Pergamon

PII: S0042-6989(97)00097-7

Vision Res., Vol. 37, No. 19, pp. 2717–2723, 1997

Printed in Great Britain
0042-6989/97 $17.00 + 0.00

Human Stereo Matching is not Restricted to Epipolar Lines

SCOTT B. STEVENSON,*‡ CLIFTON M. SCHOR†

Received 2 September 1996; in revised form 13 February 1997

Computational approaches to stereo matching have often taken advantage of a geometric constraint which states that matching elements in the left and right eye images will always fall on "epipolar lines". The use of this epipolar constraint reduces the search space from two dimensions to one, producing a tremendous saving in the computation time required to find the matching solution. Use of this constraint requires a precise knowledge of the relative horizontal, vertical and torsional positions of the two eyes, however, and this information may be unavailable in many situations. Experiments with dynamic random element stereograms reveal that human stereopsis can detect and identify the depth of matches over a range of both vertical and horizontal disparity. Observers were able to make accurate near/far depth discriminations when vertical disparity was as large as 45 arcmin, and were able to detect the presence of correlation over a slightly larger range. Thus, human binocular matching sensitivity is not strictly constrained to epipolar lines. © 1997 Elsevier Science Ltd

Correspondence Disparity Matching Stereopsis Psychophysics

INTRODUCTION

The demonstration by Julesz (1964) that observers are able to perceive depth relations in stimuli with perfectly camouflaged monocular image information revealed that a sophisticated binocular matching operation must occur in a relatively early stage of visual processing. This stereo matching process quickly became the focus of many computational models of vision, and a number of matching constraints were proposed in order to reduce the complexity of the problem (Marr & Poggio, 1976). The epipolar constraint simplified the search for matching points considerably by taking advantage of the fact that, for an image point in the left eye, all the possible matching points in the right eye must lie on a line representing the projection of the left eye's line of sight to the object onto the right retina. However, in order to know where on the right retina this line falls, the exact relative horizontal, vertical and torsional positions of the eyes must be known and there is considerable uncertainty as to how precisely this information is registered by the visual system. Rogers & Bradshaw (1996) examined whether changes in eye position might cause a change in binocular retinal correspondence, so that the epipolar constraint might be employed over a large range of convergence angles. If binocular correspondence is stable, vertical disparity increases with convergence owing to the perspective distortion of the retinal image. While it may be possible for correspondence to change with convergence and thus to accommodate the perspective distortion, the results of Rogers & Bradshaw (1996) indicate that this compensation is incomplete so that there is still an increasing need to match vertically disparate features as viewing distance is reduced. [For a review of the geometry of epipolar lines the reader is referred to Howard & Rogers (1995) Section 7.1.2 and Fig. 7.6. For a review of the use of epipolar constraints in models of stereopsis the reader is referred to Poggio & Poggio (1984).]

Use of an epipolar constraint in binocular matching implies that match detection and, by extension, stereoscopic depth judgments, should be impaired by small amounts of vertical disparity. Previous studies examining the effect of vertical disparity offsets on depth judgments have produced mixed conclusions. If targets are monocularly identifiable objects, such as isolated lines or dots, evidence indicates that depth judgments from horizontal disparity are possible, even with up to 4 deg of added vertical disparity (Ogle, 1955; Mitchell, 1970). However, experiments with random dot stereograms indicate that depth judgments (Nielsen & Poggio, 1984) and correlation detection (Prazdny, 1985) are impossible if even 10 arcmin(0.17 deg) of vertical disparity is present. The conclusion from both the Nielsen & Poggio (1984) and

---

*University of Houston College of Optometry, Houston, TX 77204-6052, U.S.A.

†University of California–Berkeley School of Optometry, Berkeley, CA 94720-2020, U.S.A.

‡To whom all correspondence should be addressed [*Fax* +1-713 743 2053; *Email* sbstevenson@uh.edu].

