What is the Depth of a Sinusoidal Grating?

Suzanne P. McKee

Preeti Verghese

Bart Farell

Stereo matching of a wide textured surface is, in principle, ambiguous because of the quasi-repetitive nature of texture. Here, we used a perfectly repetitive texture, namely a sinusoidal grating, to examine human stereo matching for repetitive patterns. Our observers matched the depth of a vertical grating segment, 6 deg wide and presented in a rectangular envelope, at or near the disparity specified by the segment edges. Their depth matches were also influenced by the interocular phase of the carrier relative to the plane specified by the edges. The limiting disparity for the edge matches was 40 – 60 arcmin, independent of the spatial frequency of the carrier. One explanation for these results is that 1st-order disparity energy mechanisms, tuned to lower spatial frequencies, respond to the edge disparities, while showing little response to the interocular phase of the carrier. In principle, these 1st-order low frequency mechanisms could account for edge-based stereo matching at high contrasts. But, edge matching is also observed at carrier contrasts as low as 5%, where these low frequency mechanisms are unlikely to detect the grating stimulus. This result suggests that edge matching for gratings depends on coarse-scale 2nd-order stereo mechanisms, similar to the 2nd-order mechanisms that have been proposed for encoding two-dimensional texture. We conclude that stereo matching of gratings (or any other texture) depends on a combination of responses in both coarse-scale 2nd-order and fine-scale 1st-order disparity mechanisms.

Keywords: binocular vision, stereopsis, stereo matching, 2nd order mechanisms

INTRODUCTION

A sinusoidal grating of limited extent is perceived in depth at or near the disparity specified by its edges (Cumming & Parker, 2000). Indeed, the depth of any periodic stimulus, e.g. a row of regularly spaced dots or lines, is determined by the disparity of its endpoints (Mitchison & McKee, 1987ab; McKee and Mitchison, 1988). The elements of a periodic pattern can be potentially matched in a number of different depth planes. For example, as shown by Figure 1, the bars of a grating can be matched in the fixation plane, or forward or backward in depth. Apparently, to resolve this matching ambiguity, the stereo system selects the disparity of the outer edges, i.e., the disparity of the rectangular envelope of the segment, as the appropriate depth for the whole pattern. A grating is an artificial stimulus, but many natural scenes contain regions or surfaces covered with quasi-repetitive textures. Stereo matching based on
the edges of these patterns could resolve the ambiguities associated with repetitive textures. Here, we measured depth matching and stereoeacuity for grating segments to examine the stereo mechanisms responsible for edge matching.

In primate area V1, there are disparity-tuned neurons that respond strongly to the local disparity of the bars that make up the grating (carrier). According to the “disparity energy” model (Ohzawa, DeAngelis & Freeman, 1990; 1996; Cumming & DeAngelis, 2001), the interocular phase of the grating controls the disparity responses of these neurons. Thus, disparity neurons, tuned to the carrier frequency, respond to a grating shifted by one full period (diagrammed in Figure 1) as though the grating were perfectly matched in the fixation plane, i.e., with no shift. Why do humans and monkeys see this grating at the edge-defined disparity, when local disparity neurons are signaling that the grating is in the fixation plane? One answer is that these neurons are not signaling zero disparity, but rather zero phase, a signal that is consistent with many disparities. As Fleet, Wagner and Heeger (1996) have pointed out, because of their spatial frequency tuning, disparity-energy neurons give a response to quasi-periodic patterns, such as random dots, that is necessarily ambiguous. Disparities corresponding to any multiple of the neuron’s characteristic period (1/spatial frequency) will result in the same interoculular phase; thus, the response of these neurons oscillates as a function of disparity with a periodicity that is related to their spatial frequency tuning. There has been some controversy about whether disparity-tuned neurons respond to interocular phase differences or to interocular position differences. The current consensus is that many disparity-tuned neurons respond to both phase and position differences. Fleet et al. (1996) showed that both phase and position disparity neurons would generate an oscillating response as a function of texture disparity, because this oscillation is generated by the neuron’s spatial frequency tuning, not by whether disparity is encoded by phase or position. Neurons tuned to low spatial frequencies show fewer oscillations within a given disparity range than neurons tuned to high spatial frequencies, because their period is larger. In this sense, the ‘coarse’ scale response is less ambiguous than the ‘fine’ scale response, but it is still ambiguous. The stereo system can only resolve this ambiguity by combining disparity responses across scales in some fashion.

Fleet et al. (1996) suggested that pooling over scales could produce a neural response uniquely tuned to a particular disparity. Suppose that, at every scale, there are detectors specifically tuned to a particular disparity, e.g. 4 arcmin, which corresponds to different phases at different scales. When the stimulus disparity equals this tuned disparity, all scales will respond. Each detector will also respond periodically to other stimulus disparities, but the periodic response as a function of disparity will be different at every scale. So, if the responses across scales are pooled, the peaks and troughs will typically cancel at all disparities except the tuned disparity. Does pooling over scales explain edge matching for an extended grating segment? To answer that question, we first must consider what kind of mechanism responds to the edge disparities.

The edge disparities may be encoded by primary (1st-order) disparity energy mechanisms tuned to lower spatial frequencies. Spatial mechanisms in primate vision are known to be broadband in their spatial frequency tuning, responding to frequencies covering about 1.5 octaves. At high contrasts, a grating segment generates a substantial response in mechanisms tuned to spatial frequencies as much as one octave lower than the carrier frequency. A disparity mechanism tuned to this lower frequency responds strongly to the edges, and weakly to the center of the grating itself. But the response of the mechanism tuned to the carrier frequency is typically much stronger than that of the mechanism responding to the edges, even in the spatial neighborhood of the edges. Thus, the lower frequency responses will not dominate the pool unless they are heavily weighted. Even if these responses are weighted to increase their contribution to the pool, edge matching should not occur at low contrasts because the responses of the low spatial frequency mechanisms would be sub-threshold. No matter how the low frequency responses are weighted, they cannot contribute to the pool unless they exceed the internal noise level.

