N90-22920

# SPATIAL CONSTRAINTS OF STEREOPSIS IN VIDEO DISPLAYS

Clifton Schor
University of California
School of Optometry
Berkeley, California

Recent development in video technology, such as the liquid crystal displays and shutters, have made it feasible to incorporate stereoscopic depth into the three-dimensional representations on two-dimensional displays. However, depth has already been vividly portrayed in video displays without stereopsis using the classical artists' depth cues described by Helmholtz (1866) and the dynamic depth cues described in detail by Ittleson (1952). Successful static depth cues include overlap, size, linear perspective, texture gradients, and shading. Effective dynamic cues include looming (Regan and Beverly, 1979) and motion parallax (Rogers and Graham, 1982).

Stereoscopic depth is superior to the monocular distance cues under certain circumstances. It is most useful at portraying depth intervals as small as 5-10 arc seconds. For this reason it is extremely useful in user-video interactions such as in telepresence. Objects can be manipulated in 3-D space, for example, while a person who controls the operations views a virtual image of the manipulated object on a remote 2-D video display. Stereopsis also provides structure and form information in camouflaged surfaces such as tree foliage. Motion parallax also reveals form; however, without other monocular cues such as overlap, motion parallax can yield an ambiguous perception. For example, a turning sphere, portrayed as solid by parallax, can appear to rotate either leftward or rightward. However, only one direction of rotation is perceived when stereo-depth is included. If the scene is static, then stereopsis is the principal cue for revealing the camouflaged surface structure. Finally, dynamic stereopsis provides information about the direction of motion in depth (Regan and Beverly, 1979). When optical flow patterns seen by the two eyes move in phase, field motion is perceived in the fronto-parallel plane. When optical flow is in antiphase (180°) motion is seen in the saggital plane. Binocular phase disparity of optical flow as small as 1° can be discriminated as changes in visual direction of motion in a 3-D space (Beverly and Regan, 1975). This would be a useful addition to the visual stimuli in flight simulators.

Several spatial constraints need to be considered for the optimal stimulation of stereoscopic depth. The stimulus for stereopsis is illustrated in figure 1. Each peg subtends a visual angle at the entrance pupils of the eyes, and this angle is referred to as binocular parallax. The difference in this angle and the angle of convergence forms an absolute disparity. In the absence of monocular depth cues, perceived distance of an isolated target, subtending an absolute disparity is biased toward 1.5 meters from the physical target distance. Gogle and Teitz (1973) referred to this as equidistance tendency. If the target moves abruptly from one distance to another, convergence responses signal the change of depth (Foley and Richards, 1972); however, smooth continuous changes in binocular parallax, tracked by vergence eye movements do not cause changes in perceived distance (Erkelens and Collewijn, 1985; Guttmann and Spatz, 1985). Once more than one disparate feature is presented in the field, differences in depth (stereopsis), stimulated by retinal image disparity become readily apparent. Stereothresholds may be as low as 2 sec arc, which ranks stereopsis along with vernier and bisection tasks among the hyperacuities.

Stereo-sensitivity to a given angular depth interval varies with the saggital distance of the stimulus depth increment from the fixation plane. Sensitivity to depth increments is highest at the horopter or fixation plane where the disparity of one of the comparison stimuli is zero (Blakemore, 1970). This optimal condition for stereopsis was used by Tschermack (1930) as one of four criteria for defining the empirical longitudinal horopter. The Weber fraction describing the ratio of increment stereothreshold (arc sec) over the disparity pedestal (arc min) (3 sec/min) is fairly constant with disparity pedestal amplitudes up to 1°. This fraction was derived from figure 2, which plots stereothreshold in seconds of arc at different saggital distances in minutes arc from the fixation point for targets consisting of vertical bars composed of coarse or fine features. A two-alternative, forced choice is used to measure a just-noticeable difference between a depth increment between an upper test bar and a lower standard bar, both seen at some distance before or behind the fixation plane. The bar used was a narrow-band, spatially filtered line produced from a difference of Gaussians (DOG) whose center spatial frequency ranges from 9.5 to 0.15 cycles/deg (Badcock and Schor, 1985). When these thresholds are plotted, the slopes of these functions found with different width DOGs are the same on a logarithmic scale. However, thresholds for low spatial frequencies (below 2.5 cpd) are elevated by a constant disparity which illustrates they are a fixed multiple of thresholds found with higher spatial frequencies. These results illustrate that depth stimuli should be presented very near the plane of fixation, which is the video screen.

