
The 'uniqueness constraint' and binocular masking

Suzanne P McKee, Mary J Bravo, Harvey S Smallman

Smith-Kettlewell Eye Research Institute, 2232 Webster Street, San Francisco, CA 94115, USA

Gordon E Legge

Department of Psychology, University of Minnesota, Minneapolis, MN 55455, USA

Based on paper presented at the Conference on Binocular Stereopsis and Optic Flow, Toronto, Canada, 22–26 June 1993

Abstract. In stereo-matching algorithms, the 'uniqueness constraint' requires that a feature in one stereo half-image be matched to, at most, one similar feature in the other half-image. Experiments are reported in which binocular contrast thresholds and depth-discrimination judgments have been used to determine whether the human stereo system makes unique matches. A single high-spatial-frequency target in the left eye was paired stereoscopically with two identical targets, presented near retinal correspondence (± 3.5 min of disparity), in the right eye. Contrast-increment thresholds were measured for each of the targets in the right eye, and it was found that the target in the left eye masked both. Indeed, the amount of binocular masking for each member of the double target nearly equaled the masking observed when only a single target was presented to the right eye. Depth judgments confirmed that the target in the left eye had been matched to both targets in the right eye. It is concluded that uniqueness is not an absolute constraint on human stereo matching.

1 Introduction

In many computational models of stereo matching, each feature in one eye is allowed to match at most one feature in the other eye (Marr and Poggio 1976; Marr 1982; Frisby and Pollard 1991). This stereo-matching rule is known as the 'uniqueness constraint'. Although this constraint greatly simplifies computer matching in natural images and in dense random-dot stereograms, there is no compelling evidence that the human stereo system is similarly constrained. In fact, Panum's limiting case is generally regarded as evidence to the contrary. Panum's limiting case is a stereogram consisting of a single vertical line presented to one eye and a pair of identical vertical lines, at and near retinal correspondence, presented to the other eye.⁽¹⁾ If the separation between the pair is small, observers generally perceive two lines at different depths.

The most common explanation for this phenomenon, attributed to Hering by Ogle (1962), is that the human binocular system combines the single line in one eye with both lines in the other eye—a case of double matching. Gettys and Harker (1967) explicitly tested this explanation by asking observers to match a probe stimulus to the apparent depth of the disparate extra line in the Panum stereogram. For the most part, their observers matched the extra line to the fixation plane, ie to the same plane as the line presented at correspondence. Only when their observers were allowed to change convergence did they match the extra line to the depth predicted by its disparity. Since the change in convergence could have shifted the match from one line to the other, these results cannot be considered strong support for double matching.

⁽¹⁾In his first great paper on binocular disparity, Wheatstone (1838) included a stereogram showing a single vertical line in one eye paired with a vertical line and a second tilted line in the other eye. This stereogram also produces the percept of two lines with different depths, one tilted in depth across the other.

Westheimer (1986) asked observers to judge the relative depths of the two lines forming a Panum stereogram. He found that the extra line was reliably perceived at a depth consistent with double matching, but only for quite small disparities (≤ 6 min).

Apparent depth is not always an infallible indicator of stereo matching. A monocular feature, lacking an appropriate mate in the other eye, will also appear at some depth relative to the binocularly matched features. Kaufman (1976) argued that unmatched monocular features were localized to the depth of the fixation plane, and that one of the lines in Panum's stereogram was unmatched. This unmatched line appeared at a different depth from the binocularly matched line because small shifts in convergence changed the disparity of the matched line and, thus, the relative depth of the unmatched line. By systematically varying convergence, Howard and Ohmi (1992) showed that the apparent depth of a monocular line was predictably affected by changes in the disparity of a binocularly matched feature, in support of Kaufman's hypothesis. Other factors can also affect the apparent depth of unmatched monocular features. Julesz (1971) contended that the depth seen in the Panum stereogram arose because one of the lines was interpreted as occluded by the other (see also Helmholtz 1910). Ono et al (1992) and Chang et al (1993) have provided evidence supporting Julesz's contention. Nakayama and Shimojo (1990) have also shown that occlusion is important in determining the depth of unmatched features in random-dot stereograms.

Stereopsis is not the only measure of whether signals from the two eyes have been combined. Binocular masking is also a strong indicator of binocular interactions (Levi et al 1979; Legge 1984). For example, we can measure a contrast-increment threshold for a monocular feature, and then add a similar feature at correspondence in the other eye. If the monocular responses to the two features are summed together binocularly, the increment threshold will increase; the neural signal from the feature in the other eye masks the monocularly presented increment. The presence of such a masking effect is diagnostic of binocular interaction.

Dichoptic masking is a special case of binocular masking in which the absolute contrast threshold⁽²⁾ for a monocular feature is measured in the presence of a higher-contrast feature at retinal correspondence in the other eye. Recently, we demonstrated that dichoptic masking is contingent on stereo matching (McKee et al 1994). In agreement with previous studies, we found that a high-contrast bar in one eye strongly masked the absolute threshold for a test bar presented in the other eye at retinal correspondence. If, however, another high-contrast bar was presented adjacent to the test bar, the masking effect nearly disappeared. Apparently, the superior match between the two high-contrast features prevented masking at retinal correspondence. This result shows that masking effects can act as a probe for stereo matching. It also suggests that the 'uniqueness constraint' may apply to binocular masking. Or does it? Incidental evidence from our study suggested that stereo matching has a threshold—that the binocular system will not combine the contrast signals of features in the two eyes if the contrasts differ by much more than a factor of five, except by default. The high-contrast mask in one eye was combined with the very-low-contrast test bar in the other eye only when there was no other suitable match. What happens if there are two suitable matches, both above this putative matching threshold? Will the binocular system match a single feature in one eye to two similar high-contrast features in the other eye?

In the present study we used binocular masking to examine the question of double matching. Using a variant of Panum's limiting case, we measured monocular

(2) Threshold is defined as the smallest test contrast that can be reliably detected in the presence of a contrast pedestal. By convention, the contrast threshold measured when the pedestal is zero is called the 'absolute' threshold.

contrast-increment thresholds for each of two targets in the right eye that were paired stereoscopically with a single target in the left eye (see double target in figure 1a). Westheimer's results (1986) indicated that double matching was possible only at small disparities. So, to maximize the change of double matching, we presented the left-eye target at a retinal locus halfway between the loci of the two identical targets in the right eye, and kept the disparities small (± 3.5 min). Also, we used high-spatial-frequency band-limited targets to minimize the intrusion of low-frequency mechanisms into our contrast measurements.

Under conditions of stable convergence, this double target provides the binocular system with two equally valid matches. If the human binocular system always makes unique matches, the single target in the left eye will mask only one of the targets in the right eye. But which one? Most likely, the choice would be random, and would vary over time or from trial to trial. Given this type of probabilistic matching, the left-eye target would have a smaller masking effect on the thresholds for the double target than it would if there were only one valid match available, as in the single target shown in figure 1b. On the other hand, if the target in the left eye always combines with both targets in the right eye, the increment thresholds for the double and single target configurations would be the same. Our experimental question is quite simple—are the increment thresholds for the single-target and double-target configurations the same?