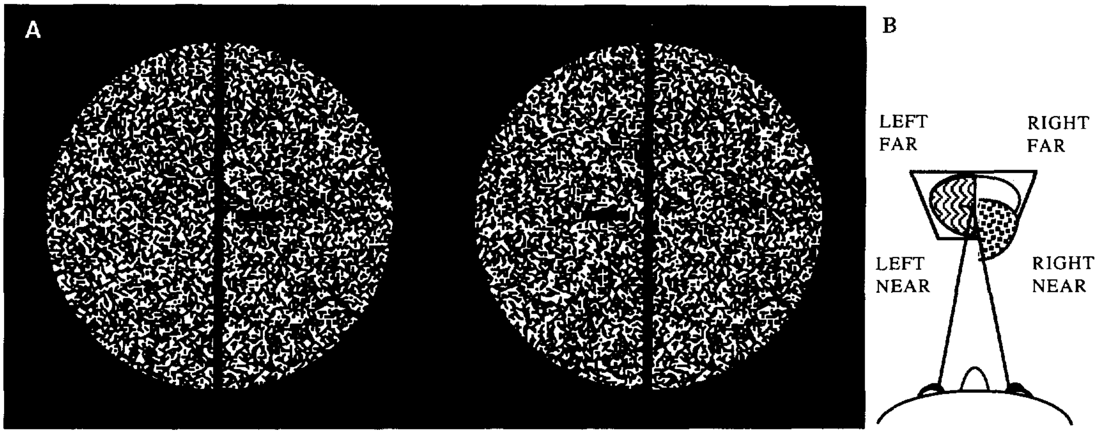

FIGURE 1. Stimulus configuration used in these experiments. (A) Random dot stereogram pair which, when cross-fused, portray a correlated surface on the right side and an uncorrelated surface on the left side. The correlated surface in this example has crossed (near) disparity. In the experiments reported, both horizontal and vertical disparity of dynamic random dots were varied relative to the aperture and lines. (B) The observers were asked in a four-alternative forced-choice paradigm to indicate, by pressing one of four buttons, both which side of the display was correlated (matching task) and whether the surface was near or far relative to fixation (depth task).

Prazdny (1985) studies was that the human visual system does, in fact, take advantage of the epipolar constraint. However, these studies used a relatively small patch of static random dots which probably limited the overall matching sensitivity. Recent experiments with dynamic random dot stereograms have revealed that matching sensitivity improves considerably if the number of matching elements is relatively large, since the visual system is capable of significant averaging over space and time to improve the signal to noise ratio (Cormack *et al.*, 1994).

We have, therefore, re-examined the tolerance of human stereo matching to vertical disparity, using dynamic random element stereograms and a relatively large display. Observers viewed a display in which half of the field contained uncorrelated random dots, with matches distributed equally over all disparities, and the other half contained fully correlated dots in which all dots had a match at the same disparity. The horizontal and vertical disparity of this correlated half was manipulated as well.

We measured both correlation detection and near/far depth discrimination and here present evidence that both abilities are robust to significant amounts of vertical disparity, indicating that the visual system is not strictly constrained to epipolar lines in binocular matching. We discuss a variety of schemes by which this might be realized in the visual system.

METHODS

Details of the display configuration are shown in Fig. 1. Observers viewed a circular field of dynamic random dots in which either the left or right half was fully correlated and was offset with some combination of horizontal and vertical disparity, while the other half was uncorrelated. Observers were asked to make two judgments on each trial: which half of the display contained correlated dots (matching task), and whether these dots were near or far relative to fixation (depth task). The matching task reveals the ability of observers to detect correlation over the range of disparities used, while the depth task reveals the ability of observers to identify the direction of the horizontal disparity component. Observers were given feedback for both judgments through a pair of tones which sounded after each response. Each trial was initiated by the response to the previous trial. Except during the stimulus presentation, the display showed continuous, uncorrelated dynamic noise so that the zero disparity aperture and bisector were the only stimuli to vergence. All changes in disparity occurred while the noise was binocularly uncorrelated to eliminate the possibility of any motion cues.