The pooling model proposed by Fleet et al. (1996) is a kind of coincidence detector, a simple and elegant scheme for choosing the disparity where all scales show a response peak. To predict edge matching by pooling across scales, one must also set some plausible limits on the spatial extent of the pool. Responses might be pooled over the receptive field of the largest 1st-order mechanism with a significant signal, i.e., the mechanism responding to the edge disparity. The difficulty with this approach is that spatial pooling across scales, if obligatory, sacrifices the stereo information that permits the perception of surface relief and transparency. Instead of an obligatory combination, the stereo system may simply use the distributed pattern of responses across scales to estimate the most likely disparity at each spatial location. The Tsai and Victor (2003) template-matching model uses this population-response approach and can predict the perceived depth of the grating edges; however, it is unable to predict the depth at the center of the grating segment (Tsai, personal communication).

Rather than a low frequency 1st-order stereo mechanism, the edge disparities could be mediated by responses in the “2nd-order” stereo mechanism that has received so much attention in the past decade (Hess and Wilcox, 1994; Wilcox and Hess, 1995; 1996; 1997; Schor, Edwards and Pope, 1998; Edwards, Pope and Schor, 1999; Langley, Fleet and Hibbard, 1998; 1999). It is assumed that this putative mechanism responds to the
disparity of the stimulus envelope, rather than to the disparity of features within the envelope. The evidence for this mechanism comes from studies showing that observers can combine half-images having opposite polarity (Pope, Edwards & Schor, 1999), different spatial frequencies (Langley, Fleet and Hibbard; 1998), orthogonal orientations (Schor, Edwards and Sato, 2001), uncorrelated random carriers (Elder & Wilcox, 2000) and envelopes of different sizes (Schor et al., 2001). Like the 2nd-order systems for texture and motion, the non-linearity that produces the envelope signal is thought to be a type of neural rectification (Elder and Wilcox, 2000; Schor et al., 2001). Langley et al (1999) have argued persuasively that this non-linearity is cortical, and Wilcox and Hess (1996) have provided evidence that it occurs before the monocular signals are combined binocularly.

As an explanation for edge matching, this 2nd-order mechanism has a distinct advantage. Full or partial rectification would enhance the strength of the edge signals. If the rectified monocular signals were subsequently combined by a conventional disparity energy mechanism tuned to a low spatial frequency (<1.0 cpd), the response to the edge disparity could be significant even at low contrast contrasts, and would extend over a substantial spatial extent.

In this study, we shall explore the limitations of edge matching for gratings of various spatial frequencies and contrast levels. We shall also examine whether the luminance modulation at the carrier frequency, i.e., the bars of the grating itself, make any contribution to the perceived depth of the pattern or whether its disparity signal is completely overridden by the mechanism that reads the disparity of the edges. Our primary objective, however, is to identify the nature of the edge-disparity mechanism and to explore how it operates to control the depth plane of an extended grating or surface.

**METHODS**

**The Depth Matching Procedure**

Observers viewed the stereoscopic half-images of a grating target presented in mirror stereoscope. Each half-image consisted of a segment of a sinusoidal grating, 6 degrees wide and two degrees high. Except where indicated, both half-images of the carrier were presented with 0 deg phase (sine phase) coinciding with the edges, as shown in Figure 2B, to preclude an abrupt transition in luminance from the background to the grating. For this study, the ‘edges’ of the grating were simply the position where the grating began or ended abruptly on the screen, i.e. each half-image was presented in a rectangular envelope. No additional aperture or frame surrounded the grating. The edge of the grating was at least 2.8 degrees away from the nearest edge of the screen.

A thin dark probe, 30 arcmin in length and positioned 1 deg below the horizontal midpoint of the grating, was presented with one of seven disparities (Figure 2A). On each trial, the observer judged whether the grating was in front or behind the probe. No feedback was given. The range of probe disparities was chosen to bracket the anticipated depth of the grating based on preliminary test runs. We used large disparity steps to guarantee that we could obtain a complete psychometric function centered on or near the grating depth. The smallest step, used for the fixation plane measurements, was 1.4 arcmin, but typically, much larger step sizes
were used for other depth planes. Occasionally, the chosen range was far from the perceived grating depth; then, the observer stopped the test run, and re-centered the probe so that the grating appeared within the seven-disparity probe range. Most of the psychometric functions shown in the graphs are based on 100 trials taken in a single experimental run, with 4 practice trials. When the perceived depth of the grating was unstable, additional runs with different probe ranges were added to the first run.

We were somewhat concerned that the probe itself might influence the depth of the grating (Zhang, Edwards and Schor, 2001) so, for a few conditions, we repeated the measurements with the whole probe range shifted by one large disparity step. The resulting psychometric functions for this shifted range could be superimposed on those obtained with the initial probe range, showing that the probe did not affect the depth of these high contrast gratings.

To control for vergence shifts, we asked the observer to align nonius lines before initiating a trial. The nonius lines were separated by two degrees vertically and presented above and below the midpoint of the screen; a small dark fixation point was presented between the nonius lines to assist the observer with alignment. The nonius lines and the fixation point were turned off during the test trial. During a test trial, the horizontal grating was presented 30 arcmin above the vertical midpoint of screen, with the test probe displayed 1 deg below the middle of the grating. The duration of the test trial was 200 msec, a time too brief to permit the completion of a voluntary vergence movement for these midline targets (Rashbass & Westheimer, 1956).