Stereo-sensitivity remains high within the fixation plane over several degrees about the point of fixation. Unlike the rapid reduction of stereo-sensitivity with overall depth or saggital distance from the horopter, stereo-sensitivity is fairly uniform and at its peak along the central 3° of the fixation plane (Blakemore, 1970; Schor and Badcock, 1985). Figures 2 and 3 illustrate a comparison of stereo-depth increment sensitivity for this fronto-parallel stereo and the saggital off-horopter stereothreshold. Also plotted in figure 3 are the monocular thresholds for detecting vernier offset of the same DOG patterns at the same retinal eccentricities. Clearly, stereopsis remains at its peak at eccentricities along the horopter and there is a percipitous fall of visual acuity (Wertheim, 1894) and, as shown here, of vernier acuity over the same range of retinal eccentricities where stereo increment sensitivity is unaffected (Schor and Badcock, 1985). Thus, stereoacuity is not limited by the same factors that limit monocular vernier acuity because the two thresholds differ by a factor of 8 at the same eccentric retinal locus.

In addition to the threshold or lower disparity limit (LDL) for stereopsis, there is an upper disparity limit (UDL), beyond which stereo depth can no longer be appreciated. This upper limit is small, being approximately 10 arc min with fine (high-frequency) targets, and somewhat larger (several degrees) with coarser (low spatial frequency) fusion stimuli (Schor and Wood, 1983). This depth range can be extended either by briefly flashing targets (Westheimer and Tanzman, 1956) or by making vergence movements between them (Foley and Richards, 1972) to a UDL of approximately 24°. The UDL presents a common pitfall for many stereo-camera displays that attempt to exaggerate stereopsis by placing the stereo-cameras far apart. Paradoxically, this can produce disparities that exceed the UDL and results in the collapse of depth into the fronto-parallel plane.

Diplopia is another problem that accompanies large disparities. The diplopia threshold is slightly smaller than the UDL for static stereopsis, and depth stimulated by large flashed disparities is always seen diplopically. Normally, this diplopia can be minimized by shifting convergence from one target to another. However, this is not as easily done with a stereo-video monitor. In real space the stimulus for vergence is correlated with the stimulus for accommodation. With video displays, the stimulus for accommodation is fixed at the screen plane while vergence is an independent variable. Because there is cross-coupling between accommodation and vergence, we are not completely free to dissociate these motor responses (Schor and Kotulak, 1986). With some muscular effort, a limited degree of vergence can be expected while accommodation is fixed, depending on the accommodative-convergence ratio (AC/A). When this ratio is high, a person must choose between clearness and singleness.

Additional problems for stereoscopic depth occur with abstract scenes containing high spatial frequency surface texture. This presents an ambiguous stimulus for stereopsis and fusion which can have an enormous number of possible solutions as illustrated by the wallpaper illusion or by a random-dot stereogram. The visual system uses various strategies to reduce the number of potential fusion combinations and certain spatial considerations of targets presented on the visual display can help implement these strategies. A common technique used in computer vision is the coarse-to-fine strategy. The visual display is presented with a broad range of spatial frequency content. The key idea here is that there is little confusion or ambiguity with coarse features like the frame of a pattern. These can be used to guide the alignment of the eyes into registration with finer features that present small variations in retinal image disparity. Once in registration, small disparities carried by the fine detail can be used to reveal the shape or form of the depth surface. An essential condition for this algorithm to work is that sensitivity to large disparities be greatest when they are presented with coarse detail and that sensitivity to small disparities be highest with fine (high spatial frequency) fusion stimuli. This size-disparity correlation has been verified for both the LDL and UDL by Schor and Wood (1983). Figure 4 illustrates the variation of stereo-threshold (LDL) and the UDL with spatial frequency for targets presented on a zero disparity pedestal at the fixation point. Stereothresholds are lowest and remain relatively constant for spatial frequencies above 2.5 cycles/deg. Thresholds increase proportionally with lower spatial frequencies. Even though stereothreshold varies markedly with target coarseness, suprathreshold disparities needed to match the perceived depth of a standard disparity are less dependent on spatial frequency. This depth equivalence constitutes a form of stereo-depth constancy (Schor and Howarth, 1986). Similar variations in the diplopia threshold or binocular fusion limit are found by varying the coarseness of fusion stimuli (Schor, Wood, and Ogawa, 1984b).