Because changes in ocular vergence would have altered the retinal disparities from their intended values, several measures were taken to insure stable and accurate binocular fixation. The combination of a stationary circular aperture and vertical bisecting line were designed to stabilize all three axes of vergence: horizontal, vertical and cyclovergence. The target focus was set to each observer's far point and target vergence was adjusted to each observer's phoria, in order to minimize fixation disparity. The sign of vertical and horizontal disparity presented on a given trial was chosen from a shuffled order, so that observers could not benefit from anticipatory eye movements. Nonius lines were present to allow the observer to monitor vertical vergence, and observers were instructed to verify alignment before initiating a trial. Only on rare occasions did observers notice offsets in the nonius lines indicating a vertical vergence response had occurred to the previous trial. These offsets were not evident until after the correlation had returned to zero, and it is unlikely that the vertical vergence substantially altered the disparities presented. On each trial, the correlated patch was presented for just

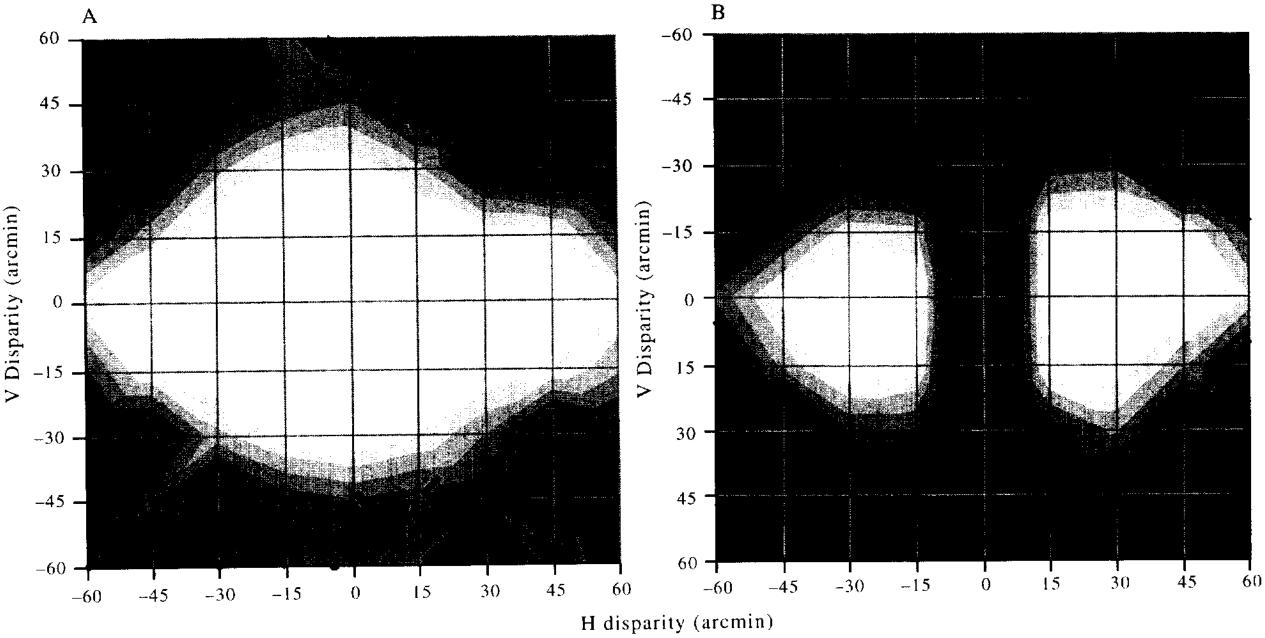

FIGURE 2. Percent correct in the correlation detection (A) and depth identification (B) tasks for observer SBS, plotted against horizontal disparity on the horizontal axis and vertical disparity on the vertical axis. Positive disparity values indicate horizontal crossed disparity and vertical left hyper-disparity. Intersections of grid lines show locations of measured values used to generate the contour plot. Full white areas indicate combinations of horizontal and vertical disparity for which performance was 100% correct. Black areas indicate chance (50% correct) performance. The dashed contour line indicates 75% correct. This observer was able to correctly identify the depth of the random dot surface, even with vertical disparities of up to 30 arcmin. The range of disparities which support stereo judgments is somewhat smaller than that which supports correlation detection, but has the same general shape. Performance in the stereo task falls to chance in the center as expected, where the horizontal disparity is zero.

