near the disparity of the test (although it was not always the only dip in the masking function). This traveling dip represents a minimum in the masking function, in that there is always less masking when the mask was at the test disparity than when it was at adjacent disparities. The dip therefore implies a component of masking due to some disinhibitory influence between mechanisms at neighboring disparities, rather than to local masking by two targets at the same disparity.

In addition to these two structural regularities, there seemed to be some idiosyncrasies in the masking behavior for the uncrossed (positive) test disparities. At a test disparity of 10 arc min, an additional minimum appeared at −10 arc min mask disparity. This feature is confirmed more weakly in the −20 arc min curve, but a symmetric dip did not appear in the crossed (negative) test disparity cases. Further characterization of this idiosyncratic dip was not possible due to the unavailability of this particular observer, however.

4.5. Polarity-specificity of disparity masking

The presence of two components found in the previous experiment (the broad component that is independent of the test disparity and the facilitatory/inhibitory component that tracks the test disparity) was tested further in the next experiment. For this case we probed whether these masking components operate only within the same luminance polarity or may influence the test probes of opposite polarity. The polarity of the masking stimulus was inverted to become a dark Gaussian blob in both eyes. The test stimulus was the same polarity as in the previous experiments. Thus, to the extent that the masking is determined by the (polarity-invariant) 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 and does not generate facilitatory behavior in test detectability, as shown in Fig. 9A, although the general effect of the polarity-inverted mask is still masking rather than its inverse, facilitation. In terms of the simplistic interpretative dichotomy, therefore, we can conclude that the generic component of masking that is test-disparity invariant is also polarity invariant; a net masking effect remains throughout the previous masking range. On the other hand, the batwing form of the test-disparity-selective component seems to be inverted to first approximation, as depicted in the conceptual model of Fig. 9B. In this variant, the plateau component of the masking model of Fig. 8 (right) is unchanged, but the wavelet component is inverted (and centered, to match

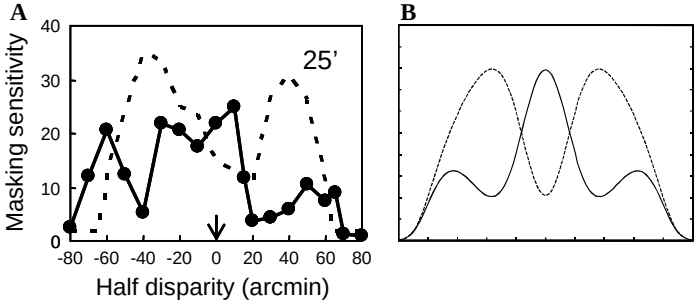

Fig. 9. (A) Comparison of masking by a dark 25 arc min bar (full curve) with that of a light 25 arc min bar (dashed curve, reproduced from Fig. 5 for test and mask with positive polarity). Note the inversion of the two main peaks and the central trough around a moderate masking level to generate two troughs flanking the central peak. Arrow at bottom indicates test disparity. Mean standard error was ±0.03 log units over the 3–5 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 stimulus disparity). It can be seen that this 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 arc min disparity now exhibit local minima on either side of the peak in the opposite-sign masking effect. Although it could not be said to be a clean inversion of the exact shape of the same-sign function, the data share many of the features of the conceptual model of a disparity-specific wavelet added to a polarity-invariant masking plateau (Fig. 9B). This initial assay reveals that a polarity-specificity masking mechanism exists and may be employed in evaluating the multi-component hypothesis of disparity masking. To establish this interpretation more definitively, however, one would have to measure the same-sign masking for both light and dark stimuli, and the opposite pair of opposite-sign stimuli, and do so as a function of both spatial-frequency and test disparity.

5. Discussion: The nature of disparity masking

The original motivation for this masking study, as implied in Section 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 in Fig. 3A conforms to the self-masking model, for example. However, the pattern of the present results makes it evident that the local masks reveal a complex structure of inhibitory interrelations among channels at different disparities. In particular, some of the inhibitory pathways are polarity-specific, others are polarity-independent.

The conceptual structure of the disparity interactions is illustrated in Fig. 10 for the case of the same-polarity mask varying in disparity around the test location. The panels depict a Keplerian array in the format of Fig. 1 with the visual axes for two eyes schematized as lying along to 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 (Fig. 10, first panel). The polarity-specific depth-inhibiting component produces two inhibition areas in front and behind the mask (second panel). The generic polarity-insensitive binocular component spreads broadly both across depth and lateral positions (third panel). These three components combined produce an inhibitory pattern of disparity interactions (fourth panel) similar to the experimentally measured behavior shown in Fig. 6. These three components would then vary with stimulus conditions to predict the masking behavior of all the other experiments. For example, varying the disparity of the mask would "scissor" the two components in the Monocular panel of Fig. 10 without affecting the locations of the Polarity-specific component, and vice-versa for variations in the disparity of the test. Polarity inversion of the mask would invert the profile of the Polarity-specific component in the sum of the three components and so on.