Figure 5 illustrates that the classical vertical and horizontal dimensions of Panum's fusion limit (closed and open symbols, respectively) are found with high spatial frequency targets, but the fusion limit increases proportionally with the spatial width of targets at spatial frequencies lower than 2.5 cycles/deg. When measured with high-frequency DOGs, the horizontal radius of PFA (Panum's fusional area) is 15 min; and when measured with low-frequency stimuli, PFA equals a 90° phase disparity of the fusion stimulus.

The increase in Panum's fusion limit appears to be caused by monocular limitations to spatial resolution. For example, if the same two targets that were used to measure the diplopia threshold are both presented to one eye to measure a two-point separation threshold, such as the Rayleigh criterion, then the monocular and binocular thresholds are equal when tested with spatial frequencies lower than 2.5 cpd. At higher spatial frequencies we are better able to detect smaller separations between two points presented monocularly than dichoptically. This difference at high spatial frequencies reveals a unique binocular process for fusion that is independent of spatial resolution. With complex targets composed of multiple spatial frequencies, at moderate disparities such as 20 min arc, a diplopia threshold may be reached with high spatial frequency components while stereopsis and fusion may continue with the low spatial frequency components. An example of this simultaneous perception can be seen with the diplopic pixils in a random dot stereogram whose coarse camouflaged form is seen in vivid stereoscopic depth (Duwaer, 1983).

In addition to target coarseness, there are several other aspects of spatial configuration that influence stereopsis and fusion. The traditional studies of stereopsis, such as those conducted by Wheatstone (1838), mainly consider the disparity stimulus in isolation from other disparities at the same or different regions of the visual field. It is said that disparity is processed locally in this limiting case, independent of other possible stimulus interactions other than the comparison between two absolute disparities to form a relative disparity. However, recent investigations have clearly illustrated that in addition to the local processes, there are global processes in which spatial interaction between multiple relative disparities in the visual field can influence both stereopsis and fusion. Three forms of global interactions have been studied. These are disparity crowding, disparity gradients, and disparity continuity or interpolation. These global interactions appear to influence phenomena such as the variation in size of Panum's fusional area, reductions and enhancement of stereo-sensitivity, constant errors or distortions in depth perception, and resolution of a 3-D form that has been camouflaged with an ambiguous surface texture.

Spatial crowding of visual targets to less than 10 arc min results in a depth averaging of proximal features. This is manifest as an elevation of stereothreshold as well as a depression of the UDL (Schor, Bridgeman, and Tyler, 1983). The second global interaction, disparity gradient, depends upon spacing between disparate targets and the difference in their disparities. (Schor and Tyler, 1981). The disparity gradient represents how abruptly disparity varies across the visual field. The effect of disparity gradients upon the sensory fusion range has been investigated with point targets by Burt and Julesz (1980), and with periodic sinuosidal spatial variations in horizontal and vertical disparity by Schor and Tyler (1981). Both groups demonstrate that the diplopia threshold increases according to a constant disparity gradient as the separation between adjacent fusion stimuli increases. Cyclofusion limits are also reduced by abrupt changes in disparity between neighboring retinal regions (Kertesz and Optican, 1974). Stereothresholds can also be described as a constant disparity gradient. As target separation decreases, so does stereothreshold, up to a limit of 15 arc min separation. Further reduction in separation results in crowding, which elevates the stereothreshold. The UDL is also limited by a constant disparity gradient (fig. 5). As spacing decreases, there is a proportional decrease in the UDL. These gradient effects set two strict limitations on the range of stereopscopic depth that can be rendered by the video display. As crowding increases, the UDL will decrease. The effect is that targets exceeding the UDL will appear diplopic and without depth. For example, a top-down picture of a forest which has trees of uneven height will not be seen as uneven depth if the trees are imaged too closely. To remedy this problem, the depth should be reduced by moving the stereocameras closer together. In the other extreme, a shallow slope will not be seen in depth unless it exceeds the gradient for stereothresholds. Even if it does, it may still not be seen if it extends across the entire visual display. Normally there can be unequal optical errors of the two eyes which produce unequal magnification of the two retinal images. This aniso magnification produces an apparent tilt of the stereoscopic frame reference referred to as the fronto-parallel plane. However, this constant depth error is normally corrected or compensated for perceptually (Morrison, 1977). This perceptual compensation could reduce sensitivity to wide static displays of a shallow depth gradient.

