Nature of coarse-to-fine constraints on binocular fusion

Ann Marie Rohaly* and Hugh R. Wilson

Department of Ophthalmology and Visual Science and Committee on Neurobiology, Visual Sciences Center, University of Chicago, 939 East 57th Street, Chicago, Illinois 60637

Received March 29, 1993; accepted June 21, 1993

SeveraI stereo-matching algorithms posit that processing on coarse (low spatial frequency) scales constrains processing on fine (high spatial frequency) scales by shifting the disparity range over which fine mechanisms operate. If this were the case, stereo increment thresholds for high frequencies in the presence of low frequencies would be constant, regardless of base disparity. In contrast, we find that in the presence of low spatial frequencies, increment thresholds for high spatial frequencies increase with base disparity, as do increment thresholds for high frequencies presented alone. As a further test of whether information on coarse scales enhances processing on fine scales, diplopia thresholds were measured. In the presence of low frequencies, fusion ranges for high frequencies were severely reduced. These constraints were found to exhibit a power-law dependence on low frequency contrast and to operate over relatively localized areas of the visual field.

INTRODUCTION

The fact that visual information is processed by a set of mechanisms, each tuned to a modest range of spatial frequency,1-4 is now generally accepted. Furthermore, these mechanisms, or channels, are usually regarded as operating independently. Evidence exists, however, that there are significant interactions among the various mechanisms. For example, DeValois5 found both physiological and psychophysical evidence of reciprocal inhibition between low and high spatial frequency channels, while Olzak and Thomas6 demonstrated that widely separated spatial frequencies are not processed independently in certain discrimination tasks. In addition, Wilson et al.8 found inhibitory interactions between low and high spatial frequency mechanisms involved in the processing of binocular depth information.

Interactions among channels tuned to different ranges of spatial frequency have also been proposed theoretically in the context of various models of stereo matching.5-11 The models of Marr and Poggio,8 Nishihara,10 and Quam11 all employ coarse-to-fine stereo-matching algorithms. The general idea behind this class of algorithms is illustrated schematically in Fig. 1. In these models, the left and right eye images are first passed through a bank of bandpass spatial frequency filters. (For clarity, only three filters are shown.) Matches among the images are then sought beginning on the coarsest or lowest frequency scale, the search being conducted over a large disparity range. Next, the coarse matches are used to shift the images at the next level into closer correspondence and thereby to restrict the range of disparity over which matches are sought. This correspondence is then refined by a repeat of the process at finer scales, the processing at each level being driven by the output of the previous level.

The idea of spatial mechanism interactions operating in stereopsis was recently tested by Wilson et al.8 These investigators measured diplopia thresholds (limits of binocular fusion) for high spatial frequency test patterns in the presence of cosine gratings of various frequencies. They found that diplopia thresholds for high frequency test stimuli were reduced on average by a factor of 3.9 in the presence of cosine gratings whose spatial frequencies were 2.0 octaves below the test frequency. Furthermore, cosine gratings with frequencies 2.0 octaves above the test frequency had no effect. Diplopia thresholds were also unaffected by cosine masks with frequencies 4.0 octaves below the test frequency. On the basis of these results Wilson et al. concluded that the fusion range for any spatial frequency is constrained by lower spatial frequencies lying within 2.0 octaves of that frequency. These results support the general idea embodied in the coarse-to-fine models of stereo matching, namely, that low spatial frequency (coarse) information constrains the processing of high spatial frequency (fine) information. On the other hand, the results do not indicate whether the effect of this constraint is to shift the range of disparity sensitivity carried by high spatial frequencies.

One method of testing the shifting hypothesis is to determine the effects, if any, of low spatial frequency information on the stereoacuity and stereo increment thresholds of high spatial frequencies. Stereoacuity represents the smallest depth increment, measured in terms of binocular disparity, that can be reliably distinguished from zero depth, i.e., the fixation or frontoparallel plane. Stereo increment thresholds represent the smallest disparity increment that can be reliably distinguished from a particular standing (base) disparity. Thus stereoacuity simply refers to a stereo increment threshold at a standing disparity of zero.