200 msec, short enough to prevent vergence responses from altering the retinal disparity.[*]

Stimuli were viewed haploscopically through the stimulator optics of a binocular dual-Purkinje image eye-tracker in order to take advantage of the galvanometer-mounted mirrors for precise control of horizontal and vertical disparity. Eye movements were not recorded. The circular apertures, bisector and nonius lines were printed on 35 mm slides which were placed just behind the first lens in the optical path, so that mirror movement displaced the dot patterns behind stationary apertures. Because the dots were dynamic and uncorrelated across frames, lateral shifts of the image could not be detected with either eye alone, insuring that detection of disparity depended on binocular combination and correlation extraction. The apertures were 12 deg in diameter. Pixels were 2.5 arcmin tall by 5 arcmin wide at the 43 cm viewing distance.

Dynamic random dots were generated by special purpose hardware which produced a separate pseudorandom bit sequence for each of two display monitors. Correlation of the dots in each display was determined by a video image under computer control so that any part of the display could be made correlated (identical dots in each image) or uncorrelated (independent, random set of dots in each image), depending on the brightness level in this video image. [This is similar to the method used by Fox *et al.* (1978) for controlling disparity in dynamic random dot displays, but in our case the video signal controlled interocular correlation.] The position of the correlated dots (left/right, near/far) was randomized from trial to trial.

A total of 81 combinations of vertical and horizontal disparity were presented, ranging from +60 to −60 arcmin in 15 arcmin steps for both disparity directions. A block of trials consisted of all 81 combinations, presented in a shuffled order. Each observer ran a total of 60 blocks. The proportion correct for each of the 9 by 9 combinations of horizontal and vertical disparity was fit with a surface using commercial 3D plotting software (DeltaGraph Pro) to generate iso-performance contour lines.

RESULTS

Results for both matching and depth tasks are presented in Figs 2 and 3. These gray scale plots indicate

---

[*] Eye position records from our laboratory taken in the context of another study indicate that neither horizontal nor vertical vergence responses are measurable until more than 200 msec after stimulus onset. These vergence measurements were made with a similar stimulus configuration, except that the entire field of random dots was correlated, instead of just half. Cyclovergence measurements with this stimulus have not been made, but the stimulus used in the current study present vertical disparity in just one-half the field and therefore could stimulate cyclovergence as well because it includes a component of vertical shear disparity (Kaneko & Howard, 1994; 1997). However, reports of cyclovergence measured with comparable stimuli indicate that the latency is longer than 200 msec (Kaneko & Howard, 1994) and the gain is in the order of 0.1 (Howard *et al.*, 1994), making it unlikely that cyclovergence would have appreciably altered the disparities presented in our display.