Some previous studies have noted changes in the perceived depth of periodic patterns at longer durations. For example, McKee & Mitchison (1988) noted that a sparse row of widely spaced dots changed from an initial depth match at the disparity of the edges to a match in the fixation plane at longer durations (>1 second). To check whether the depth of the gratings was similarly unstable, we added a continuously visible fixation point (where the observer was asked to hold fixation) and repeated some depth matches for gratings of 3 and 6 cpd at durations of 500 and 1000 msec. The edge-matching pattern was identical to that observed for the shorter 200 msec duration.

Disparity Threshold Procedures

In addition to depth matching, we also measured stereoacuity for these grating targets. In a typical stereoacuity paradigm, the test and reference stimuli are presented in nearly the same depth plane; this plane may be the fixation plane, or a depth plane other than the fixation plane (left side of Figure 2C). The disparity of the test stimulus is changed from trial-to-trial in small increments centered on the disparity of the reference stimulus, and the observer judges whether the test stimulus is in front or behind the reference stimulus. Of course, the labels ‘test’ or ‘reference’ stimulus are totally arbitrary, and merely mean that the disparity of one target (the reference) is constant over trials, while the other (the test) varies from trial-to-trial. For the experiments shown in Figure 6, the probe line served as the test stimulus, and the disparity of the grating was fixed. For the experiments in Figure 7, the disparity of the probe was fixed and the interocular phase of the carrier was varied in small increments from trial-to-trial. We chose this latter arrangement to examine the contribution of the interocular phase of the carrier to stereoacuity.

We used two different paradigms to measure stereo sensitivity. The first was the stereoacuity paradigm described above. For the stereoacuity measurements, we presented the grating immediately above the probe in or near the same depth plane (left side of Figure 2C). In the other paradigm, we presented the probe with a mean standing disparity separating it from the grating -- an arrangement we call ‘depth interval’ (right side of Figure 2C). For the depth interval discriminations, the observer judged small incremental changes in the standing disparity separating the targets, i.e., whether the depth between the grating and the probe was larger or smaller than their mean separation in depth. Previous studies have shown that depth interval judgments are generally more difficult than stereoacuity judgments made with test and reference in nearly the same plane (McKee, Levi, & Bowne, 1990; Vreven, McKee & Verghese, 2002); the configuration on the left side of 2C produces better sensitivity than the configuration on the right side of 2C. For all threshold measurements, the observer was given feedback and practice trials to stabilize her estimate of the mean and range of the test set. Note that the incremental disparity steps used to estimate sensitivity were generally much smaller, e.g. 0.2 arcmin, than the disparity steps used for the depth matching measurements.

From all threshold procedures, we obtained a psychometric curve based on the proportion judged ‘larger’ than the mean depth as a function of the probe or phase disparity. A cumulative normal function was fitted to the data with probit analysis to determine the change in disparity that produced a change in response rate from the 50% to the 75% level, equal to a d’ of 0.675. Each threshold shown in figures 6 and 7 is based on a minimum of 200 trials.

General Stimulus Arrangements

The displays were programmed on a Macintosh computer and displayed on the screens of two 14” Sony Trinitron monitors, model 110GS. We used only the central 6 deg of each screen, where screen curvature was slight, for our displays. Nevertheless, we assessed the effect of screen curvature on visual direction from the
perspective of the observer’s head position. We mounted a flat transparent grid of horizontal and vertical lines immediately in front of each screen and then aligned a thin cursor by eye with each grid intersection for each screen separately. The cursor alignment measurements were incorporated into the software to correct each screen for any systematic deviations in x-y locations induced by screen curvature or raster imperfections. The monitors were run at 75 Hz, using a 1024 x 768 resolution level. At our viewing distance of 1.22 meters, each pixel subtended 0.71 arcmin. The grating displays were created by dithering adjacent pixels to produce a sinusoidal luminance variation on the two monitors. We also used dithering to create sub-pixel shifts in the location of the probe target.

For all studies except the explicit study on the effect of contrast on stereo matching, the contrast of the grating targets was 50%. The gratings were presented against a homogeneous background equal to the mean luminance of the grating, which was 68 c/\(\text{m}^2\). A Pritchard photometer was used to measure luminance as a function of poke value, and these values in turn were used in the contrast and dithering calculations. Overhead fluorescent lights illuminated the apparatus and the surrounding room at photopic levels.

**Observers**

The observers in these experiments were two of the authors and three young adult female volunteers. For each observer, we explained the purpose and methods used in this study, and then obtained a signed consent form. All five were experienced psychophysical observers with good stereoacuity. They wore optical corrections as needed for image clarity.
Figure 3. Psychometric functions showing depth matching judgments for five different observers. Each curve shows the percentage of trials on which the grating segment was seen in front in the probe as a function of probe disparity. The colors indicate different edge disparities for the range covering ± 1 period. The midpoints of these functions indicate that observers matched the grating at or near the disparity of the edges. Grating is 3cpd, 6 deg wide and 2 deg high. Duration = 200ms; Contrast = 50%

Results

Edge-Based Matching of Extended Sinusoidal Gratings

In our first experiment, observers matched the depth of 6 deg wide, 50% contrast, 3cpd gratings whose edges were offset over a range of ± 1 period (± 20 arcmin). The separate graphs in Figure 3 show the data from each of the five observers. Each curve gives the percentage of trials on which the observer judged that the grating was in front of the test probe as a function of probe position; note that the probe steps were very large (minimally 1.4 arcmin) to ensure that the perceived depth of the grating fell within the disparity range of the probe. The different colors of the curves correspond to different edge disparities. For example, the red curves show the data generated by gratings that have been shifted by one full period so that the edge disparity is 20 arcmin and crossed. As we noted above, the “edges” refer merely to the starting and ending positions of the grating half-images; the grating segment is presented in a rectangular envelope.