A third form of global interaction is observed under conditions where disparity differences between neighboring regions occur too gradually to be detected, such as in the 3-D version of the Craik-Obrien Cornsweet illusion (fig. 6 by Anstis, Howard, and Rogers, 1978), when stereo patterns are presented too briefly to be processed fully (Ramachandran and Nelson, 1976; Mitchison and McKee, 1985), or when several equally probable, but ambiguous, disparity solutions are presented in a region neighboring an unambiguous disparity solution (Kontsevich, 1986). Under all of these conditions, the depth percept resulting from the vague disparity is similar to or continuous with the depth stimulated by the more visible portion of the disparity stimulus. This illustrates the principle of depth continuity formulated by Julesz (1971) and restated later by Marr and Poggio (1979), which recently was shown by Ramachandran and Cavanaugh (1985) to include the extension of depth to subjective contours in which no physical contour or disparity exists.

Clearly there are many spatial constraints, including spatial frequency content, retinal eccentricity, exposure duration, target spacing, and disparity gradient, which—when properly adjusted—can greatly enhance stereodepth in video displays.

## REFERENCES


Anstis, S. M., Howard, J. P., and Rogers, B. (1978) A Craik-Cornsweet illusion for visual depths. Vision Res. 18: 213-217.

Badcock, D. R., and Schor, C. M. (1985) Depth-increment detection function for individual spatial channels. J. Opt. Soc. Am. A. 2: 1211-1216.

Beverley, K. I. and Regan, D. (1975) The relation between sensitivity and discrimination in the perception of motion-in-depth. J. Physiol. 249: 387398.

Blakemore, C. (1970) The range and scope of binocular depth discrimination in man. J. Physiol. 211: 599-622.

Burt, P. and Julesz, B. (1980) A disparity gradient limit for binocular fusion. Science 208: 615-617.

Duwaer, A. L. (1983) Patent stereopsis and diplopia in random-dot stereograms. Perception and Psychophysics 33: 443-454.

Erkelens, C. J. and Collewijn, H. (1985) Motion perception during dichoptic viewing of moving random-dot stereograms. Vision Res. 25: 583-588.

Foley, J. M. and Richards, W. (1972) Effects of voluntary eye movement and convergence on the binocular appreciation of depth. Perception and Psychophysics 11: 423-427.

Gogle, W. G. and Teitz, J. D. (1973) Absolute motion parallax and the specific distance tendency. Perception and Psychophysics 13: 284-292.

Guttmann, J. and Spatz, H. (1985) Frequency of fusion and of loss of fusion, and binocular depth perception with alternating stimulus presentation. Perception 14: 5-13.

Helmholtz, H. C. (1866) Handbuch der physiologische optik. Hamburg, Voas.

Ittelson, W. H. (1952) The Ames demonstration in perception. Princeton: Princeton University Press.

Ittelson, W. H. (1960) Visual space perception visual space cues: Binocular stereopsis. 123-127.

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

Kertesz, A. E. and Optican, L. M. (1974) Interactions between neighboring retinal regions during fusional response. Vision Res. 14: 339-343.

Kontsevich, L. L. (1986) An ambiguous random-dot stereogram which permits continuous change of interpretation. Vision Res. 26: 517-519.

Marr, D. and Poggio, T. (1979) A computational theory of human stereo vision. Proc. Royal Soc. 204: 301-328.

Mitchison, G. J. and McKee, S. P. (1985) Interpolation in stereoscopic matching. Nature 315# 6018: 402-404.

Morrison, L. C. (1977) Stereoscopic localization with the eyes asymmetrically converged. Am J. Optometry & Physiol. Optics 54: 556-566.

Ramachandran, V. S. and Cavanagh, P. (1985) Subjective contours capture stereopsis. Nature 314: 527-530.

Ramachandran, V. S. and Nelson, J. I. (1976) Global grouping overrides point-to-point disparities. Perception 5: 125-128.

Regan, D. and Beverley, K. I. (1979) Binocular and monocular stimuli for motion-in-depth: changing disparity and changing size inputs feed the same motion in depth stage. Vision Res. 19: 1331-1342.

Rogers, B. and Graham, M. (1982) Similarities between motion parallax and stereopsis in human depth perception. Vision Res. 22: 261 270.

Schor, C. M. and Badcock, D. R. (1985) A comparison of stereo and vernier acuity within spatial channels as a function of distance from fixation. Vision Res. 25: 1113-1119.

Schor, C. M., Bridgeman, B., and Tyler, C. W. (1983) Spatial characteristics of static and dynamic stereoacuity in strabismus. Investigative Ophthalmol. Vis. Sci. 24: 1572-1579.

Schor, C. M. and Howarth, P. A. (1986) Suprathreshold stereo-depth matches is a function of contrast and spatial frequency perception.