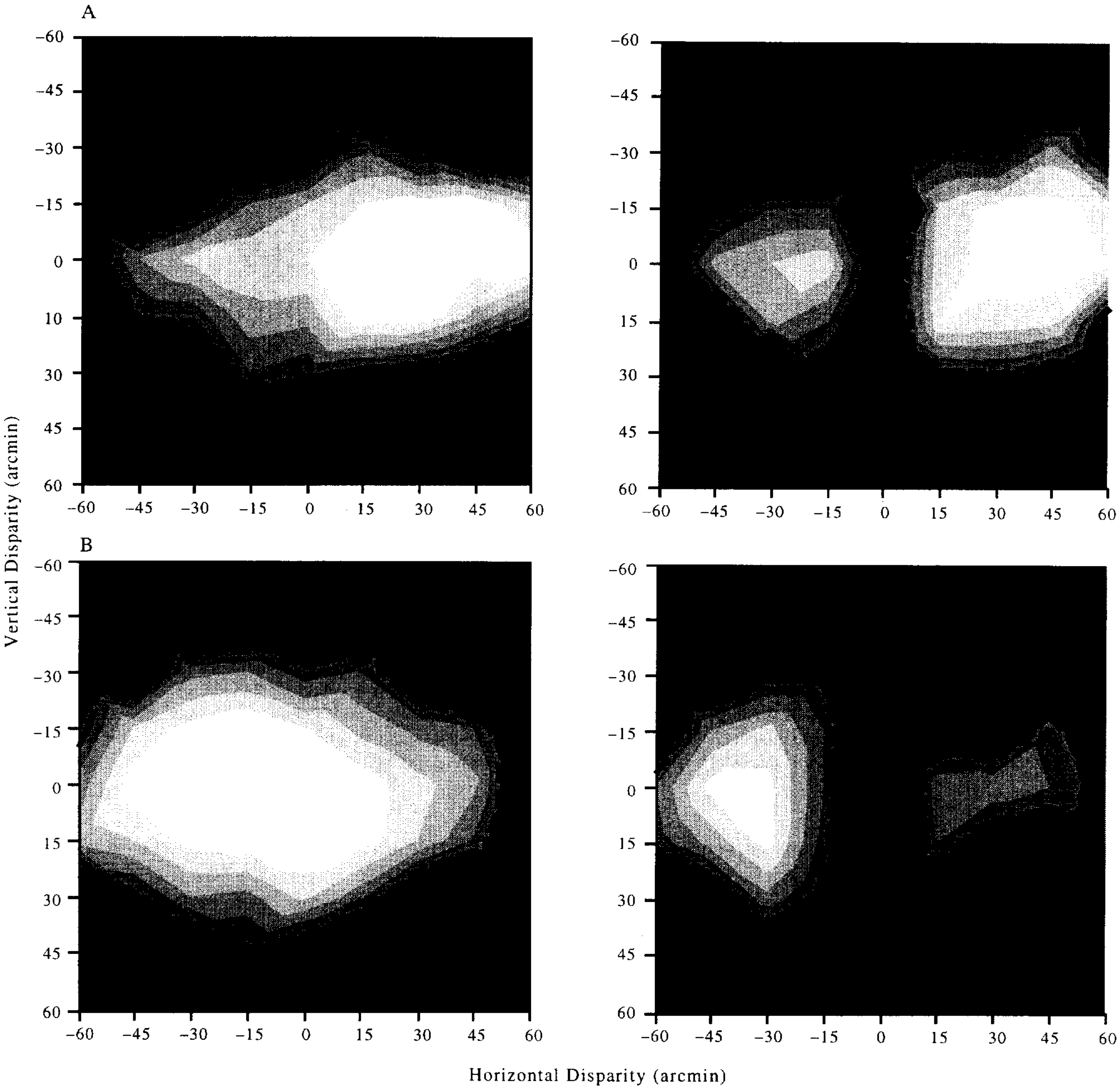

FIGURE 3. Data for observers CMS (A) and MC (B). Note that CMS shows a bias for crossed disparity in both tasks, while MC shows a bias for uncrossed disparity.

the level of performance achieved by observers in each task for combined levels of horizontal and vertical disparity. Brighter areas reflect better performance in the task and the dashed contour lines indicate those combinations of horizontal and vertical disparity which yield 75% correct performance. The intersections of vertical and horizontal grid lines indicate the locations of the measured values which were used to generate the surface.

The matching range is clearly elliptical with respect to disparity, with accurate identification of the correlated region being robust to roughly 1 deg vertically and >1 deg horizontally for observer SBS (Fig. 2). Observers CMS (Fig. 3A) and MC (Fig. 3B) show somewhat smaller ranges, and in addition these observers show a bias for near (CMS) and far (MC) disparities, respectively. A fourth observer (SM, not shown) showed range comparable to SBS, but with a slight bias for far disparity. Despite the individual biases, there is no systematic bias across subjects for either near or far disparity. Nowhere in these results is there evidence for a bias for one or another direction of vertical disparity.

Depth identification is supported over a range of disparities similar to correlation detection, with somewhat reduced limits on both disparity axes. Tolerance to vertical disparity is greatest for the smallest horizontal disparities used for all observers. Note that the depth task necessarily produces chance performance at zero hor-