The psychometric functions in Figure 3 provide striking evidence that, at 3cpd, an extended grating is matched at or near the disparity specified by its edges. For each function, the probe disparity corresponding to the 50% point (the ‘Point of Subjective Equality’) falls close to the disparity of the edges. Edge matching is not confined to integer multiples of the grating period; observers match the grating to edge disparities that correspond to fractions of a period. As one example, the 50% point of the purple curves lie at or very near 10 arcmin behind the fixation plane, i.e., at the disparity specified by the edges. If the grating were so wide that the peripherally viewed edges were indistinct or invisible, this 10 min shift in the grating, corresponding to a 180-degree phase shift, would produce an unstable percept that would appear either forward and back of the fixation plane at random from trial-to-trial. The signal generated by the edges (or envelope) resolves this ambiguity so that a grating shifted by 180 degrees is seen at (or very near) the same position in depth on every trial.

The same pattern of edge matching is found at higher spatial frequencies. In Figure 4, the graphs on the left show the data from three observers for a 6 deg wide grating of 6 cpd, while the graphs on the right show the data for a grating of 10 cpd. High spatial frequency gratings shifted by one full period are matched near the disparity of the edges.

Figure 4. Psychometric functions showing depth matching judgments at higher spatial frequencies (6 cpd on the left and 10 cpd on the right) for three observers. The colors indicate different edge disparities for the range covering ± 1 period. Grating is 6 deg wide, 2 deg high. Duration= 200ms; Contrast = 50%.

Stereo matching based on the disparity of the edges breaks down at low spatial frequencies, although the exact limit depends on the observer (Figure 5). Observer MPM matched the 2-cpd grating, with an uncrossed edge disparity, either in the fixation plane or in the plane specified by the edges, indicating that her response to the edges is too weak to determine carrier location consistently. Observer SPM matched the 1 cpd grating, shifted by 1 period, in the fixation plane, while observer PV saw the 1 cpd grating at depths off the fixation plane, but short of the edge disparities. As shown, observer DV...
matched the 0.75 cpd grating, shifted by 1 period, in the fixation plane; her matches (not shown) for 1 cpd grating, shifted by 1 period, were firmly in the plane of the edges (± 60 arcmin).

interocul phase disparity of the carrier have any effect on perceived depth at moderate spatial frequencies (>2 cpd)?

Figure 5. Psychometric functions showing depth matching judgments at low spatial frequencies for four observers (2cpd for MHM; 1cpd for SPM and PV; 0.75cpd for DV). Edge-based matching tends to fail at low spatial frequencies (2 cpd or lower). Arrows show location of edges. Grating is 6 deg wide, 2 deg high. Duration = 200 msec; Contrast = 50%.

One explanation for the difference in edge matching between low and moderate spatial frequencies is that the human stereo system cannot process frequencies higher than ~ 1 cpd. If so, a high spatial frequency grating would be appear much like a fuzzy homogeneous blob to the stereo system, and would naturally be matched at the disparity of its edges. This explanation is countered by data showing an improvement in stereoaucity for spatial frequencies above 1 cpd (Schor & Wood, 1983; Legge & Gu, 1989; Hess & Wilcox, 1994). Alternatively, the stereo system may respond to the carrier, but as the number of cycles in the grating increases with the increasing spatial frequency, it may be unable to resolve the matching ambiguity diagrammed in Figure 1 and so may assign a depth to the grating segment that is consistent with the unambiguous match associated with the edges (Hess & Wilcox, 1994; Prince & Eagle, 2000). Of course, this ambiguity also exists for a 1 cpd grating (6 cycles in a 6 deg window), but to a lesser extent. These considerations raise an interesting question: does the

What is the contribution of the carrier to depth processing?

For these experiments, we measured stereo sensitivity, rather than perceived depth, to assess the role of the carrier in stereo matching. We measured stereo sensitivity for two conditions: 1) a homogeneous luminous box equal in extent to the grating target and presented in the fixation plane; 2) a 3cpd grating with an interocular phase disparity of 144 deg, equal to a crossed disparity of -8 arcmin. As diagrammed in Figure 6, the edges were presented in the fixation plane with zero disparity for both conditions (see ‘edge outline’ in diagram). A
black test probe was presented at a mean crossed disparity of -8 arcmin, in front of the fixation plane. From trial-to-trial, we varied the disparity of the probe backward and forward in small increments, and asked the observer to judge whether the disparity separating the probe from the fixation plane was larger or smaller than the mean disparity (-8 arcmin). In the first condition, the observer had to judge the depth interval separating the probe from the zero-disparity box. In the second condition, the grating phase specified a shift forward to the mean position of the probe, so in principle, the carrier could act as an adjacent reference and improve stereo sensitivity for the probe (right diagram in Figure 6). As we noted in the methods section, depth interval judgments are generally more difficult than stereoacuity judgments made with a test and a reference target presented in the same plane.