A putative explanation for the disinhibition near the test location in the polarity-specific component is suggested by the release from masking described by McKee et al. (1994), who found that dichoptic masking between a mask in one eye and a corresponding test in the other was almost eliminated by introducing an ancillary masking stimulus in the second eye. This stimulus provided a correspondence match to the mask in the first eye that moved it to a different perceived (or effective) location in space and reduced the original masking effect. If this monocular disinhibition plays any role in the masking ef-

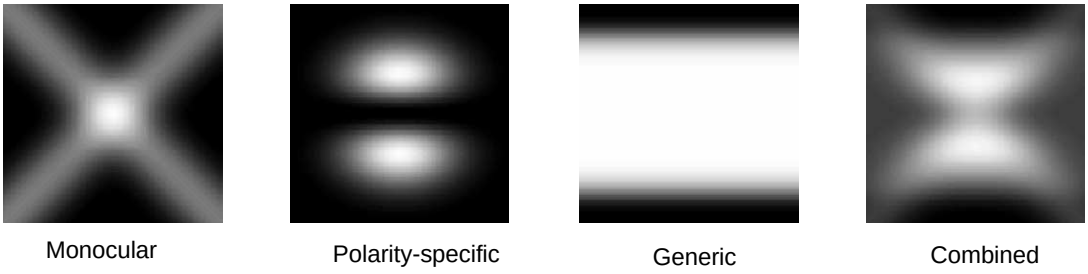

Fig. 10. Masking components in the Keplerian array coordinates of Fig. 1 (lateral position on the horizontal axis and near/far disparity on the vertical axis). The summed Monocular, Polarity-specific binocular, and Generic binocular components of masking are depicted separately in the respective panels. They combine to produce the signature "diaboloid" shape of masking in the Combined panel, which matches the masking pattern measured experimentally (see Fig. 6).

fects studied here, it would have to be in the disinhibitory lobes of the polarity-specific component, since all the other effects are masking (i.e., inhibitory). To apply this explanation to the present stimuli, one would have to suppose that they consisted of a transient version of the double-nail illusion (Kaufman, 1974) in which the two stimuli aligned in depth (on the z axis) have their monocular pairings reassign to be seen as two stimuli side-by-side, and are hence more easily detected. However, this explanation does not account for the pattern of masking obtained, since it would predict maximum masking when the test and mask were at the same disparity (which generates minimum masking in the presents data) and minimum masking when the polarity is reversed (which increases the masking in the present data).

The goal of the present study was to lay out the 4D space of disparity masking effects as a function of the lateral position and disparity of the mask, the disparity of the test, and their joint spatial-frequency content (width). (It may be noted that the full space of these disparity interactions is ten-dimensional, since it involves variations in the 2D position, disparity, width, and orientation separately for the test and mask, even before taking temporal factors into account.) We have therefore not extended the analysis to the issue of monocular masking effects. In terms of the spatial tuning, the present results are not directly compatible with the evaluation of sustained stereoscopic channels by Kontsevich and Tyler (1994). Based on data for interocular differences in spatial frequency and contrast, they found that there was no evidence for any spatial-frequency selective channels for stereopsis tuned to frequencies below 2.5 cycles per degree in the fovea. A similar conclusion was reached for transient stereoscopic stimuli by Schor, Edwards, and Pope (1998). Although the present masking behavior is found for spatial frequencies down to 0.45 cycles per degree in Fig. 6, this result is not contradictory to the earlier findings for two reasons. One is that the present study was conducted at 5° eccentricity, where the relevant positional acuity magnification function (Levi, Klein, & Aitsebaomo, 1985) would imply that the lowest channel might be as much as 5 times lower, or 0.5 cycles per degree. Moreover, the cited studies may be interpreted as reflecting excitatory channel structure, while the present masking results seem to derive predominantly from inhibitory interactions among channels. This interpretation would resolve any discrepancy between the results, although the main evidence for it is the pattern of the results in the two cases rather than any external considerations.

6. Conclusion

It is evident from the data that there is a large degree of disparity-domain masking that cannot be explained by the masking of its monocular constituents. This masking is, in turn, specific to the position, disparity, size (spatial frequency), and contrast polarity of the mask. 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. The masking range seems to have two disparity-specific components; one has a fixed disparity range and is polarity independent, while the other is keyed to both the disparity and polarity of the mask and has a complex center/surround form in the disparity domain. The structure of disparity interactions is thus found to be more elaborate than envisaged from previous studies, and is captured qualitatively by a three-component computational model. It is likely to be rewarding to build such masking behavior into future models of disparity encoding of depth information, for both computational and neurophysiological goals.