Various investigators have determined that stereo increment thresholds for single spatial frequency stimuli increase exponentially as a function of base disparity.12-16 Recall that most coarse-to-fine models postulate that low spatial frequency (coarse) information is used to shift high spatial frequency (fine) information into correspondence. Another fundamental assumption of these models is that the magnitude of the shift is equal to the disparity carried by the low frequencies. If this is so, in the case of an increment threshold the effect of shifting high spatial frequency information would be to remove the base disparity. This hypothesis implies that, in the presence of
ity of low spatial frequencies (arrows and dashed-line func-

This prediction is illustrated schematically in Fig. 2.

In general, each stereogram comprised two D6 patterns: one located at fixation and the other located 1.33 deg peripherally (Fig. 3). In some cases, subjects had difficulty at this eccentricity, so the D6's were placed on either side of the fixation point at ±0.67-deg eccentricity.

In order to test the validity of the shifting hypothesis, we measured stereoacuity and stereo increment thresholds for high spatial frequency test stimuli in the presence of low spatial frequencies lying 2.0 octaves below the test frequency. We find that, under these conditions, increment thresholds are not constant as a function of base disparity; rather, they increase. This result implies that low spatial frequencies do not shift the processing range for high spatial frequency information, contradicting the assumption of many coarse-to-fine models. Diplopia thresholds were also measured to demonstrate the existence of coarse-to-fine constraints under our experimental conditions. Further experiments were carried out to explore the contrast dependence of the constraints. In particular, we find that the constraints exhibit a power-law dependence on the contrast of the low spatial frequency stimulus component and that the constraints can be evoked by relatively localized areas of low-frequency stimulation.

METHODS

All stereoscopic stimuli used in the experiments were generated by an Apple Macintosh II computer and were displayed on a pair of matched monochrome monitors with 8 bits per pixel resolution. Left and right eye images were presented on separate monitors (controlled by a single computer) and were viewed through a mirror stereoscope. Subjects were seated 1.57 m from the monitors with their heads positioned in a chin and forehead rest. At this viewing distance the monitor screens subtended 5.85 deg (height) by 7.80 deg (width), with each pixel subtending 43.9 arcsec. Disparities smaller than the width of one pixel, however, could be obtained by calculation of stimulus patterns shifted by fractions of a pixel. The monitors had previously been calibrated so that a linear gray scale centered on a mean luminance of 74 cd/m² was achieved. The mirror stereoscope transmitted approximately 70% of the light from the monitors, resulting in a mean luminance of 49.5 cd/m² as seen by the subject.

In all experiments, stimuli consisted of vertical bars whose vertical luminance profiles were described by a Gaussian and whose horizontal luminance profiles were described by the sixth derivative of a Gaussian (D6). D6 patterns were chosen for this study because they are well localized in both space and spatial frequency, having full bandwidths in spatial frequency of 1.0 octave at half-amplitude. This characteristic is desirable, as it tends to minimize the effects of the spatial inhomogeneity of the visual system and restricts stimulation to a narrow range of spatial frequency-tuned mechanisms. Further details concerning properties of these functions may be found elsewhere.

In general, low spatial frequencies, stereo increment thresholds for high spatial frequencies should be constant regardless of the base disparity, as opposed to increasing exponentially. This prediction is illustrated schematically in Fig. 2. High-spatial-frequency processing (solid-line function) is always shifted to a disparity range centered on the disparity of low spatial frequencies (arrows and dashed-line func-

Fig. 1. Illustration of the idea embodied in coarse-to-fine stereo-matching algorithms. Left and right eye images are passed through a bank of bandpass filters tuned to restricted ranges of spatial frequency. Only three filters are shown in the diagram for clarity; in practice there could be any number of filters. Matches between the filtered left and right eye images are first sought on the coarsest (lowest frequency) scale. The matches assigned on the coarse scale serve as constraints on the matching process at the next-finer scale (labeled medium in the diagram). The coarse matches are used to shift the images at the next level into closer correspondence and thereby to restrict the range of disparity over which matches are sought. This process is repeated at each successive level until the finest scale is reached.