If the stereo system cannot process the 3cpd carrier or is unable to resolve the matching ambiguity of the carrier, both conditions should produce about the same sensitivity. However, if the shift in the interocular phase of the grating changed the apparent depth of the grating, centering it above the mean position of probe, the offset grating should yield lower thresholds. We found that stereo judgments for the second condition with the grating phase shifted to a crossed disparity of 144 deg (red columns) were more precise, by about a factor of two, than the depth interval judgments with the box in the fixation plane (blue columns). For completeness, we also measured depth interval thresholds with a grating, instead of the luminous box, presented in the fixation plane (no phase shift). The thresholds for the grating in the fixation plane were identical to the thresholds for the box in the fixation plane. These results indicate that the interocular phase disparity of the carrier contributes to stereo processing for extended grating segments. From the picture in Figure 6, it might appear that the edges are separate from the grating. But there is no separation: the edges are only the starting and ending positions of the grating. Changes in the interocular phase of the carrier, with no change in edge position, necessarily alters the luminance profile at the edges. In
the left half of Figure 7, we have diagrammed two ways to produce the same phase shift in the carrier. In the upper diagram, the two half-images start at the same position, but the starting phase of one half-image is shifted by 135 degrees, producing an interocular phase difference of 135 degrees. In the lower diagram, the starting position of the grating is shifted by 7.5 arcmin in one half-image, also producing an interocular phase difference of 135 degrees in the 3 cpd carrier as well as an edge disparity of 7.5 arcmin. Except for the first half-cycle, these two configurations are identical. We wondered if this subtle difference in the edges would affect thresholds for depth interval judgments measured at small pedestals (<180 degrees of interocular carrier phase). For this set of experiments, the probe position was fixed at zero disparity, and the pedestal disparity of the carrier was varied parametrically to create a depth interval between the probe and grating. Note that in this experiment, a pedestal shift moved the grating away from the plane of the probe, whereas in the experiment shown in Figure 6, a pedestal disparity moved the grating into the same depth plane as the probe. For any block of trials, this pedestal disparity was constant. We introduced small incremental changes in the phase disparity from trial to trial, and asked the observer to judge whether the depth separating the grating from the fixed probe was larger or smaller than the mean depth interval. For the phase shift data, the edge (envelope) disparity was zero and the interocular phase of the carrier was shifted to create a pedestal disparity (upper diagram on left of Figure 7). For the edge shift data, the edges were shifted to the pedestal disparity (lower diagram on left of Figure 7), but there were no trial-to-trial variations in the edge disparities; the edges (envelope) remained fixed at the pedestal disparity while the interocular phase of the carrier was varied in small incremental steps around the disparity of the edges.

The data shown in the graph in the lower right of Figure 7 reveal that this difference in the edge configuration is quite important in depth interval judgments. When the edges of the half-images are fixed at zero disparity (“phase shift” condition), increment thresholds rise precipitously once the pedestal phase exceeds about 60 degrees (open circles). In their measurements of phase increment thresholds, Farell, Li and McKee (2003ab) observed exactly the same accelerating pattern for gratings presented in a Gaussian envelope (envelope disparity = zero). On the other hand, the ‘edge shift’ condition (filled squares) shows a much slower rise in the thresholds. The difference in the thresholds for these two edge configurations is about 0.5 log unit at an interocular phase of 90 degrees. Consistency between the disparity of the edges (envelope) and the pedestal disparity specified by the interocular phase of the carrier enhances sensitivity in this depth interval task.

The Upper Limit for Edge Matching of Extended Gratings

These stereoacuity results show that the interocular phase disparity of the carrier influences the perceived depth of grating segments at moderate spatial frequencies. The edge disparity matches shown in Figures 3 and 4 cannot be attributed to the inability of the stereo system to respond to these frequencies. What then accounts for the differences between the matches at moderate spatial frequencies, and the matches at low spatial frequencies (Figure 5)? Why do the edges (envelope) of the grating determine the depth at moderate spatial frequencies, while the interocular phase of the carrier appears to have more influence at low spatial frequencies?

In the introduction, we speculated that the edges (envelope) might be processed by a stereo mechanism tuned to a lower spatial frequency than the one that responds to the center of the grating. When different scales process the edges and the center, the scales can be adjusted differentially; for example, lower spatial frequencies can be more highly weighted (amplified) in the

![Figure 8. Psychometric functions showing depth matching judgments for a 1 cpd grating presented at an edge disparity of 30 arcmin, equal to half-period of carrier. Compare these functions to those in Figure 5. Edge matching is not limited by the carrier frequency, but rather by the size of the edge disparity. Arrows show location of edges. Grating is 6 deg wide, 2 deg high. Duration= 200msec; Contrast = 50%.](image-url)
neural calculation of stereo matching. However, when the carrier is itself a low spatial frequency, both the edge and carrier disparity may be processed by a mechanism tuned to the same spatial scale. Given that disparity detectors respond only to fairly small regions of visual space, the local detectors that respond to the edges could signal one disparity, while those that respond to the center of the grating could signal another. Why would the stereo system selectively weight the responses from one location more than the responses from another, when both detectors are tuned to the same spatial scale? If each location were processed independently, one would expect that the grating segment would appear bowed in depth, which is not the case. Instead, the center of the low frequency grating segment, shifted by one period, is stably matched in the fixation plane (Figure 5, observers DV, SPM), alternates in depth between the fixation plane and the plane defined by the edges (Figure 5, observer MHM) or appears to lie at a position between the planes (Figure 5, observer PV).

There is an alternative explanation for the low spatial frequency matches. The mechanism that responds to the grating envelope may have an upper disparity limit of roughly 60 arcmin, with the exact value varying from subject to subject. If so, a low frequency grating would be matched at the depth corresponding to the edge disparity provided that the edge disparity was below this limit. We found evidence supporting this second alternative -- a 1 cpd grating was firmly and stably matched at an edge-specified disparity of ± 30 arcmin (see Figure 8). There is nothing special about the 180 deg phase shift. Observer SPM matched a 1-cpd grating shifted at the edges by 45 arcmin (270 deg of phase) also at the disparity of the edges.