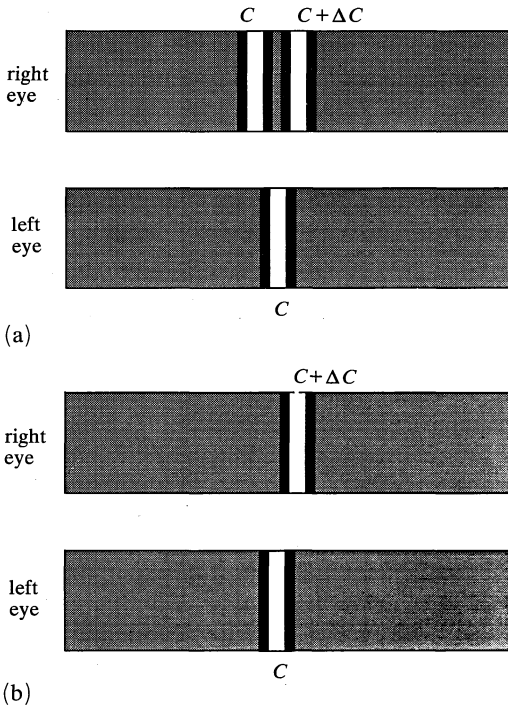


Figure 1. Diagram of high-spatial-frequency triphasic targets used for the binocular masking study. (a) Double target: a single target in the left eye is presented stereoscopically with two identical targets in the right eye. (b) Single target: a single target in the left eye is presented with a single target in the right eye; the right eye's target occupies the location of one of the two targets in the double configuration. For each of the targets (pedestal contrast value C) contrast-increment (ΔC) thresholds are measured in the right eye only. Are the contrast increment thresholds for the double and single configurations the same?

2 Methods

2.1 The stimulus configuration

The actual dimensions of the double-target configuration are shown diagrammatically in figure 2. Luminance-balanced triphasic targets, designed to confine target energy to a high-spatial-frequency bandwidth, were presented on a uniform background, 1.5 deg wide and 0.8 deg high. Each triphasic target consisted of a bright vertical bar, 2.75 min wide, flanked by two thin dark bars, each 1.375 min wide; the height of each bar equaled the height of the background (0.8 deg). In the double-target condition, the center-to-center separation between the pair of triphasic targets in the right eye was 7 min. In the single-target configuration, only one of the two triphasic targets was presented to the right eye at either a crossed or an uncrossed disparity of 3.5 min with respect to the target in the left eye. For all experiments, short black nonius lines were superimposed on the upper and lower edges of the uniform fields, 7 min from the targets (see figure 2). The nonius lines were actually small rectangular pieces of black chart tape, about 6 min in width, that had been applied to the center of each monitor at the midpoint of the background field; they were visible whenever the uniform background was illuminated. The locations of the stereo half images were defined with respect to the physical location of the nonius lines. Therefore, if the nonius lines were perfectly aligned (specifying the same visual direction in the two eyes), the disparity of each target was given by the relative distances between each half image and its respective nonius line.

The integrated luminance of the center and surround of physically balanced triphasic targets is equal to the uniform background. However, there is no guarantee that physically balanced stimuli are balanced for the human visual system. For example, the response to a luminance decrement may not equal the response to a luminance increment because of early intensive nonlinearities in luminance processing. As Burton (1973) noted, a nonlinearity of this type would introduce a small contrast signal at low spatial frequencies. To ensure that the center and flanks were psychophysically balanced, we measured detectability (d') for a fixed contrast as we systematically varied the luminance ratio of the center to the flanks. Since human

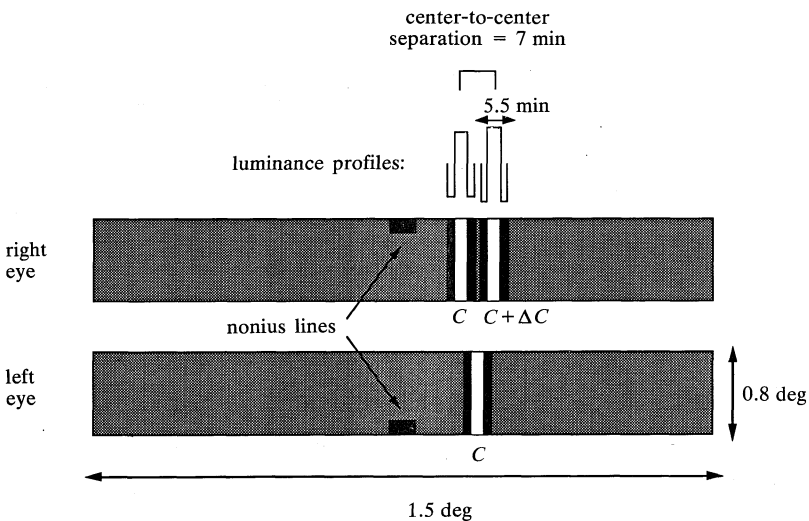


Figure 2. Diagram of the dimensions of the double target, showing the luminance profiles of the triphasic targets. Nonius lines were visible whenever the background was illuminated. Subjects waited until nonius lines appeared to be in alignment before initiating a 200 ms presentation of the triphasic stimuli.

contrast sensitivity falls monotonically with increasing spatial frequency beyond 2 or 3 cycles deg^{-1} , minimum detectability should correspond to the ratio that best isolates a high-spatial-frequency bandwidth. At the luminance ratio that produced minimum detectability, the dark flanks were relatively lower in luminance than at the ratio that produced physical balance, confirming similar results from Bravo and Blake (1992). Nevertheless, the difference was quite small; the integrated luminance of the psychophysically balanced triphasic targets was 98% of the background. The psychophysically balanced targets were invisible at long viewing distances (>4 m), even at high contrasts—additional evidence of their high-spatial-frequency bandwidth. A Fourier transform of the luminance distribution of these targets indicated that their peak spatial frequency was roughly 12 cycles deg^{-1} with a bandwidth of 1.6 octaves.

The targets were drawn by computer-generated signals on the screens of two Tektronix 608 monitors, each equipped with a P-31 phosphor. A custom-designed function generator was used to convert the point-addressable X - Y display monitors into a raster-scan mode. A high-speed (125 kHz) triangle wave was fed into the Y -axis to create a vertical raster which was moved horizontally by a slower sawtooth ramp (244 Hz) fed into the X -axis. There were 1024 lines drawn within a small region in the center of each monitor. 'Ringing' and other nonuniformities at the edge of the background fields were obscured by placing a black mask with a rectangular aperture at the same location in front of each of the two monitors. At the 1.25 m viewing distance, the center-to-center distance between each raster line was 8.25 s arc. The luminance of each raster line could be set independently by varying the Z -axis poke values. The Z -axis had 8 bits of resolution, and the gain could be changed to achieve a larger or smaller luminance range as needed for greater flexibility in measuring contrast thresholds. For most experimental conditions, the stimulus bars were filled with raster lines of the same poke value, creating a physically uniform bar, but for some test conditions the poke values did not allow sufficient luminance resolution. In these cases, smaller steps in luminance were achieved by regularly alternating between two poke values of adjacent raster lines to create an apparently uniform field a half step between the two luminance values.