Fig. 2. Prediction of coarse-to-fine models regarding the shape of stereo increment threshold functions. According to the models, high spatial frequency processing (solid-line function) is always shifted to a disparity range centered on the disparity of low spatial frequency information (arrows and dashed-line functions), in effect subtracting out the base disparity. This shifting causes stereo increment thresholds for high spatial frequencies in the presence of low spatial frequencies to be processed as effectively as they would be with a standing disparity of zero (circles), and stereo increment thresholds are constant, regardless of base disparity (dotted line).
Two general types of experiment were performed: (1) determination of stereoacuity and stereo increment thresholds and (2) determination of diplopia thresholds. Stereoacuity and stereo increment thresholds were measured with the method of constant stimuli. When the subject was maintaining fixation on the dot (2 arcmin square) in the center of the display, she initiated a trial with a button press. Two 500-ms stimulus intervals followed, one containing the test stimulus and one containing the standard. The stimulus intervals were separated by a 500-ms blank at the mean luminance of the stimuli. For the stereoacuity experiments [Fig. 3(a)] the standard consisted of two D6's in the plane of fixation (fronto-parallel plane). For the test stimulus the left-hand D6 was in the plane of fixation and the right-hand D6 was either nearer (crossed disparity) or farther (uncrossed disparity) than the plane of fixation. For the stereo increment threshold experiments [Fig. 3(b)] the standard consisted of a left-hand D6 in the plane of fixation and a right-hand D6 of a particular base disparity. For the test stimulus the disparity of the right-hand D6 was equal to the base disparity plus a disparity increment. In both cases after each trial the subject indicated with another button press which interval contained the test stimulus.

In each experimental run all stimuli had exclusively crossed or exclusively uncrossed disparity. Four different disparities (or disparity increments) were presented, with forty repetitions of each, in random order. The disparities used were selected on the basis of pilot experiments. The resulting data, in the form of percentage correct versus disparity, were fitted with a Quick^19 or a Weibull^19 function by use of a maximum-likelihood estimation technique. The 75%-correct point on the fitted function was taken as threshold for that particular run. The data presented here represent means of 2 or 3 thresholds determined on separate days.

Diplopia thresholds were measured with the same stimulus configuration that we used to measure stereoacuity [Fig. 3(a)]. Once again the subject initiated each trial with a button press. In this case, however, each trial consisted of a single 500-ms interval. The subject’s task was to indicate whether the right-hand D6 appeared fused or diplopic (double). In each experimental run five different disparities, either all crossed or all uncrossed, were presented, with 25 repetitions of each. As before, the disparities were selected on the basis of pilot experiments. The resulting data, in the form of percentage diplopic versus disparity, were fitted as described above. As the data ranged from 0 to 100%, the 50% diplopic point on the fitted function was taken as threshold for that particular run. Once again, the data presented below represent the means of 2 or 3 thresholds collected over the course of the study.

Three observers participated in this study. Two were the authors and the third was naïve with respect to the purpose of the study. All observers had good stereopsis and wore their normal corrective lenses during the experiments.

RESULTS
Stereoaucity and Stereo Increment Thresholds
The purpose of these experiments was to determine whether low spatial frequency information enhances pro-
processing of high spatial frequency information when processing occurs off the horopter. As discussed above, current coarse-to-fine models posit that low spatial frequency information shifts the range of processing of high spatial frequency information. This assumption implies that stereo increment thresholds for high frequencies in the presence of low frequencies should be constant, as opposed to increasing exponentially (see Fig. 2). Stereoacuity and stereo increment thresholds are shown in Fig. 4 for a low spatial frequency D6 (open squares), a high spatial frequency D6 (open circles), and the superposition of high and low spatial frequency D6's (filled circles). Subject HRW was tested with 3- and 12-cycle/deg (cpd) D6's and subject AMR with 2- and 8-cpd D6's. For all stimulus conditions, increment thresholds increase with increasing base disparity. The functions are approximately exponential (straight lines in semilog coordinates), as expected from previous studies.