Schor, C. M, and Kotulak, J. (1986) Dynamic interactions between accommodation and convergence are velocity sensitive. Vision Res. 26: 927-942.

Schor, C. M. and Tyler, C. W. (1981) Spatio-temporal properties of Panum's fusional area. Vision Res. 21: 683-692.

Schor, C. M, and Wood, I. (1983a) Disparity ranqe for local stereopsis as a function of luminance spatial frequency. Vision Res. 23: 1649-1654.

Schor, C. M., Wood, I. C. and Ogawa, J. (1984a) Spatial tuning of static and dynamic local stereopsis. Vision Res. 24: 573-578.

Schor, C. M., Wood, I. and Ogawa, J. (1984b) Binocular sensory fusion is limited by spatial resolution. Vision Res. 24: 661-665.

Tschermak, A. (1930) Beitrage zur physiologishen Optik 111; Raimsinn. In Bethe, A., Bergmann, G.V., Embolen, G. and Ellinger, A: Handbuck der normalen und pathologischen physiologic. Berlin Julius Springer vol. 12, p. 2, pp. 833-1000.

Wertheim, T. (1894) Uber die indirekte sekscharfe. Z. Psychol. Physiol. Sinnesory 7: 172-187.

Westheimer, G. and Tanzman, I. J. (1956) Qualitative depth localized with diplopic images. J. Opt. Soc. Am. 46: 116-117.

Wheatstone, C. (1838) Some remarkable phenomena of binocular vision. Phil. Trans. R. Soc. 128: 371-394.

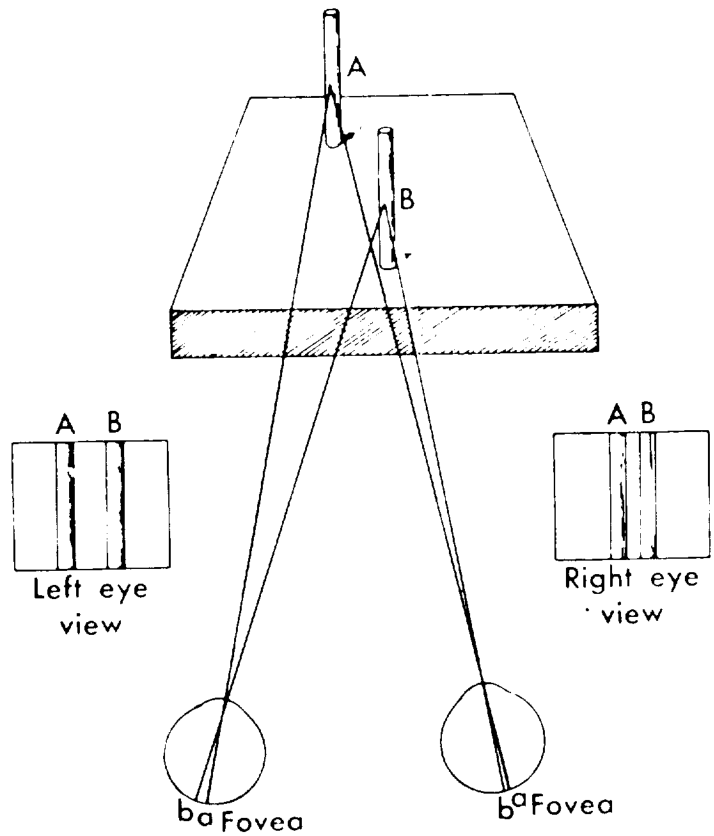


Figure 1. Retinal image disparity based on horizontal separation of the two eyes.