The luminances for each Z -axis poke value and half-step value were measured with a Pritchard photometer, by means of a 6 min probe centered on a vertical bar, 12 min wide, that was presented in the center of the uniform background. The luminance of the background field on each screen was roughly 98 cd m^{-2} on both screens for the contrast measurements, and 92 cd m^{-2} for the depth measurements. The luminance measurements were converted into Michelson contrast according to the formula $(L_{\max} - L_{\min}) / (L_{\max} + L_{\min})$. Overhead fluorescent lighting located about 6 m from the CRT monitors supplied indirect illumination of the experimental setup at a low photopic level. All luminance calibration measurements were made with this indirect lighting under the same conditions as used for the experiments.

The observers viewed the two monitors via a mirror stereoscope. The monitors were mounted in parallel on a large table so that each screen was 1.25 m from the observer's eyes. One set of mirrors was placed in front of each screen and another set in front of the eyes; the angled mirrors in front of the screens reflected the image into the mirrors in front of the eyes. These mirrors formed a steep angle in front of the observer's nose so that the image from one screen was visible to each eye. A septum, placed perpendicular to the center of the observer's forehead, ensured that each eye could see only one screen.

2.2 Procedure

At the beginning of each trial for all experimental conditions, the observer was shown only the uniform fields with the superimposed nonius lines. When the nonius

lines appeared to be aligned, the observer pressed a button to initiate a 200 ms presentation of the stimulus configuration. After 200 ms, the stimulus disappeared, but the uniform fields remained visible until the observer responded. After the response of the observer, or after a 2 s wait period, the uniform field was briefly extinguished for 100 ms, signaling the beginning of a new test trial. To preserve a stable state of luminance adaptation, observers were forced to wait 800 ms before initiating a target presentation after this signal.

For the contrast-discrimination measurements, the observer judged whether a contrast increment had been added to a test bar in the right eye. In the double-target condition, the observers were told which of the pair of test targets would contain the increment, except for in one control condition, described below, in which the location of the increment was randomized. For all contrast measurements, we used a 'yes-no' procedure to determine the contrast-increment threshold, rather than a two-alternative forced-choice procedure with two temporal intervals, because we thought observers might initiate a convergence movement prior to or during the second temporal interval. On each trial, the designated test bar was presented either at the pedestal contrast value, eg 50%, or at the pedestal value plus the increment, eg 50% + ΔC . Observers indicated by pressing one of two buttons whether the test bar contained the increment or not. Feedback was provided. Each incremental contrast level was presented in a separate block of 80 trials with equal probability of the increment being present or absent on each trial. The hits and false alarms from the 80 trials were used to estimate d' (Elliott 1964). A minimum of three contrast levels (240 trials total) were used to determine the contrast corresponding to a threshold criterion of $d' = 1.0$. This threshold criterion was estimated by fitting a straight line (least-squares criterion) to the data points plotted in $\log d'$ versus \log contrast coordinates (Foley and Legge 1981). For some conditions, we made repeated measurements, each based on at least 240 trials, to evaluate threshold variability; the error shown in the tables is the standard error from three or more measurements, or the sample standard deviation of two measurements.

We also made depth-discrimination measurements for each of the two test targets in the double-target condition, and for the two conditions in which a single test target was presented to the right eye at either a crossed or uncrossed disparity of 3.5 min. For these measurements, a reference target of variable disparity was presented 27 min to one side of the test targets shown in figure 2. To facilitate comparison, the reference target was placed on either the left or the right side, adjacent to the test target that was being judged in the double-target condition. In the single-target conditions, the reference target was also placed 27 min to the left or right, depending on the location of the single test target. The reference target, like the test targets, was a high-frequency triphasic stimulus. All targets were presented at a contrast of 50% for the depth measurements. As in the contrast-discrimination tests, observers aligned the nonius lines before initiating a 200 ms presentation of test and reference targets. This duration is somewhat longer than convergence latency, but is certainly too short to permit the completion of a voluntary convergence movement during a test trial.

Observers judged whether the designated test target was in front of or behind the reference target in depth. They indicated their choice by pressing one of two buttons on a response box. No feedback was provided. On each trial, the reference target was presented at random at one of nine disparities. The nine disparities were chosen so that they spanned a sizeable range (6 min) of either uncrossed or crossed disparities, always including the fixation plane; the observer did not know the mean disparity of the range prior to testing. The percentage of trials on which the test target was judged "in front" was measured as a function of the reference disparity. A cumulative

normal curve was fitted by probit analysis to the data from 200 trials. The median depth of the test stimulus, ie the disparity corresponding to the 50th percentile of the psychometric function, was estimated from the fitted function. On occasion, the reference target was also presented for a range of disparities far from the anticipated position of the test target, eg 0–6 min of crossed disparity when the test target was apparently matched behind the fixation plane. In this case, the observer always gave the same response, eg 0% “in front”—an indication that the reference target did not influence the perceived depth of the designated test target.

2.3 Observers

Four experienced psychophysical observers participated in these experiments. All had normal, or corrected-to-normal, visual acuity (20/20) for the viewing distance of 1.25 m, and normal stereopsis. All four observers were aware of the purpose of the experiments, and they had practised in many of the experimental conditions before the data presented here were collected.

3 Results

3.1 Monocular contrast discrimination for single and double targets

Our purpose in using small triphasic targets was to isolate the contrast signals associated with each of the two adjacent test targets in the double-target condition. Since the triphasic bars were localized in space and confined to a high-frequency band, we assumed that the increment threshold for each bar depended on the contrast signal of that bar, and not on a pooled low-frequency signal from both bars. We checked this assumption by comparing the contrast-increment threshold for a single triphasic bar on its own and in the presence of a second triphasic bar, ie the double-target configuration. If the same mechanism limits discrimination in both cases, a failure of isolation would be revealed by an increased increment threshold in the double-target configuration owing to a greater effective masking signal. All measurements were made with the right eye, while the left eye viewed the uniform background field. We compared the average increment threshold obtained separately from each member of the double-target configuration to the average threshold obtained separately from single triphasic targets presented at the same locations as each of the two targets in the double-target configuration. The results of this comparison are shown in table 1 for a pedestal contrast of 50%. There was no significant difference between the monocular increment thresholds for the single-target and double-target configurations. We thus concluded that the increment threshold for each bar in the double-target configuration was determined by a local detector or filter selectively responding to only one of the two bars.

Table 1. Monocular contrast thresholds (% contrast). Single targets: each value is the average of the increment thresholds (ΔC) measured separately for single triphasic targets presented on the left and on the right at the same locations as each of the two targets in the double-target configuration. Double target: each value is the average of the increment thresholds (ΔC) measured separately for each of the two triphasic targets. No target was present in the left eye. Pedestal contrast (C) = 50% for all targets; target duration = 200 ms.