The fact that increment thresholds for the compound D6 stimuli (filled symbols) are not constant as a function of base disparity implies that low spatial frequency information does not shift the processing range for high spatial frequency information.

Because the high and low frequency D6's were always displaced in depth as a unit (locked condition, Fig. 5(a)), the lack of a significant improvement in increment thresholds may be due to the fact that processing on the low-frequency scale resulted in a poor estimate of the depth of the test stimulus. This poor estimate would affect the accuracy of depth judgments on the high-frequency scale as a result of the propagation of constraints from coarse to fine scales. To control for this possibility, some of the increment threshold measurements were repeated with the high and low frequency components of the test stimulus unlocked from each other [Fig. 5(b)]. In this case, the low frequency D6 was always presented at the base disparity, while the high frequency D6 was displaced in depth around it. Thus the low spatial frequency component offered no depth cues, as it was not displaced on successive trials; only the high spatial frequency component was displaced in depth.

Figure 6 shows increment threshold functions for both the locked (open symbols) and unlocked (filled symbols) conditions. The data for the eccentric, locked condition were taken from Fig. 4. The eccentric, unlocked data for subject HRW were measured at an eccentricity of 0.67 deg, as he found the task extremely difficult at the usual 1.33-deg eccentricity. It can be seen from Fig. 6 that unlocking the high and low frequency D6's never caused the increment threshold functions to be flat; the functions still increase with base disparity. In some cases, this increase was so dramatic that the task became impossible to perform with nonzero base disparities. These cases are indi-
Fig. 6. Stereo increment thresholds for locked versus unlocked condition at two eccentricities. Data for the locked condition are displayed with open symbols and data for the unlocked condition with filled symbols. The data for the eccentric, locked condition are from Fig. 4. Increment thresholds for the unlocked conditions increase with base disparity just as thresholds for the locked conditions do. In some instances, thresholds were too large to measure (points with upward-pointing arrows). Data for eccentric and foveal viewing are similar; this rules out any effects of eccentricity or eye movements.

cated by the upward-pointing arrows on the symbols at the tops of the panels. Some of the unlocked data indicate a slight improvement over the locked condition (e.g., 4-arcmin uncrossed disparity), but overall the thresholds still increase as a function of base disparity. Thus poor depth localization of the low spatial frequency D6 is not responsible for the increase in increment thresholds with base disparity.

To control for the effects, if any, of eccentricity and eye movements, thresholds were also measured at fixation with a stimulus presentation time of 167 ms. These data are also shown in Fig. 6 (circles). The same pattern of results was obtained under these conditions as with eccentric viewing and a 500-ms duration. These data demonstrate that neither eccentricity nor eye movements affected the previous results in any way. In addition, because the high and low frequency D6's carried different disparities in the unlocked case, this condition is a partial control for the effect of the phase relationship between them. Wilson et al. also found no phase dependence of low frequency constraints on high frequency diplopia thresholds.

Diplopia Thresholds
Because the addition of the low frequency D6 did not have a significant effect on the increment threshold of the high frequency D6 (see Fig. 4), we measured diplopia thresholds to determine whether there were any interactions at all between the two frequencies. This experiment was intended to extend the findings of Wilson et al. They found that diplopia thresholds for high frequency D6's were reduced on average by a factor of 3.9 in the presence of low frequency cosine gratings. In our experiment both components of the stimulus were D6's, and therefore our stimuli were more spatially localized than those of the previous study. Baseline diplopia thresholds first were measured for high and low spatial frequencies presented alone. As shown in Fig. 7, thresholds averaged 19.9 arcmin for the low spatial frequency alone and 14.0 arcmin for the high spatial frequency alone, in agreement with the results of previous studies.