Acknowledgment

Supported by NEI Grant EY 7890.

References

Barlow, H. B., Blakemore, C., & Pettigrew, J. D. (1967). The neural mechanism of binocular depth discrimination. *Journal of Physiology (London), 193*, 327–342.

Blake, R., & Wilson, H. R. (1991). Neural models of stereoscopic vision. *Trends in Neuroscience, 14*, 445–452.

Cumming, B. G. (2002). An unexpected specialization for horizontal disparity in primate primary visual cortex. *Nature, 418*, 633–636.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: Masking experiments require a new model. *Journal of the Optical Society of America A, 11*, 1710–1719.

Graham, N., & Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: A comparison of single-channel and multiple-channel models. *Vision Research, 11*, 251–259.

Halpern, D. L., Wilson, H. R., & Blake, R. (1996). Stereopsis from interocular spatial frequency differences is not robust. *Vision Research, 36*, 2263–2270.

Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.

Julesz, B. (1978). Global stereopsis: Cooperative phenomena in stereoscopic depth perception. In R. Held, H. W. Leibowitz, & H.-L. Teuber (Eds.). *Handbook of sensory physiology, perception* (Vol. VII). Berlin: Springer-Verlag.

Kaufman, L. (1974). *Sight and mind*. Oxford UK: Oxford University Press.

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

Kontsevich, L. L., & Tyler, C. W. (1994). Analysis of stereothresholds for stimuli below 2.5 c/deg. *Vision Research, 34*, 2317–2329.

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research, 39*, 2729–2737.

Kontsevich, L. L., & Tyler, C. W. (2004). Local channel structure of sustained peripheral vision. In B. E. Rogowitz, T. N. Pappas (Eds.), *Human vision and electronic imaging IX*; Proc SPIE 5292, 26–33.

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

Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research, 25*, 963–978.

Mansfield, J. S., & Parker, A. J. (1993). An orientation-tuned component in the contrast masking of stereopsis. *Vision Research, 33*, 1535–1544.

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

McKee, S. P., Bravo, M. J., Taylor, D. G., & Legge, G. E. (1994). Stereo matching precedes dichoptic masking. *Vision Research, 34*, 1047–1060.

Ogle, K. N. (1950). *Researches in binocular vision*. London: Saunders.

Poggio, G. F., Motter, B. C., Squatrito, S., & Trotter, Y. (1985). Responses of neurons in visual cortex (V1 and V2) of the alert macaque to dynamic random-dot stereograms. *Vision Research, 25*, 397–406.

Rohaly, A. M., & Wilson, H. R. (1993). Nature of coarse-to-fine constraints on binocular fusion. *Journal of the Optical Society of America A, 10*, 2433–2441.

Rohaly, A. M., & Wilson, H. R. (1994). Disparity averaging across spatial scales. *Vision Research, 34*, 1315–1325.

Rohaly, A. M., & Wilson, H. R. (1998). The effects of contrast on perceived depth and depth discrimination. *Vision Research, 39*, 9–18.

Schor, C. M., Edwards, M., & Pope, D. R. (1998). Spatial-frequency and contrast tuning of the transient-stereopsis system. *Vision Research, 38*, 3057–3068.

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

Stiles, W. S. (1939). The directional sensitivity of the retina and the spectral sensitivities of the rods and cones. *Proceedings of the Royal Society London B, 127*, 64–105.

Stromeyer, C. F., & Julesz, B. (1972). Spatial-frequency masking in vision: Critical bands and spread of masking. *Journal of the Optical Society of America, 62*, 1221–1232.

Tyler, C. W. (1971). Stereoscopic depth movement: Two eyes less sensitive than one. *Science, 174*, 958–961.

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. *Perception & Psychophysics, 17*, 225–230.

Tyler, C. W. (1983). Sensory processing of binocular disparity. In C. Schor & K. J. Ciuffreda (Eds.), *Basic and clinical aspects of binocular vergence eye movements* (pp. 199–295). London: Butterworths.

Tyler, C. W. (1991). Cyclopean vision. In D. Regan (Ed.), *Vision and visual disorders (Vol. 9). Binocular vision* (pp. 38–74). New York: Macmillan.

Wilson, H. R., Blake, R., & Halpern, D. L. (1991). Coarse spatial scales constrain the range of binocular fusion on fine scales. *Journal of the Optical Society of America A, 8*, 229–236.