ORIGINAL PAGE IS
OF POOR QUALITY

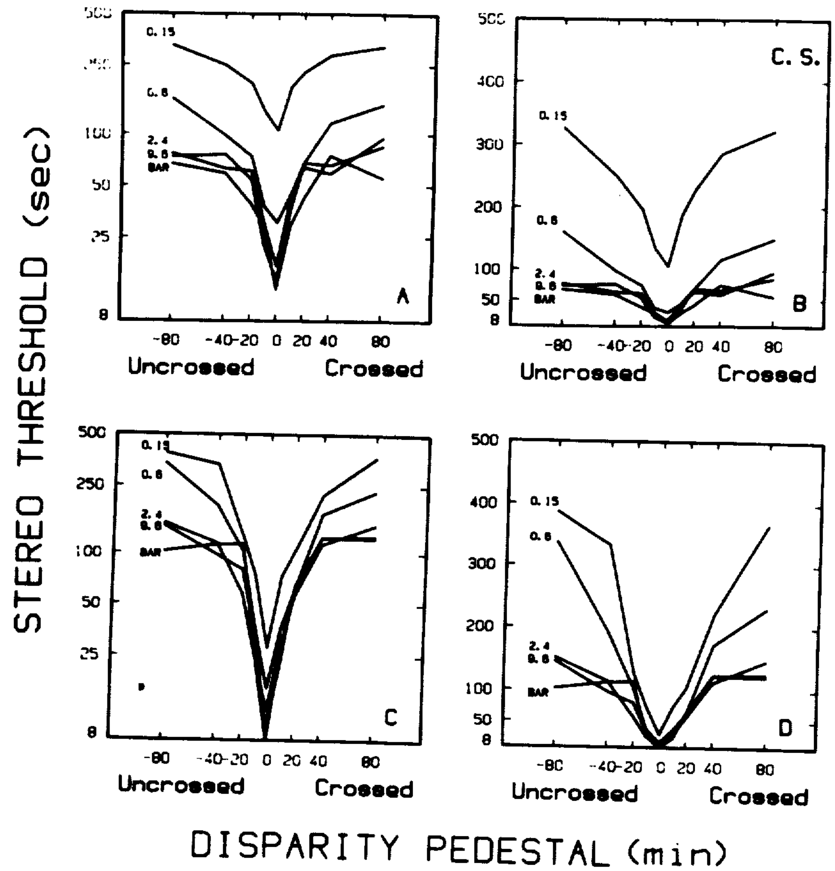


Figure 2. Threshold depth increments obtained, for observer D.B., as a function of pedestal size in both the convergent and divergent directions. Functions illustrate results obtained with a thin bar and DOGs whose center spatial frequencies ranged from 0.15 to 9.6 c/deg. Panels C and D plot the performance measured when the comparison stimulus was a thin bright bar and the test stimulus was a DOG. Panels A and B show the results obtained when a DOG was used both as a comparison and as a test stimulus. Panels A and C plot stereothreshold on a log scale. The data are replotted on a linear scale in panels B and D.

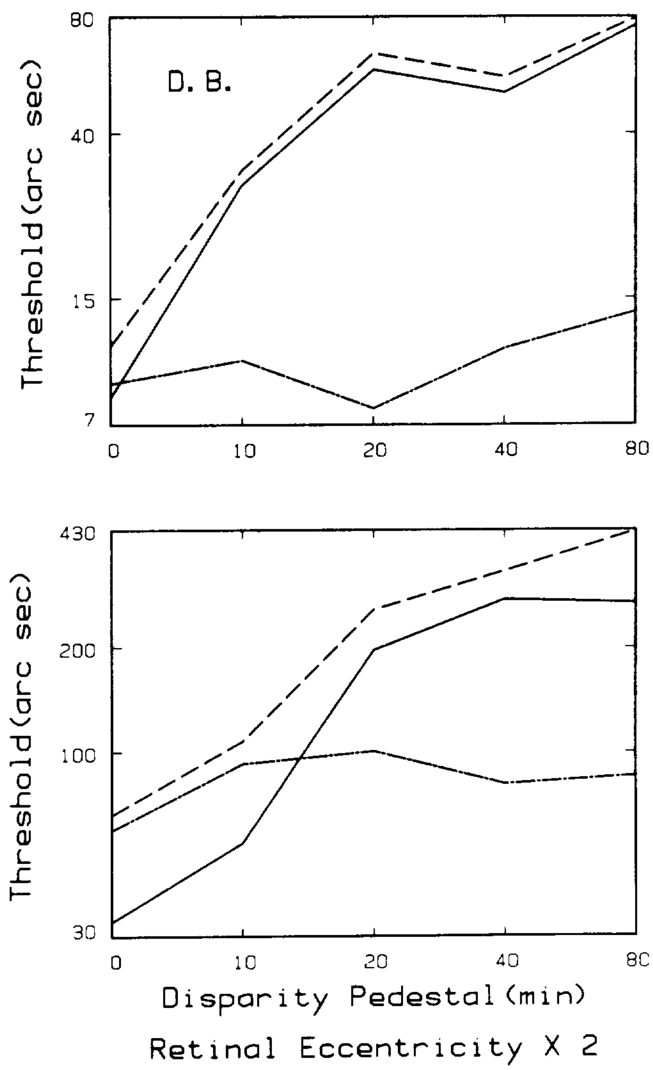


Figure 3. A comparison is made of extra-foveal vernier threshold (solid line) with extra-foveal (mixed dashed line) and extra-horopteral (long dashed line) stereothresholds for a high spatial frequency stimulus (upper plot) and a low spatial frequency stimulus (lower plot). Note that retinal eccentricity has been doubled to be comparable to disparity pedestal. Over a 40 arc min range of retinal eccentricity, stereoacuity remained unchanged and vernier acuity increased moderately. A marked increase in stereothreshold occurred over a comparable (80 arc min) disparity pedestal range.