Observer	Threshold	
	single target	double target
MB	8.1 ± 1.8	8.2 ± 0.5
DT	4.8 ± 0.9	5.1 ± 0.6
SM	7.9 ± 0.5	6.6 ± 0.8

3.2 Binocular contrast discrimination for single and double targets

In our next experiment, we repeated the increment-threshold measurements described in the section above, but now with an added triphasic target in the left eye. The left-eye target was presented at a retinal locus halfway between the positions of the two triphasic targets in the right eye; its contrast was set equal to the right-eye pedestal value (C). Although these thresholds were measured with a monocular increment (ΔC) presented to the right eye only, they were really *binocular* contrast thresholds, since they reflected, at least potentially, the combined contrast signal from the two eyes. As before, we measured thresholds for each of the two targets in the double configuration, and also for single targets presented at the locations corresponding to each of the targets in the double configuration.

The diagrams at the top of figure 3 show these target configurations. 'Binocular left double' refers to the threshold measured for an increment presented on the left member of the double configuration whereas 'binocular right double' is the increment threshold for the right member. 'Binocular left single' refers to the increment threshold for a single target occupying the position of the left member of the double-target configuration, ie the binocular contrast threshold for a single target presented at a crossed disparity of 3.5 min, while 'binocular right single' is the threshold for a single target occupying the position of the right member of the double-target configuration. The histograms below the diagrams of the target indicate the thresholds measured for the five conditions. The black bars in figure 3 show the purely monocular thresholds averaged from the two values in table 1; all the other thresholds shown in figure 3 were measured with the additional target in the left eye. If the left-eye target were

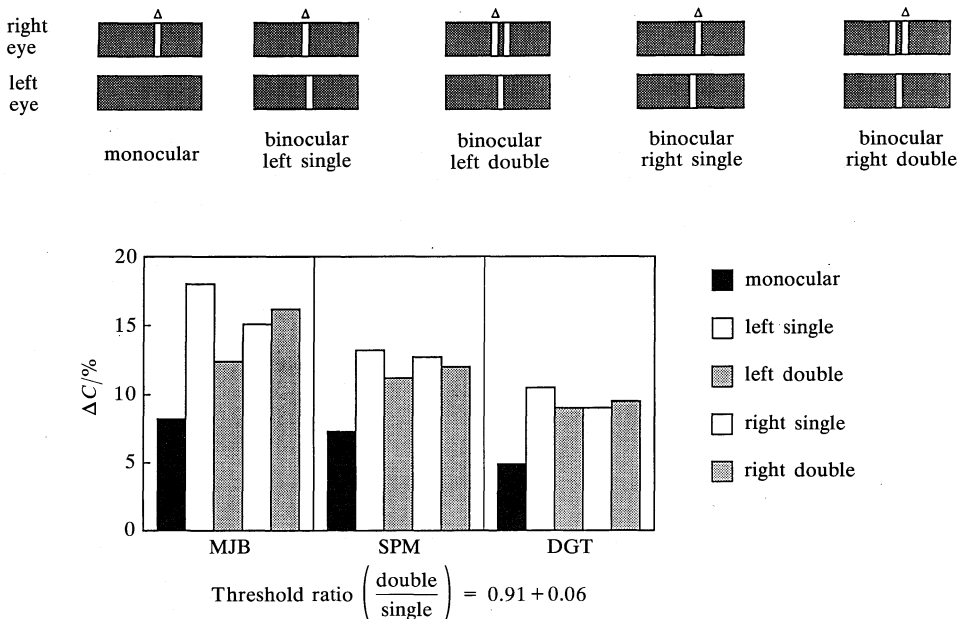


Figure 3. Each of the five bars (in sequence from left to right) in the histogram for each observer represents the contrast increment (ΔC) thresholds (% contrast) for one of the five conditions shown at the top of the figure, in the sequence corresponding to the bars in the histograms. Monocular increment thresholds, shown by black bars, were measured when the left eye viewed a uniform background. All other thresholds were measured with a single target present in the left eye. Note that the thresholds (gray bars) for the double-target configurations are nearly identical to the thresholds (white bars) for the single-target configurations, indicating that the target in the left eye masks both targets in the right eye in the double configuration. Pedestal contrast = 50%; target duration = 200 ms.

combined binocularly with the target or targets in the right eye, these increment thresholds should be higher than the purely monocular thresholds. Indeed, the binocular thresholds were almost twice the monocular thresholds for most of the conditions shown in figure 3.

Our primary interest was whether the binocular contrast thresholds were the same for the single-target and double-target configurations. Consider three possible outcomes for the double target conditions:

(i) *Preferential matching*. The brain makes only one match. It has a consistent preference for one side because the eyes are consistently converged very slightly in front of or behind the fixation plane (a sustained fixation disparity). This type of matching would produce a characteristic pattern of thresholds for the double-target configuration. The increment threshold for the target on one side would be elevated owing to binocular masking, while the threshold for the other side would be unaffected by the left-eye target and would thus equal the monocular threshold.

(ii) *Probabilistic matching*. The brain makes only one match, but the side of the match is chosen at random from trial to trial. On some trials, the increment is masked by the binocularly summed contrast from the targets in both eyes, but, on other trials, it is affected only by the monocular contrast in the right eye. If probabilistic matching occurs, the increment thresholds for the double-target condition should be consistently lower than the thresholds measured for the single-target configuration. In the simplest version of this hypothesis, the increment thresholds for the double targets should equal the average of the monocular and single-target thresholds.

(iii) *Double matching*. The brain makes two matches on every trial. In this case, the binocular thresholds for the single-target and double-target configurations should be the same.

The first possibility, preferential matching, was easily rejected. As shown in figure 3, all of the double-target thresholds were substantially higher than the monocular thresholds. The second possibility, probabilistic matching, was slightly more difficult to refute. Compare the thresholds of observer MJB for 'monocular', 'left single', and 'left double' (the first three bars of her histogram). Her 'left double' threshold was much lower than her 'left single' threshold, and was about equal to the average of her 'left single' and 'monocular' thresholds, as predicted by probabilistic matching. However, the probabilistic hypothesis failed when we considered her thresholds for the right-hand targets; her 'right double' threshold was now higher than her 'right single' threshold (last two bars of her histogram). On the whole, the results from this experiment best supported the third possibility—double matching. The ratio at the bottom of figure 3 was generated by comparing the six corresponding double and single binocular thresholds, ie 'left double' with 'left single', and 'right double' with 'right single' for each subject. The ratio of double to single thresholds is not significantly different from 1.0, which is consistent with the double-matching hypothesis.

We next systematically measured contrast-increment thresholds (ΔC) as a function of pedestal contrast (C). Figure 4 shows the thresholds from three conditions: monocular increment thresholds for the double-target configuration (labeled 'monocular'), binocular increment thresholds for the single-target configuration (labeled 'binocular single'), and binocular increment thresholds for the double-target configuration (labeled 'binocular double'). The contrast of the triphasic target in the left eye always matched the pedestal contrast of the targets in the right eye in the binocular conditions. As we found in the previous experiment, the binocular thresholds are systematically higher than the monocular thresholds, demonstrating that the added target in the left eye is masking the contrast increment in the right eye. Of more significance, the functions

for 'binocular single' and 'binocular double' are essentially the same⁽³⁾—further support for the double-matching hypothesis.