We also measured the perceived depth of higher spatial frequency gratings (3 and 6 cpd) at edge disparities corresponding to multiple periods of the carrier. Observers PV and SPM matched these higher frequency gratings at the depths corresponding to the edges up to a disparity of approximately 45 arcmin, (equal to 2.25 and 4.5 periods of 3 and 6 cpd respectively). Beyond that disparity, either the gratings were always matched at 40-50 arcmin independently of the edge disparity, or appeared at a distant depth so indeterminate that no stable match could be made. Note, however, that there was a difference between the matches made to high frequency and low frequency grating segments: at large edge disparities, the high frequency gratings were never seen in the fixation plane.

Broadly speaking, our results suggest that edge matching for gratings or other textured surfaces is limited to a range of about ± 60 arcmin in the central visual field. Similar limitations have been noted in other studies of extended gratings or random dot displays (Wilcox & Hess, 1995; Prince & Eagle, 1999; McKee, Watamaniuk, Harris, Smallman & Taylor, 1997). In their measurements of disparity sensitive neurons in area V1, Prince, Cumming & Parker (2002) found that few neurons had a preferred disparity greater than ± 1 deg in the central visual field of their awake-behaving primates. The upper limit of depth matching observed here may reflect limitations imposed by the range of these neurons.

A 2nd-order Mechanism to encode edge disparities

Because of its rectangular envelope, a grating segment presented at a high contrast level produces a substantial signal at spatial frequencies both higher and lower than the carrier frequency. The pointed green curve in the upper left graph in Figure 9 shows the smoothed amplitude spectrum for a 3cpd grating, 6 deg wide, presented in a rectangular envelope. The superimposed red curve shows the tuning function for a Gabor filter, with a frequency bandwidth of 1.5 octaves, tuned to the carrier frequency (3cpd). A disparity mechanism with this spatial frequency tuning would integrate the contrast energy of the target that fell within its broad bandwidth. The blue curve shows a filter tuned to a frequency an octave below the carrier (1.5 cpd). By examining the overlap between the stimulus (green curve) and each of these filters, one can determine whether a detector with this tuning could respond to the stimulus.

As is apparent from inspection of the blue and green curves, the lower frequency filter is quite insensitive to the carrier frequency (3cpd), but is strongly responsive to the integrated signal produced by the envelope sidebands in the low spatial frequency range. A disparity mechanism with this tuning function might be responsible for edge matching, because it responds strongly to frequency components introduced by the envelope and weakly to the carrier frequency. To make this point clearly, we have convolved the grating segment with these two filters to show the relative response amplitude after filtering from one edge of the grating to the other (upper middle graph in Figure 9.)

What happens to the signal produced by the sidebands if the grating contrast is reduced? As is apparent from the convolution, the mechanisms tuned to 3 cpd generally have a stronger response than the mechanisms tuned to 1.5 cpd. Thus, at sufficiently low contrasts, the 1.5 cpd mechanism would not detect the edges of the grating consistently above the internal noise level of the human visual system, at levels where the 3-cpd mechanisms could still detect the carrier. If the envelope (edge) disparity is not detected, observers should match the grating to the disparity specified by the carrier. A grating offset by one full period should either appear matched in the fixation plane, or at some randomly selected multiple of the period.

In our next experiment, we measured stereo matching for the extended grating as a function of contrast. As shown by the lower two graphs in Figure 9, observers
consistently match a grating offset by one full period (+20 arcmin) at the disparity of the edges, even when the grating contrast is reduced to 5%.

Can the 1.5-cpd mechanism detect this grating at 5% contrast above the internal noise of human visual system? It is difficult to estimate the physiological response of this hypothetical mechanism, but we can estimate an upper bound of the response. Given the assumption that the initial operation prior to the physiological calculation of disparity energy is similar to linear filtering, we can multiply, in the Fourier domain, the filters shown in the upper left graph of Figure 9 with the stimulus – the grating at 5% contrast. The graph on the upper right of Figure 9 shows the result of this multiplication. The dotted arrow points to the lowest contrast threshold measured in the fovea for an extended grating, equal a value of about 0.3 % (e.g. Foley, 1994). Most studies show that the best human contrast sensitivity is found at a spatial frequency of 3cpd (Graham, 1989), so this dotted arrow also marks the contrast threshold for our grating target. The function showing the hypothetical amplitude of the 1.5 cpd mechanism to the 5% grating lies completely below the dotted line. It seems unlikely that the integrated response of this mechanism would exceed the noise of human visual system. For one thing, this curve shows the maximum estimated amplitude. We have not scaled the two filters shown in the left graph (Figure 9) by the contrast sensitivity function. Typically, the sensitivity at 1.5 cpd is somewhat less than the sensitivity at 3 cpd, so the amplitude of the 1.5 cpd mechanism could be justifiably reduced. We are also assuming that the filtering process does not reduce the amplitude or increase the noise in this mechanism – two effects that would reduce the signal/noise ratio. Given these considerations, we conclude that at 5% contrast, this mechanism does not detect the grating.