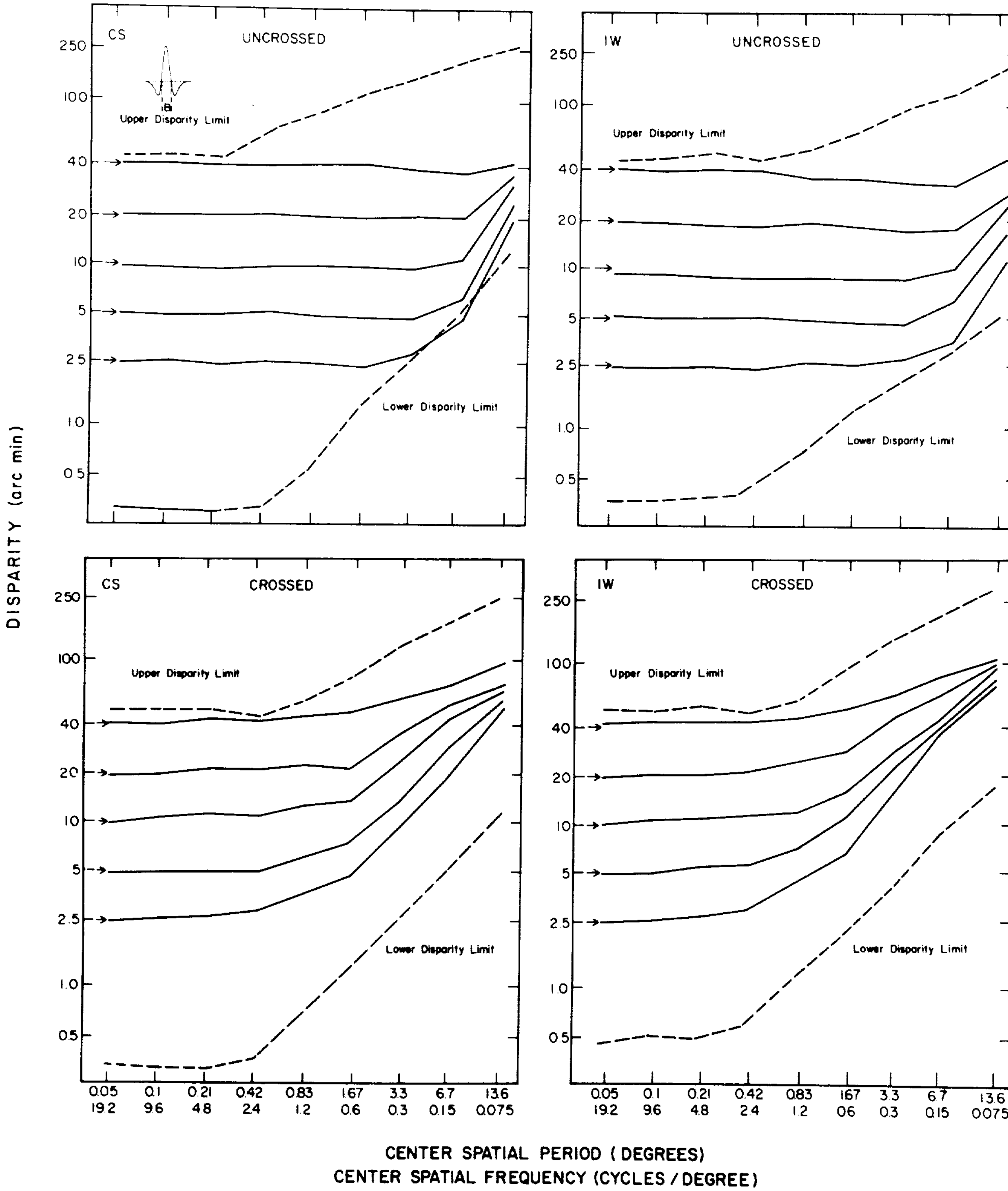


Figure 4. Upper and lower limits for stereopsis are plotted for two subjects as a function of DOG center spatial period along dashed curves at the top and bottom of data sets for uncrossed and crossed disparities respectively. Stereothreshold was lowest at small spatial periods (<0.42 arc min) and increased according to a 6° phase disparity between stereo-half images as spatial period increased. The upper limit increased proportionally to the square root of spatial period over the same range of broad spatial periods. Depth matching curves (solid lines) for several standard suprathreshold disparities (horizontal arrows) have flatter frequency responses than the upper and lower dashed threshold curves. Their breakaway point occurs at a higher spatial period for crossed than for uncrossed disparities. The luminance profile of the difference of two Gaussian functions is inset in the upper left corner.