Could our experimental procedure have somehow generated evidence for double matching, despite the existence of a physiological uniqueness constraint in human stereopsis? We have been treating each of the triphasic targets shown in figure 1 as a unit—as a single high-frequency target much like the difference-of-Gaussian targets often used in other vision studies. In fact, each triphasic target is composed of three thin bars. Perhaps the stereo system treats each bar of the triphasic target as a separate feature, and matches each bar in one eye to the nearest appropriate bar in the other eye. In the double-target configuration, the individual bars of the left-eye single target could be matched to individual bars of both triphasic targets in the right eye. For example, the left-hand flanking bar of the single target in the left eye could be matched to the right-hand flanking bar of the left target in the right eye, while the right-hand flank of the single target could be matched to the left-hand flank of the right target. The point of this digression is that components of the single triphasic target in the left eye could mask components of both triphasic targets in the right eye. However, if uniqueness constrains human stereo matching, three of the six bars in the right eye would not have a match. As shown in figure 2, the contrast increment is distributed uniformly across the triphasic target; the center gets brighter and the flanks get darker when the increment is added to the contrast pedestal. Even if the stereo system is matching the individual bars of the left-eye target to individual bars

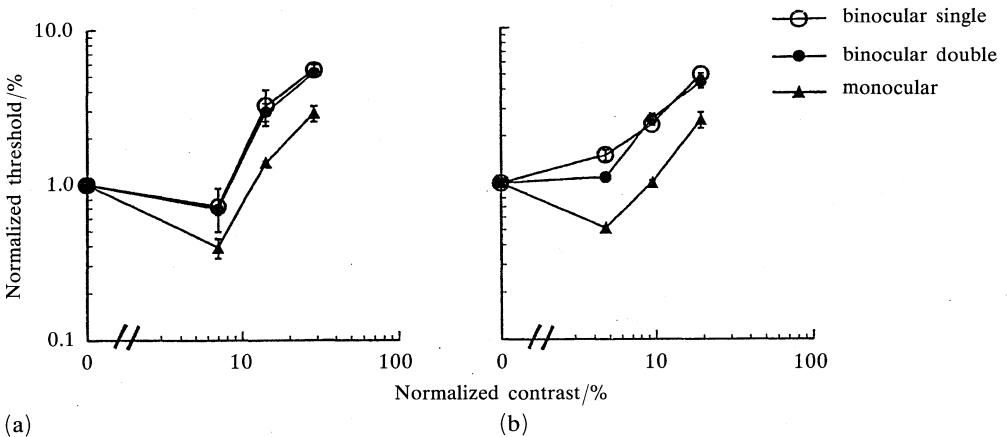


Figure 4. Contrast-increment thresholds for observers DGT (a) and SPM (b) as a function of the pedestal contrast for three conditions: monocular (two targets in right eye, no target in left eye); binocular single (single target in both left and right eye); binocular double (single target in left eye, two targets in right eye). Pedestal contrast (C) for the single triphasic bar in the left eye was always matched to pedestal contrast of the target or targets in the right eye. Thresholds (ΔC) and pedestal contrast (C) have been divided by the absolute threshold for each subject, ie, 'normalized'.

⁽³⁾ For subject SPM, the normalized increment threshold (1.1) for 'binocular double' is significantly below the threshold (1.5) for 'binocular single' at a normalized pedestal contrast of 5%. While this difference could be viewed as support for probabilistic matching, we think that there is another explanation. At these faint contrast levels, subject SPM found she could better detect the incremental change in one target by comparing it with the pedestal contrast of the adjacent target—a strategy that could only be used for the double targets. She used the same strategy for her monocular thresholds. At this low contrast, her normalized monocular threshold was 0.5 for the double target, but was 0.8 for the single target.

of both targets in the right eye, the contrast increment added to the unmatched bars would not be masked. Given a uniqueness constraint, bar-by-bar matching still predicts better contrast detection for the double-target configuration than for the single-target configuration—contrary to our results.

We considered one other extreme explanation for our results. Recall that the observers were told which of the two targets in the double configuration contained the incremental change. Perhaps the observers converged slightly in the direction of the designated test target on every trial, producing a single match with the tested side, and thus the same amount of binocular masking as for the single-target configuration. Or perhaps attention, *per se*, forced a match with the tested side. In this convoluted scenario, knowing which side of the double target contained the increment increased the observer's threshold; if the test side were unknown, matching would be probabilistic, and the observer's thresholds for the double configuration would be better than for the single-target configuration. As a control for convergence or attention, we presented the increment at random on the left or right side of the double-target configuration, and the observer judged whether the increment was present or not. Table 2 shows binocular increment thresholds for a single target, double targets with side known, and double targets with side randomized. Randomizing the location of the increment did not improve the thresholds for the double-target configuration. Predictably, the uncertainty increased thresholds. Convergence does not explain the equivalence of the binocular contrast thresholds for the single-target and double-target conditions.

Table 2. Binocular contrast thresholds (% contrast). Each value is the average of thresholds for the left and right positions. A single target was presented to the left eye in all conditions. In the single-target condition, only a single target was presented to the right eye. In the double-target, increment-known condition, the observer knew which of two targets in the right eye contained the increment for a given block of trials. In the double-target, increment-random condition, the increment was presented at random from trial to trial on the left or right member of the double target in the right eye.

Observer	Threshold		
	single target	double target, increment known	double target, increment random
DGT	9.7±0.7	9.3±0.3	11.8±0.7
SPM	13.0±0.3	11.6±1.1	14.7±1.6

3.2 *Depth discrimination*

Thus far, our results indicate that uniqueness is not a constraint on binocular contrast summation. The contrast from a single target in one eye can be combined with the contrasts of two similar targets in the other eye, provided that the disparities are small and that the contrasts of the targets in the two eyes are comparable in magnitude. Most computational models of stereo matching are composed of at least two stages, a preliminary stage in which all potential matches for every feature are identified, and a secondary stage in which the best match for each feature is chosen and other competing matches are eliminated. If the human matching process is similar to the computational models, binocular contrast summation may occur at the preliminary stage before a biological 'uniqueness constraint' has removed alternative matches. Our previous study (McKee et al 1994), showing that dichoptic masking is contingent on stereo matching, does argue against this arrangement. However, to check our current conclusion, we asked observers to judge whether the designated test target

was in front of or behind a probe target. As a parallel to the contrast measurement, we also measured the depth of single targets presented at a disparity corresponding to the disparity of each target in the double configuration. Once again, consider the possible outcomes for the double target configuration:

(i) *Preferential matching.* The single target in the left eye is consistently matched with one of the two targets in the right eye. The other target is an unmatched monocular feature that appears at the depth of the fixation plane. Figure 5a shows the psychometric functions that could theoretically result from this type of matching. In this graph, the left-eye target has been consistently matched with the right side of the double-target configuration in the right eye, so the hypothetical psychometric function is centered on an uncrossed disparity of 3.5 min. Of course, a real observer may exhibit a preference for the left side of the double target, producing a psychometric function centered on or near 3.5 min of crossed disparity. Whatever the preference, the psychometric function for the other target should appear roughly centered on the fixation plane.