For the compound-frequency case, thresholds first were measured with the stimulus components locked together. In this case, thresholds averaged 20.6 arcmin, a value that is somewhat larger than the average threshold for the high spatial frequency alone (14.0 arcmin). Thus, at first glance, it seems that the fusion range for the high frequency D6 was extended in the presence of the low frequency. Note, though, that 20.6 arcmin is not significantly different from the average threshold for the low spatial frequency alone (19.9 arcmin). This result is exactly what one would expect given the facts that the high

Fig. 7. Diplopia thresholds for three subjects under various conditions. Subject AMR (light hatch) was tested with 2- and 8-cpd D6's and subjects HRW (dark hatch) and PS (stipple) with 3- and 12-cpd D6's. Thresholds for the high spatial frequency in the presence of the low spatial frequency in the locked configuration are not significantly different from those for the low spatial frequency presented alone. However, in the unlocked configuration, the presence of the low spatial frequency reduced diplopia thresholds by a factor of 4.5 relative to the threshold for the high spatial frequency presented alone.
and low frequency D6's were always at the same disparity and that the fusion ranges of high frequencies are constrained to be centered on the local disparity of low spatial frequency information. Therefore diplopia thresholds were remeasured with the stimulus components in the unlocked configuration. For this experiment subjects were instructed to make their responses on the basis of the appearance of the high spatial frequency component of the stimulus. Figure 7 shows that when the high and low frequencies were unlocked, thresholds were lower than those for either the high or the low frequency alone. Thresholds averaged 3.11 arcmin, or a factor of 4.5 lower than the average threshold for the high frequency alone. Thresholds were also determined with the low spatial frequency 4.0 octaves below the test frequency (as opposed to 2.0 octaves lower in the previous case). In this case diplopia thresholds were identical to those for the high spatial frequency presented alone. These results are in agreement with the finding of Wilson et al. that low spatial frequencies constrain the fusion range of high spatial frequencies lying within 2.0 octaves of that frequency. Furthermore, this agreement demonstrates that the failure to find a spatial frequency interaction in the stereocuity and stereo increment threshold experiments was not due to differences between the stimuli used in the study of Wilson et al. and those used in the present study. Because we used D6's as opposed to cosine gratings, our result also extends the previous study in demonstrating that coarse-to-fine constraints are relatively local phenomena; full-field stimulation is not necessary for evoking them.

We repeated the diplopia threshold measurements at fixation with a 167-ms stimulus duration to control for the effects of eccentricity and eye movements. For the compound-frequency stimulus, measurements were made only for the unlocked condition because of the lack of any frequency interaction in the locked condition. Thresholds for the compound stimuli were normalized with respect to those for the high frequency component alone and are shown in Fig. 8. In this figure a value of ±1.0 on the ordinate corresponds to the case in which the addition of the low frequency D6 has no effect on the diplopia threshold of the high frequency D6. The data show that there is a reduction in the fusion range at fixation with the shorter stimulus presentation time and that the magnitude of this reduction is approximately the same as that obtained with peripheral viewing. Thresholds were reduced on average by a factor of 5.1 at fixation, compared with a factor of 4.5 at 1.33-deg eccentricity.

**Contrast Dependence of Coarse-to-Fine Constraints**

The purpose of this experiment was to determine the dependence of the frequency interactions found in the previous experiment on the contrast of the low spatial frequency D6. In that experiment both the high- and the low spatial frequency components were maintained at 50% contrast. In this experiment diplopia thresholds were measured for the compound D6 stimulus with the high spatial frequency component at 50% contrast and the low spatial frequency component at contrasts of 6.25%, 12.5%, and 25%. Data for the various contrast conditions were collected in separate experimental runs. The high and low frequency D6's were always in the unlocked configuration.

Figure 9 shows diplopia thresholds for high spatial frequency D6's in the presence of low spatial frequency D6's of various contrasts. Thresholds for crossed and uncrossed disparities were similar, so they were averaged together to produce each datum. The points at 0 and 50% contrast are the thresholds from Fig. 7 for the high frequency alone and for the low plus high frequencies (unlocked), respectively. As the contrast of the low spatial frequency increases, the diplopia threshold for the high spatial frequency falls rapidly and then approaches an asymptote.