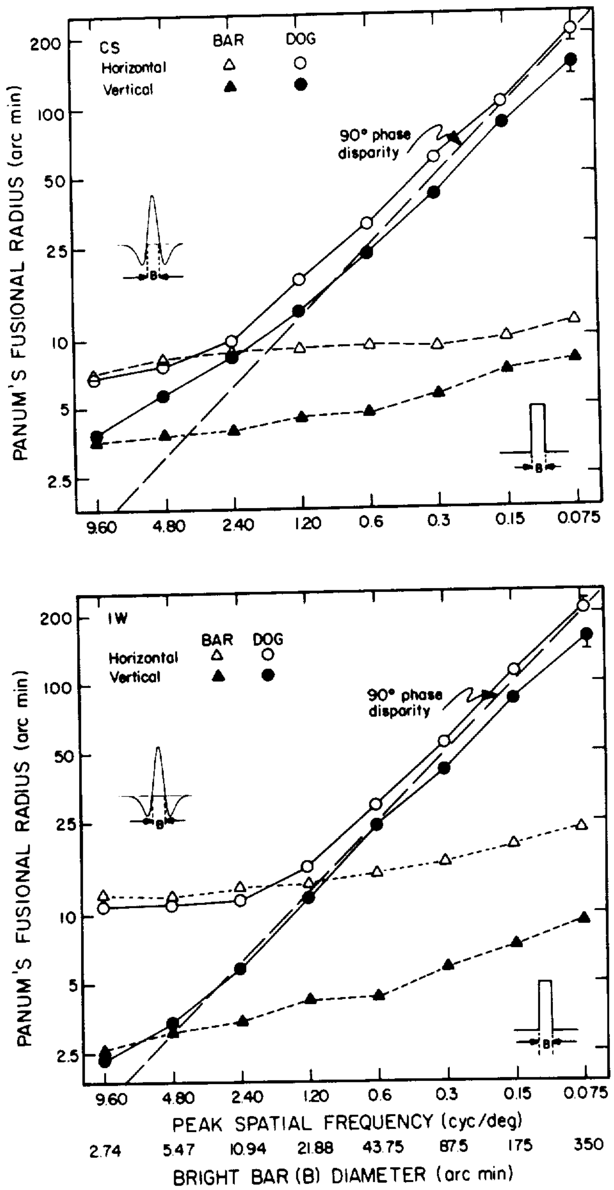


Figure 5. Diplopia thresholds for two subjects are plotted as a function of bright bar width (B) of bar and difference of two Gaussian functions (DOG). Luminance profiles of these two test stimuli are inset below and above the data respectively. A constant phase disparity of 90° is shown by the dashed diagonal line. Horizontal and vertical Panum's fusion ranges (solid lines) coincide with the 90° phase disparity for DOG widths greater than 21 arc min. At the broadest DOG width, the upper fusion limit equals the upper disparity limit for stereoscopic depth perception (bold dashed line). The standard deviation of the mean is shown for the broadest DOG stimulus. At narrow DOG widths, both horizontal and vertical fusion limits approach a constant minimum threshold. Panum's fusion ranges remain fairly constant when measured with bar patterns (dotted lines) and resemble values obtained with high spatial frequency DOGs.

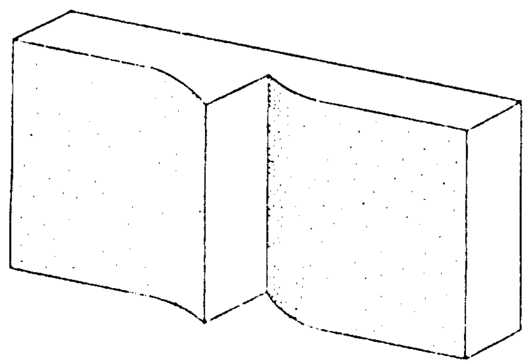

Figure 6. Perspective sketch of the illusory depth surface. Left part looks apparently nearer than the right part.