(ii) *Probabilistic matching.* The single target in the left eye is matched at random to one of the two targets in the right eye. On half the trials, each of the two targets is matched at the appropriate disparity and on the other half, each is unmatched, and defaults to the depth of the fixation plane. The sinuous curves in figure 5b show the theoretical psychometric functions that would result from this type of matching; they were generated by combining a steep function centered on the fixation plane with an equally steep function centered on either +3.5 or -3.5 min, ie a combination of the curves in figure 5a. For a real observer, the psychometric functions would undoubtedly be very shallow, covering a 6–8 min range and extending across the fixation plane.

(iii) *Double matching.* The single target in the left eye is matched to both targets on every trial. This type of matching would theoretically produce the steep psychometric functions, separated by 7 min of disparity, shown in figure 5c.

The data obtained from two observers are shown in figure 6. The curves drawn in each figure are the cumulative normal functions fitted by probit analysis to the depth judgments for each of the targets in the double-target configuration (open circles). The depth judgments for the single targets are shown by filled circles. The steep functions of observer SPM resemble the predicted functions shown in figure 5c for double matching, but they are separated by less than 7 min. The uncrossed match is just in front of the predicted position, but the crossed match is substantially behind

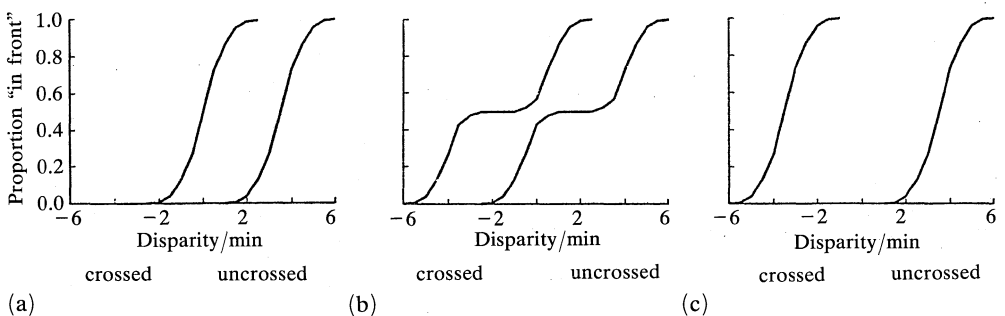


Figure 5. Hypothetical psychometric functions for observers SPM (a) and HSS (b) for depth discrimination judgments of each target in the double-target configuration. (a) Predicted functions if the observer always matches the target in the left eye to the same member of the pair of targets in the right eye (preferential matching). (b) Predicted functions if the observer matches the left-eye target to one member of the pair in the right eye chosen at random from trial to trial (probabilistic matching). (c) Predicted outcome if the observer matches the left-eye target to both members of the pair on every trial (double matching).

the predicted position (the predicted positions are shown by arrows under the abscissae in figure 6). Note, however, that her data for the corresponding single targets are indistinguishable from her data for the double targets. As shown by the median values in table 3, there was no difference between the perceived depths of the targets in the single-target and double-target configurations for this observer.

A similar pattern is apparent in the data of observer HSS (see table 3). For the crossed matches, the perceived depth of the double target is identical to the perceived depth of the corresponding single target. For the uncrossed matches, the psychometric function for the double target was somewhat shallow and it appeared slightly in front of the position occupied by the corresponding single target. Still, there was no evidence that this observer perceived either member of the double target in the fixation plane on a significant proportion of trials. In short, the depth-discrimination data from both observers best support the double-matching hypothesis.

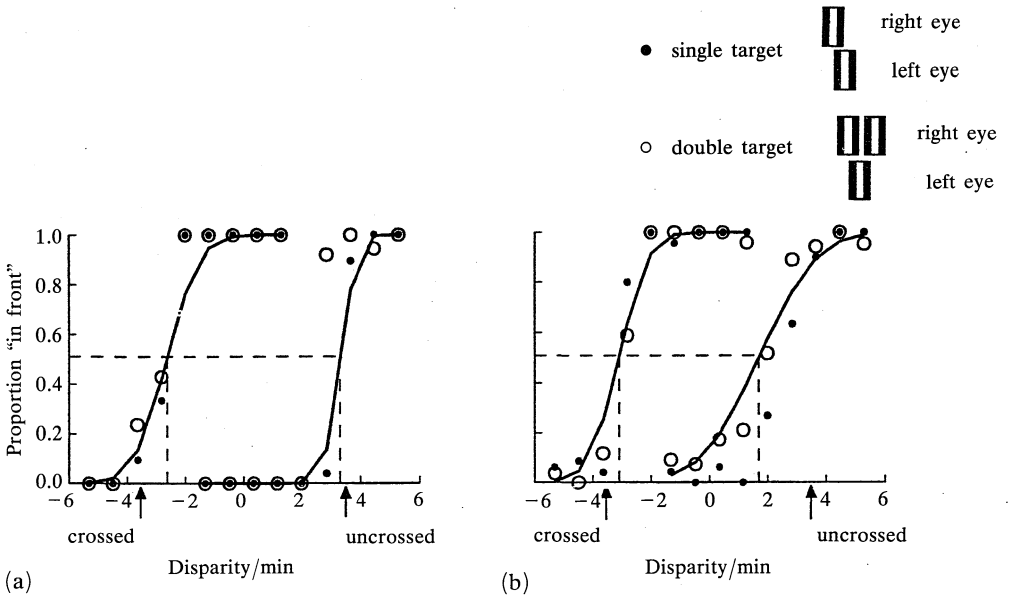


Figure 6. Psychometric functions for depth discrimination for observers SPM (a) and HSS (b): Observers judged whether the target was in front of or behind an adjacent reference target (not shown). The depth-discrimination data are essentially the same for both the single-target and the double-target configuration. See table 3 for median values for each condition and subject. Target duration 200 ms. The arrows indicate the predicted positions of the crossed and uncrossed matches.

Table 3. Medians (50th percentiles) in min of psychometric functions generated by depth-discrimination judgments. The table shows a comparison between median depth for each member of the double target, and for single targets presented at the same retinal locus as each of the double targets. Values were estimated from cumulative normal functions fitted to each data set (see figure 6) by probit analysis; each curve is based on 200 trials.