It may not be obvious from the convolution shown in the middle graph that the response amplitude of the 1.5 cpd filter is sub-threshold. The triphasic response at the edges (blue curve) appears to be about 40% of the response amplitude of the 3cpd filter to the center of the grating (red curve). If the grating segment were itself 10 times contrast threshold, e.g. 5%, wouldn’t this triphasic edge response be at least 4 times contrast threshold? Probably not. Keep in mind that the edges of the grating segment fall at an eccentricity of 3 degrees and that the triphasic response is limited in spatial extent. The threshold shown by the dotted arrow is for an extended grating. It is known that the threshold for a narrow grating, e.g., 2 periods, is substantially higher than for a wider grating, e.g., 12 periods, of the same frequency. Any comparison between the human threshold for a real target and a hypothetical response from a biological filter is inherently risky. Nevertheless, the response amplitude of the 1.5-cpd filter to the edges is unlikely to be detected when the grating contrast is 5%. A mechanism tuned to a frequency nearer to the carrier frequency, e.g. 2cpd, would have a better signal, but it would resemble a reduced version of the response of the 3cpd filter to the grating segment – the response to the center would be greater than to the edges.

Since it is implausible that any 1st-order disparity mechanism could be responsible for the edge matches at 5% contrast, a different mechanism is required to explain these observations. It is also implausible that the stereo system would use one mechanism for edge matching at high contrast and a different mechanism at low contrast. Therefore, we conclude that 1st-order mechanisms do not account for edge matching at any contrast level.

The real issue here is not whether 1st-order mechanisms can identify the disparity of the edges at high contrasts, but rather why the stereo system would override 1st-order mechanisms responding to the center of the grating, in favor of 1st-order mechanisms responding to the edges 3 deg away. Contemporary models of stereo matching combine responses from 1st-order mechanisms tuned to different spatial scales to assign a unique disparity to a circumscribed spatial location. The Tsai-Victor model can predict the edge disparities of the grating, but it does not assign these disparities to the grating center – contrary to the perceived depth reported by our subjects.

The type of stereo mechanism required to explain our edge matching either must respond to a substantial spatial region, e.g. be tuned to a very low spatial frequency, or must operate functionally in a different way from the 1st-order mechanisms, e.g. be able to influence the responses at distant locations. In fact, it may need both properties. We propose that 2nd-order disparity mechanisms that respond to the envelope of the grating are responsible for our results. The general structure of all 2nd-order mechanisms, such as those that have been proposed for texture and motion, is essentially the same: initial filtering by a mechanism tuned to moderate or high spatial frequencies, followed by a nonlinearity such as rectification, and then further filtering of the nonlinear signal by a low spatial frequency mechanism (Bergen & Landy, 1991; Graham, 1991; Sutter, Sperling & Chubb, 1995; Graham & Sutter, 1998). In this case, we speculate that the non-linear monocular signals are processed by disparity energy mechanisms tuned to low spatial frequencies, following the same set of operations that has been suggested for 2nd-order motion detectors (Wilson, 1994). A flow chart of these operations is shown in Figure 10A. Note that the initial filtering is the same for both 1st- and 2nd-order mechanisms, so, in principle, if a stimulus is above threshold for the 1st-order mechanism, it will also above threshold for a 2nd-order mechanism.
Figure 9. The upper left graph shows Fourier transforms of a 3cpd, 6deg wide grating in a rectangular envelope (green curve), and two spatial mechanisms represented by Gabor functions tuned to 3 cpd (red curve) or 1.5 cpd (blue curve). In the upper middle graph, the two filters have been convolved with the grating to show the relative response amplitude after filtering as a function of spatial position. In the upper right graph, the two filters have been multiplied, in the Fourier domain, with a grating stimulus of 5% contrast. The data in the lower two graphs show psychometric functions for depth matching judgments for a 3cpd grating presented at one of three contrast levels. The edges of the grating have been shifted forward or backward by one full period. Even at a contrast of 5%, the grating is matched at a disparity corresponding to the disparity of the edges. Grating is 6 deg wide, 2 deg high. Duration= 200msec.

In the introduction, we described the abundant evidence from other psychophysical studies for 2nd-order mechanisms in human stereopsis. There is also neurophysiological evidence that is consistent with the existence of 2nd-order disparity mechanisms in primate stereopsis. Read & Cumming (2003) have noted that the disparity frequency, i.e., response oscillations as a function of the disparity of a quasi-repetitive pattern, of many binocular cells is shifted to a coarser spatial scale than predicted from the spatial frequency tuning of their monocular components. In addition, the response of the neurons to anticorrelated stereograms is not consistent with disparity energy models. Read & Cumming suggested that a non-linearity (thresholding of monocular components) prior to binocular combination could account for both these observations (see also Read, Cumming & Parker, 2002), and we would add, edge-matching as well.

DISCUSSION

As has been recognized for a quarter of a century (Marr & Poggio, 1979), the responses of fine-scale disparity mechanisms to the quasi-repetitive structure of natural textures are ambiguous, specifying multiple depth planes. The standard solution for resolving this ambiguity is some combination of responses across scales – so-called ‘coarse-to-fine’ matching. Here we have used a perfectly periodic texture, a sinusoidal grating, to examine stereo matching for repetitive patterns. The disparities at the ‘edges’—the starting and ending positions of the grating—determine the matching depth plane for the whole grating. Nevertheless, the interocular phase of the grating also influences perceived depth. The edges define the depth plane, but the phase disparity of the carrier modulates the depth with respect to this plane. Generally, the coarse- and fine- scales play complementary roles in specifying surface structure. The coarse edge
order disparity mechanisms. B) Convolution of 1st- and 2nd-order disparity mechanisms with one half-image of 3cpd, 6 deg wide grating showing response as a function of space, prior to calculation of disparity energy. C) Diagram showing depth interpolation of fine-scale 1st-order responses with respect to plane defined by edge disparities that are encoded by 2nd-order “depth texture” mechanisms (shown by thick dotted line). The thin dotted lines are meant to indicate ‘ghost’ planes – alternative planes consistent with the phase disparity of the 1st–order mechanisms.

mechanisms extend the depth range of the mechanisms that respond to a finely textured surface, but they are necessarily imprecise, and are thus unable to supply reliable stereo information about fine surface relief (Wilcox, 1999). The fine scales can supply this information, because, in principle, their responses to the disparity of the texture elements need not be affected by the standing disparity at the edges.