In order to determine the functional form of these data, we replotted the data in double-logarithmic coordinates (excluding the points at 0 contrast). Figure 10 shows that when the data are plotted in this way, the points fall along straight lines, indicating a power-law dependence on con-
The magnitudes of the overall threshold reduction between the two cases reflect the individual differences in slope among the subjects. The slopes of the lines, representing the exponent of the power law, lie between 0.437 and 1.09. The data fall along straight lines with slopes between 0.437 and 1.09. The straight lines indicate a power-law dependence on contrast. The fitted lines have correlations (r) of 0.993 (AMR), 0.980 (HRW), and 0.850 (PS).

Fig. 10. Data of Fig. 9 excluding the points at 0 contrast replotted in log-log coordinates to illustrate the functional form of the data. The data fall along straight lines with slopes between 0.437 and 1.09. The straight lines indicate a power-law dependence on contrast. The fitted lines have correlations (r) of 0.980 (HRW), 0.993 (AMR), and 0.850 (PS).

DISCUSSION

We found that relative depth discrimination, as reflected by stereo increment thresholds, is not improved in the presence of spatial frequencies 2.0 octaves below the test frequency (see Fig. 4). Under these conditions fusion ranges (diplopa thresholds), however, are severely reduced (see Fig. 7). Taken together, these results demonstrate that coarse spatial scales do constrain processing on fine spatial scales but that a shift in fine scale processing does not occur. This finding does not support the coarse-to-fine stereo-matching models of Marr and Poggio, Nishihara, and Quam. In each of these models disparity estimates obtained on coarse scales are used to shift the filtered stereo images on fine scales into closer correspondence.

A possible explanation for the lack of significant improvement in high spatial frequency increment thresholds measured in the presence of low spatial frequencies is that the reduction in fusion range resulting from the coarse-to-fine constraints caused the stimuli to be diplopic and interfered with the subjects' performance. In some instances, the measured increment thresholds were close to the reported diplopa thresholds (compare points for the eccentric, unlocked case in Fig. 6 with those for the un-locked case in Fig. 7), so some stimuli may have been perceived as diplopic. The perception of diplopa is not a given, however, because our threshold criterion was that the subject report an unfused stimulus on only 50% of the trials. It is not clear, though, what effect, if any, diplopa would have had on the results, because diplopa stimuli still convey a robust sense of depth. A more important consideration is the fact that the majority of increment thresholds for the unlocked case were similar to those for the locked case (see Fig. 6). In the locked configuration, diplopa thresholds were considerably higher (see Fig. 7), so none of the increment threshold stimuli would have been perceived as diplopa. Thus, given the similarity of the thresholds in the two cases, it is unlikely that the reduction in fusion range caused by the coarse-to-fine constraints could be responsible for the increase in increment thresholds with base disparity.

Another factor that may have affected the results is the experimental method used to collect the increment threshold data. Westheimer demonstrated that stereo increment thresholds measured with a simultaneous presentation of base and base-plus-increment disparities in a single temporal interval are significantly smaller than those measured in a two-temporal-interval task. The results presented here confirm this finding; our thresholds are comparable with those found by McKee et al., who used a two-interval task, and are higher than those found in a number of other studies in which a single-interval task was used. Therefore the fact that stereo increment thresholds were not constant in the presence of low spatial frequencies may simply be due to the fact that the visual system is less sensitive to depth increments under the experimental conditions employed in this study.

In the unlocked stimulus configuration [see Fig. 5(b)], however, both disparities were presented simultaneously in a single interval; the low spatial frequency carried the base disparity, while the high spatial frequency carried the base-plus-increment disparity. In reality, then, the second stimulus interval was not necessary for the subjects to perform the task; the subjects simply needed to judge whether the high and low spatial frequency D6's appeared to lie at the same depth. On the other hand, the fact that thresholds measured with consecutive presentation of base and base-plus-increment disparities are elevated may have affected the data from the locked
condition, since subjects were forced to compare depth estimates across the two temporal intervals. We can rule out this hypothesis, however, by noting that the increment thresholds measured with both the locked and the unlocked stimuli were similar (see Fig. 6).