Observer	Depth/min			
	crossed single	crossed double	uncrossed single	uncrossed double
SPM	-2.83 + 0.10	-2.64 + 0.15	3.31 + 0.08	3.31 + 0.11
HSS	-3.20 + 0.13	-3.11 + 0.14	2.29 + 0.24	1.70 + 0.18

4 Discussion

Our results demonstrate that the human stereo system is not constrained to make unique matches. Evidence for double matching has also been obtained from experiments involving random-dot stereograms (Braddick, cited as personal communication in Marr 1982; Weinshall 1989, 1991). These stereograms were constructed of multiple representations of Panum's limiting case such that each point in one half image was paired to two or more points in the other half image; observers reported seeing multiple transparent surfaces. Pollard and Frisby (1990) argued that the transparency seen in these circumstances was not a violation of uniqueness. They claimed that locally each point in one half image was matched to only one point in the other half image, but that in different parts of the stereogram, different matches were selected, creating the appearance of two lacy planes. While their explanation is plausible, it is not supported by local analysis of depth matching. Shimojo and Nakayama (1994) placed a monocular probe in a sparse random-dot stereogram, and showed that, in the absence of other matching constraints (occlusion), the probe was matched to the closest horizontal neighbor in the other eye.

The failure of the uniqueness constraint in Panum's limiting case creates a conundrum for models of human stereopsis. Consider the stereogram shown at the top of figure 7. It is, of course, the experimental stimulus shown in figure 1a with an extra target in the left eye; the pair in the left eye is offset from the pair in the right eye by a small crossed disparity. Predictably, observers viewing this stereogram see only two targets located slightly in front of the fixation plane. If the human stereo system accepts double matches in Panum's limiting case, why do we not see 'ghost' matches, at the intersection of the depth map, when we view the two-pair stereogram (Cogan 1978)?

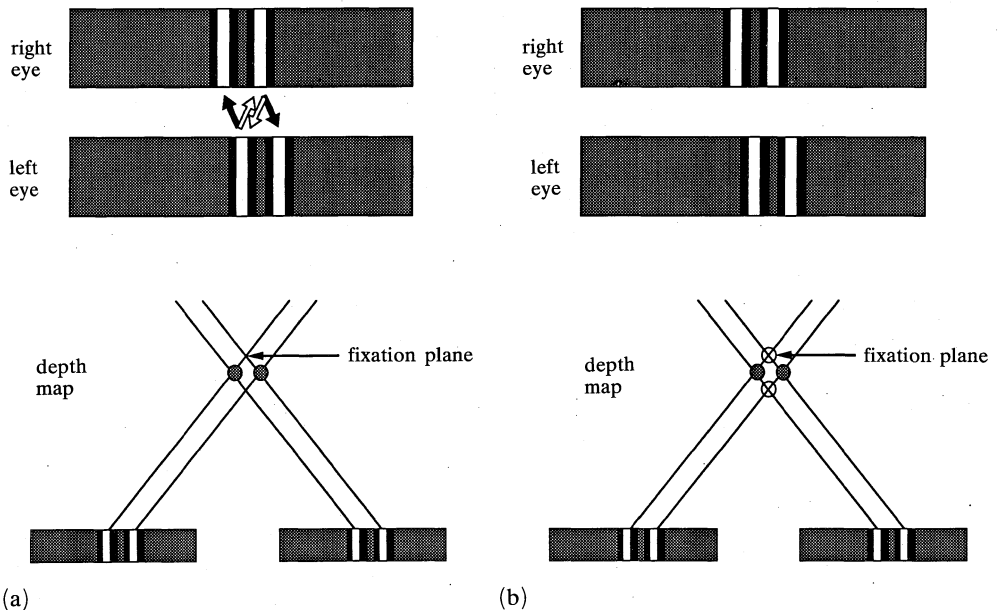


Figure 7. Diagram of the two-pair stereogram. The stereogram is identical to the stereogram shown in figure 1a except that an extra target has been added to the left eye so that each eye is presented with a pair of targets. Observers perceive two targets in the same plane. The conundrum for stereo matching is why the brain permits double matching in Panum's limiting case but enforces uniqueness for the two-pair stereogram. Two solutions are shown (a) Grimson's solution and (b) local correlation (see text for description).

Grimson (1981) proposed a simple computational solution to this puzzle. In his matching algorithm, the process of identifying appropriate matches for each feature was separately initiated from each half image. For every feature in the left half image, the algorithm searched for the best possible match within a circumscribed region of the right half image. A similar search was simultaneously performed for each feature in the right half-image. There was no constraint on double matching for a feature in either half image alone, so a feature in the left half image could be matched to two features in the right half image and vice versa. The uniqueness constraint was imposed only at the stage where the separate depth maps for the two half images were compared. At this stage, a feature could not have two matches if one of the chosen matches itself had two valid matches. In the diagram at the top of figure 7a, the left target in the left eye has made two matches, to both targets in the right eye, consistent with our data. The right target in the right eye has also made two matches, but one of its matches is 'taken' by the matching process initiated from the left eye. Since both targets have an alternative match, the overlapping choice is eliminated; the open double arrows shown in the center of figure 7a represent the forbidden match. Grimson's solution permitted double matching in Panum's limiting case, but prevented double matching for the two-pair stereogram in figure 7. This solution is easily implemented in computers, but it is not clear how it would be implemented in human stereo matching, given our current knowledge of the biological hardware.

Alternatively, the binocular system may have a distributed representation of all the disparity signals falling within a small region of binocular space. The chosen matching plane would be the one with the largest signal (Sperling 1970; Cogan 1978; see also the automap model described in Julesz 1971). One operation that can identify the plane with the largest signal is interocular correlation. In exploring the puzzle posed by failures of uniqueness, Weinshall (1991) proposed a local-correlation model to explain the two-pair stereogram of figure 7. Tyler (1983) had earlier made a similar suggestion for random-dot stereograms. Recently, Cormack et al (1991) have described a quantitative model of local correlation that explains the dependence of correlation thresholds on contrast, and Simmons and Hawken (1993) have used interocular correlation of luminance energies to predict stereoacuity for random-dot stereograms.

A local-correlation approach has the advantage of improving the signal-to-noise ratio in dense stereograms where there are many potential false matches. The disadvantage is that the disparity of individual features is lost in the pooled disparity signal of the surroundings unless the correlation is performed over quite small areas, where the benefits of correlation are necessarily smaller. Identifying 'unmatched' features (zero correlation) is also difficult if the correlation is performed over large areas. As a compromise, a local correlation could be separately computed in mechanisms tuned to different spatial scales (different spatial-frequency bandwidths) with the size of the correlation area scaling with the mechanism (Smallman and MacLeod 1994), much like the elaborated Reichardt correlators proposed for motion processing (van Santen and Sperling 1985). This hybrid approach maintains the benefits of high signal-to-noise ratios at the coarse scales, while preserving detailed disparity information at the fine scales for less-noisy conditions. There remains the question of how the scaled mechanisms would interact to produce a single solution in these circumstances.