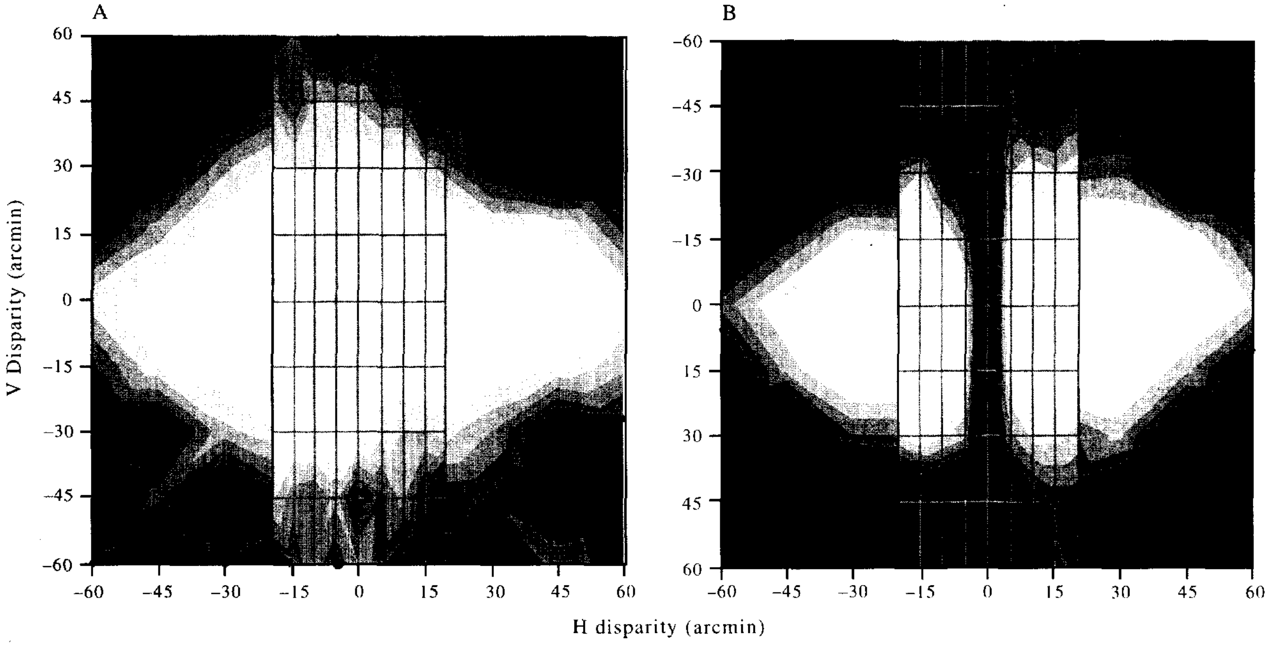

FIGURE 4. Results for observer SBS with a reduced range of horizontal disparity, plotted for comparison as an overlay with the data from Fig. 2. Measured values used to generate the contour plots are once again indicated by the intersections of grid lines. At horizontal disparities of only 5 arcmin the near/far task can be performed at or above 75% correct, with vertical disparities as large as ±30 arcmin.