In addition to those in the compound-frequency case, increment thresholds measured in the single-frequency cases are also at odds with the coarse-to-fine strategy embodied in the models of Marr and Poggio, Nishihara, and Quam. Another requirement of these algorithms is that disparity be processed over a wide range on coarse spatial scales and over a narrow range on fine spatial scales. As Badcock and Schor pointed out, this requirement implies that increment threshold functions for low spatial frequencies should increase more slowly with increases in base disparity than should functions for high spatial frequencies. In contrast, they found exactly the opposite behavior; thresholds increased faster for frequencies below -2.4 cpd. Figure 4 shows that under our experimental conditions, increment thresholds for all spatial frequencies increase at approximately the same rate. This is consistent with the finding of Badcock and Schor that above 2.4 cpd all increment threshold functions have the same shape.

Our results are also inconsistent with the assertion of Marr and Poggio that eye movements play a crucial role in the stereo-matching process. All the experiments were performed with two different presentation times: 167 and 500 ms. In all cases the pattern of results was similar under the two conditions. Therefore, since an eye movement cannot be initiated in 167 ms, we can rule out the possibility that eye movements are necessary for stereopsis.

A possible interpretation of the diplopia threshold result is that the observed reduction in fusion range of high spatial frequencies in the presence of low spatial frequencies is due simply to the presence of a zero-disparity reference in close spatial proximity. In other words, the effect is not dependent on the spatial frequency of the background and therefore does not demonstrate the existence of coarse-to-fine constraints. The frequency specificity of this reduction in fusion range, however, is demonstrated by the results of a number of control experiments performed in both the study of Wilson et al. and the present study. In particular, Wilson et al. found no reduction in fusion range when the cosine mask carried a frequency either 2.0 octaves above or 4.0 octaves below the test frequency. In addition, for more spatially localized masks, we also found no reduction in fusion range when the low spatial frequency D6 was 4.0 octaves below the test frequency. On the basis of these results, we conclude that the fusion range for any spatial frequency is constrained by lower spatial frequencies lying within 2.0 octaves of that frequency. Thus the observed reduction in fusion range is not due simply to the presence of a zero-disparity reference and clearly indicates the presence of coarse-to-fine constraints on stereo matching.

If the existing coarse-to-fine constraints do not involve a shifting of disparity information on different spatial scales, how are the constraints implemented neurally? A qualitative model that can account for the present data has been proposed by Wilson et al. In their model each spatial scale comprises pools of neurons tuned to zero, near, and far disparities, in accord with psychophysical and neurophysiological evidence. The coarse-to-fine constraints are implemented through various inhibitory connections among the spatial scales. For the purpose of explaining the present data, we need only consider one set of connections, namely, low frequency zero-disparity units inhibiting high frequency near and far disparity units. The effect of the inhibition is to cause high frequency processing to take place over a small range centered on zero disparity.

In summary, the data presented above reject the assertion of coarse-to-fine models of stereo matching that information on low spatial frequency scales constrains processing on high spatial frequency scales by shifting the disparity range of high frequency processing. Rather, the effect of coarse-to-fine constraints is to select a range of disparity processing on high frequency scales that is already hard wired into the visual system. In addition, our results provide evidence that these constraints are relatively local phenomena; full-field stimulation is not necessary for evoking them. Together, these results provide important constraints on future models of stereopsis.

ACKNOWLEDGMENTS

This research was supported in part by National Institutes of Health training grants NS07195 to A. M. Rohaly and EY02158 to H. R. Wilson.

*Present address, U.S. Army Research Laboratory, Human Research and Engineering Directorate, Attn: AMSRL-HR-SD, Aberdeen Proving Ground, Maryland 21005-5425.

REFERENCES