The results of a multiscaled correlation would not be much different from models in which explicit matches of individual features are made on the basis of the combined output from scaled spatial filters (Marr and Poggio 1979; Grossberg 1987; Jones and Malik 1991). In the Jones-Malik model, a 'winner-take-all' rule is used to select the best match. What happens if there are two 'winners' as in Panum's

limiting case? The computational model could enforce unique matching, but apparently the brain does not. If double matches are permitted, the problem posed by the two-pair stereogram in figure 7 reappears. While Grimson's solution is one way out of the difficulty, a 'continuity constraint' that facilitates matches in the same frontal-parallel plane would also work (Marr and Poggio 1976). In fact, a continuity constraint may have much the same benefit as local correlation.

Obviously, a multifaceted experimental attack is needed to determine how the brain solves the matching problem. Here we have shown that two different measurements (binocular contrast masking and depth discrimination) produced consistent evidence of double matching. Elsewhere we have argued that stereo matching has a contrast threshold, one that depends on the contrast of the matching feature in the other eye (McKee et al 1994). Taken together, these results suggest that binocular contrast summation and stereo matching depend on a common cortical pathway. If this speculation is correct, then the abundant data and sophisticated models of contrast encoding may be useful guides for creating an adequate model of human stereo matching. It is the strategy that we are currently pursuing.

Acknowledgements. This research was supported by federal funding received from the National Institutes of Health—National Eye Institute (NEI) Grant RO1-EY06644 and NEI Core Grant 5P-30-EY01186. Gordon Legge's participation in this research was supported by NEI grant RO1-EY02857. Harvey Smallman received support from the Wattis Fund of the Smith-Kettlewell Eye Research Institute.

References

- Bravo M J, Blake R, 1992 "The contributions of figure and ground textures to segmentation" *Vision Research* **32** 1793–1800
- Burton G J, 1973 "Evidence for non-linear response processes in the visual system from measurements on the thresholds of spatial beat frequencies" *Vision Research* **13** 1211–1225
- Chang J J, Anderson B L, Julesz B, 1993 "Panum's limiting case exhibits greatest stability and clearest depth when consistent with the geometry of occlusion" *Investigative Ophthalmology and Visual Science, Supplement* **34** 1438
- Cogan A L, 1978 "Fusion at the site of the 'ghosts'" *Vision Research* **18** 657–664
- Cormack L K, Stevenson S B, Schor C M, 1991 "Interocular correlation, luminance contrast and cyclopean processing" *Vision Research* **31** 2195–2208
- Elliott P B, 1964 "Tables of d' ", in *Signal Detection and Recognition by Human Observers* Ed. J A Swets (New York: John Wiley) pp 651–684
- Frisby J P, Pollard S B, 1991 "Computational issues in solving the stereo correspondence problem", in *Computational Models of Visual Processing* Eds M S Landy, J A Movshon (Cambridge, MA: MIT Press) pp 331–358
- Foley J M, Legge G E, 1981 "Contrast detection and near-threshold discrimination in human vision" *Vision Research* **21** 1041–1053
- Gettys C F, Harker G S, 1967 "Some observations and measurements of the Panum phenomenon" *Perception & Psychophysics* **2** 387–395
- Grimson W E L, 1981 *From Images to Surfaces: A Computational Study of the Human Early Visual System* (Cambridge, MA: MIT Press)
- Grossberg S, 1987 "Cortical dynamics of three-dimensional form, color, and brightness perception: II. Binocular theory" *Perception & Psychophysics* **41** 117–158
- Helmholtz H von, 1910 *Physiological Optics* volume 3 (New York: Dover, 1962); English translation by J P C Southall for the Optical Society of America (1925) from the 3rd German edition of *Handbuch der Physiologischen Optik* (Hamburg: Voss, 1910)
- Howard I P, Ohmi M, 1992 "A new interpretation of the role of dichoptic occlusion in stereopsis" *Investigative Ophthalmology and Visual Science, Supplement* **33** 1370
- Jones D G, Malik J, 1991 "A computational framework for determining stereo correspondence from a set of linear spatial filters", report UCB/CSD 91/655, Computer Science Division (EECS), University of California, Berkeley, CA
- Julesz B, 1971 *Foundations of Cyclopean Perception* (Chicago, IL: University of Chicago Press)
- Kaufman L, 1976 "On stereopsis with double images" *Psychologia* **19** 224–233
- Legge G E, 1984 "Binocular contrast summation—I. Detection and discrimination" *Vision Research* **24** 373–383

-
- Levi D M, Harwerth R S, Smith E L, 1979 "Humans deprived of normal binocular vision have binocular interactions tuned to size and orientation" *Science* **206** 852-854
- McKee S P, Bravo M J, Taylor D G, Legge G E, 1994 "Stereo matching precedes dichoptic masking" *Vision Research* **34** 1047-1060
- Marr D, 1982 *Vision* (San Francisco, CA: W H Freeman)
- Marr D, Poggio T, 1976 "Cooperative computation of stereo disparity" *Science* **194** 283-287
- Marr D, Poggio T, 1979 "A theory of human stereopsis" *Proceedings of the Royal Society of London, Series B* **204** 301-328
- Nakayama K, Shimojo S, 1990 "Da Vinci stereopsis: depth and subjective occluding contours from unpaired image points" *Vision Research* **30** 1811-1825
- Ogle K N, 1962 "The optical space sense", in *The Eye* volume 4, Ed. H Davson (New York: Academic Press) pp 374-375
- Ono H, Shimono K, Shibuta K, 1992 "Occlusion as a depth cue in the Wheatstone-Panum limiting case" *Perception & Psychophysics* **51** 3-13
- Pollard S B, Frisby J P, 1990 "Transparency and the uniqueness constraint in human and computer stereo vision" *Nature (London)* **347** 553-556
- Shimojo S, Nakayama K, 1994 "Interocularly unpaired zones escape local binocular matching" *Vision Research* **34** 1875-1882
- Simmons D R, Hawken M J, 1993 "Interocular correlation as a predictor of stereoacuity" *Investigative Ophthalmology and Visual Science* **34** 1189
- Smallman H S, MacLeod D I A, 1994 "Size-disparity correlation in stereopsis at contrast threshold" *Journal of the Optical Society of America A* **11** 2169-2183
- Sperling G, 1970 "Binocular vision: A physical and neural theory" *American Journal of Psychology* **83** 461-534
- Tyler C W, 1983 "Sensory processing of binocular disparity", in *Vergence Eye Movements: Basic and Clinical Aspects* Eds C M Schor, K J Ciuffreda (Sevenoaks, Kent: Butterworth) pp 199-295
- Van Santen J P H, Sperling G, 1985 "Elaborated Reichardt detectors" *Journal of the Optical Society of America A* **2** 300-321
- Weinshall D, 1991 "Seeing 'ghost' planes in stereo vision" *Vision Research* **31** 1731-1748
- Westheimer G, 1986 "Panum's phenomenon and the confluence of signals from the two eyes in stereoscopy" *Proceedings of the Royal Society of London, Series B* **228** 289-305
- Wheatstone C, 1838 "On some remarkable, and hitherto unobserved, phenomena of binocular vision" *Philosophical Transactions of the Royal Society of London* **128** 371-394; reprinted in *Visual Perception: The Nineteenth Century* Ed. W Dember (New York: Wiley, 1964) pp 371-377, 386-387