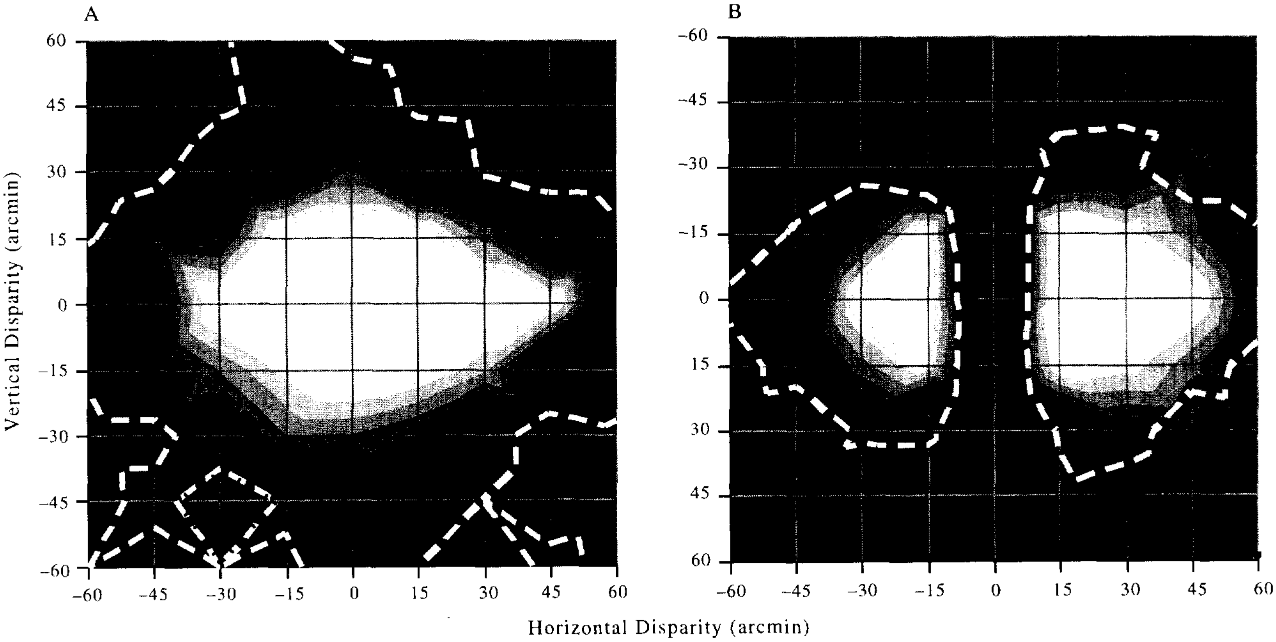

FIGURE 5. Results for observer SBS with 6 deg patch size, with 75% contour shown by black dashed line. For comparison, the 75% contour for the 12 deg patch size is shown as well by the white dashed line. The effect of reducing the patch size is to reduce both the horizontal and vertical range of matching sensitivity.

izontal disparity, producing the dark vertical band down the center of these figures.

Reduced horizontal disparity

In the results shown in Figs 2 and 3, the smallest horizontal disparity used was 15 arcmin. In order to examine the effect of vertical disparity on depth judgments with finer horizontal disparities, we repeated the experiment with one observer using the same range of vertical disparities but a more restricted range of horizontal disparities ranging from +20 to −20 arcmin in 5 arcmin steps.

Results are shown in Fig. 4 for observer SBS. With as little as 5 arcmin horizontal disparity, near/far judgments were 75% accurate with more than 30 arcmin of vertical disparity.

Reduced field size

In attempting to reconcile these results with those of Nielsen & Poggio (1984) and with those of Prazdny (1985), who found little or no tolerance for vertical disparity in similar tasks, we examined the effect of field size on performance. These studies had used much smaller patches of random dots than in the present study and previous work has indicated that sensitivity to correlation depends critically on the overall number of elements presented (Cormack *et al.*, 1994). One observer repeated the experiment with a 6 deg diameter field instead of one of 12 deg. The aperture was the same as was used in the initial measurements, but the correlated region was a 6 deg semicircular patch on either the right or left side of the central line.

Results of these measurements are shown in Fig. 5, with the 75% contour from Fig. 2 replotted for comparison. Under this condition, the overall range of disparities which supported correlation detection and depth identification was reduced similarly for horizontal and vertical disparities. Although tolerance to vertical disparity was reduced, this effect appeared to be secondary to an overall drop in performance for all combinations of horizontal and vertical disparity.

DISCUSSION

The present results indicate that the visual system analyzes binocular correspondence over a considerable range of both horizontal and vertical disparity, as revealed by the significant tolerance to vertical offset in both the matching task and the depth identification task. Certainly, an artificial binocular visual system with known camera parameters might be designed which took advantage of the epipolar constraint, matching points only along a single dimension and thereby reducing the computational load. The human visual system does not strictly follow this constraint, however, perhaps because precise eye alignment information is not available, or because eye alignment is too variable to compensate for under some conditions, such as during locomotion.

The generality of these results and the mechanisms suggested by them should be considered in the context of the stimulus and task employed. The disparity identification task required only a qualitative judgment of depth direction (Ogle, 1952) and not an estimate of the magnitude of depth perceived. However, the use of the dynamic random element stimulus probably puts the task in the category of "fine stereopsis", since the global disparity processing required to solve the correspondence problem is usually associated with fine stereopsis (Julesz, 1978; Tyler, 1991). Mitchell (1970) found substantial tolerance to vertical disparity using isolated line targets and a qualitative depth identification task, which would certainly be considered to stimulate "coarse stereopsis". Thus, tasks which fit both categories of stereoscopic judgments are shown to have a significant tolerance to vertical disparity.