Our results indicate that the edge matches depend on 2nd-order stereo mechanisms, similar to mechanisms that have been proposed for two-dimensional texture segmentation. Recently, Stelmach and Buckthought (2003) proposed a similar 2nd-order process to disambiguate the depth of noise patterns at large disparities. They imposed a contrast-modulated envelope on the noise and varied the disparity of the envelope and noise (carrier) separately. Like our results with the gratings, they found that perceived depth depended on the disparities of both the 2nd-order envelope and the noise carrier.

The 2nd-order disparity mechanisms have two advantages over 1st-order coarse mechanisms. First, they will respond significantly to periodic textures even at low contrast levels and second, they can be very low frequency. If spatial frequency tuning of the 2nd-order mechanisms were sufficiently low, their large receptive fields might cover a considerable fraction of the grating. How large? Our results show that edge matching for gratings has an upper disparity limit of roughly 60 arc-min. If 2nd-order mechanisms have the same phase-disparity organization as has been proposed for 1st-order disparity energy mechanisms, this 60 arcmin limit would correspond to an interocular phase difference of 180 deg. If so, the period (360 deg) of the 2nd-order edge mechanism would be roughly double this limit, or about 120 arcmin (2 deg). Taking the reciprocal of the estimated period, the spatial frequency tuning of this mechanism would peak at approximately 0.5 cpd.

But suppose that disparity in 2nd-order mechanisms is encoded by position, i.e., an interocular difference in retinal location, instead of by an interocular difference in phase. Unlike phase disparity mechanisms, there is no necessary relationship between the upper disparity limit and the spatial frequency tuning of the mechanism. So, if disparity is encoded by position rather than by phase, the 60-arcmin disparity limit may not indicate anything
about the spatial scale of the mechanism. To the contrary, in their survey of disparity tuning in a large sample of primate V1 neurons, Prince et al (2002) found that a neuron’s preferred disparity, whether coded by phase or position, usually fell within the range specified by a half-period of the neuron’s preferred spatial frequency. Again, if the 60 arcmin edge-matching limit corresponds to a half-period, the disparity mechanism would be tuned to about 0.5 cpd. Previous studies of the tuning of 2nd-order mechanisms in spatial vision have found similar spatial frequency tuning (Sutter et al, 1995; Langley, Fleet & Hibbard, 1996; 1999).

For the sake of the argument, we have convolved a 1st-order detector, tuned to 3cpd with one half-image of our 3cpd, 6 deg wide, grating target (shown in red in Figure 10). To create the 2nd-order detector, we convolved the half-image with the same 3cpd filter, ‘half-squared’ the response (Heeger, 1992), and then convolved the half-squared response with a filter tuned to 0.5 cpd (shown in blue in Figure 10). These convolutions are shown in Figure 10B. They can be thought as being the response of the mechanisms prior to the calculation of disparity energy in both the pathways diagrammed in Figure 10A. While the 2nd-order mechanism clearly shows a substantial response to the edges that can be used to encode the edge disparity, its response to the center of the grating is minimal. Maybe the 2nd-order mechanism is lower in frequency and thus extends over a larger segment of the grating? Or maybe it is composed of several sub-units that are pooled together to cover the whole grating? Either of these possibilities could be true, but the enterprise of constructing a special ‘wide’ mechanism to explain our results is a bit suspect. A future study may demonstrate edge matching for a larger texture or a wider grating, forcing reconsideration of these spatial dimensions.

Is it necessary that the edge disparity be explicitly represented by some type of large-scale neural mechanism that covers the center of the grating? As mentioned above, 2nd-order mechanisms have been proposed to identify the borders of either static or moving textures. These coarse texture mechanisms are not required to fill in the region within the borders. The 2nd-order mechanism diagrammed in Figure 10A is simply a variant of these other texture mechanisms – a ‘depth’ texture mechanism. Like the other texture mechanisms, there may be no need to fill in the region within the depth borders. The 2nd-order responses generated by the edge disparities specify that there is a surface lying in a given depth plane. The 1st-order fine scale disparity mechanisms respond to the interocular phase (or position) information generated by surface relief, but these disparities are interpreted by the stereo system as being arrayed around the plane specified by the 2nd-order mechanism (Figure 10C). In their studies on stereo matching for repetitive patterns, Mitchison and McKee (1987ab) proposed exactly this type of interpolation model. No one is surprised that a large surface without any markings is assigned the interpolated depth defined by the disparity of its edges. The only reason why the grating seems different from the homogeneous surface is that the local signals appear to specify a disparity different from the edges. But as we noted in the introduction, zero phase does not necessarily imply zero disparity. Almost all stereo models utilize some type of consistency across scales to solve stereo matching within a given neighborhood. The interpolation model merely utilizes consistency across space to resolve the matching ambiguities associated with texture.

ACKNOWLEDGEMENTS

We thank Maria McKee for her excellent assistance in collecting and analyzing the data in this paper. This work was supported by NEI EY06644 to SPM, by NEI EY12286 to BF, and by NASA NAG 9-1163 to PV.

REFERENCES