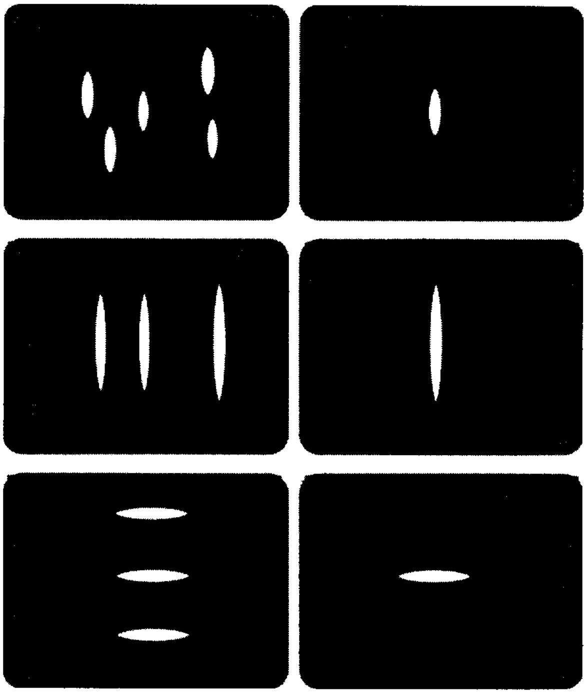

FIGURE 6. Cartoon illustrating hypothetical configurations of receptive fields which are tuned for horizontal disparity, but which have a tolerance for vertical disparity as well. Each panel shows a population of cortical cells whose right eye receptive fields coincide, but whose left eye receptive field positions and phases are scattered so as to produce a range of disparity tuning in the population. In (A), stereosensitivity is tolerant to vertical disparity because there is a scatter in the vertical disparity tuning of horizontal disparity-sensitive cells. In (B), tolerance occurs because receptive fields are vertically elongated, giving them broad tuning for vertical disparity but narrow tuning for horizontal disparity. Cyclo-disparity tuning of these cells would depend on their location in the visual field. While schemes in (A) and (B) may account for horizontal vergence and stereosensitivity, a configuration such as that shown in (C) may also be required in order to account for the precision of vertical vergence control. This receptive field scheme has sharp tuning and position scatter for vertical disparity, providing the sensitivity and disparity direction information required to drive vertical disparity vergence.

What sort of visual mechanisms might account for this observed tolerance to vertical disparity? Fig. 6(A, B) presents some hypothetical binocular receptive field arrangements which would provide sensitivity to horizontal disparity despite the presence of vertical disparity in the stimulus. Tolerance to vertical disparity might occur because of a scatter of receptive fields [Fig. 6(A)] or because of vertically elongated receptive fields [Fig. 6(B)]. The scheme depicted in Fig. 6(B) preserves the notion of an epipolar constraint in the sense that the receptive field scatter is one-dimensional, but matching sensitivity in this scheme is not truly constrained to epipolar lines because the mechanisms are sensitive to matches with a significant range of vertical disparity as well. Figure 6(C) presents an arrangement which provides sensitivity to vertical disparity for the purpose of controlling vertical vergence. It may be that the ensemble of cortical receptive fields includes some or all the types depicted here. For example, a combination of the types in Fig. 6(B, C) acting in parallel might provide enough information for controlling horizontal and vertical vergence and for making stereo judgments. Somewhat similar schemes have been suggested by Morgan & Castet (1995) to explain stereoacuity with horizontally disparate oblique gratings, which stimulate both horizontal and vertical disparity processors (the "aperture problem").

Speculations aside, the conclusion from the present results is that whatever the true arrangement of cortical binocular receptive fields is, it must provide for horizontal disparity sensitivity in the presence of substantial quantities of vertical disparity.

REFERENCES

Cormack, L. K., Stevenson, S. B. & Schor, C. M. (1994). An upper limit to the binocular combination of stimuli. *Vision Research, 34*, 2599–2608.

Fox, R., Lehmkuhle, S. & Leguire, L. E. (1978). Stereoscopic contours induce optokinetic nystagmus. *Vision Research, 18*, 1189–1192.

Howard, I. P. & Rogers, B. J. (1995). *Binocular vision and stereopsis*. New York: Oxford University Press.

Howard, I. P., Sun, L. & Shen, X. (1994). Cycloversion and cyclovergence: the effects of the area and position of the visual display. *Experimental Brain Research, 100*, 509–514.

Julesz, B. (1964). Binocular depth perception without familiarity cues. *Science, 145*, 356–362.

Julesz, B. (1978). Global stereopsis: cooperative phenomena in stereoscopic depth perception. In Held, R., Leibowitz H. W. & Teuber, H. L. (Eds), *Handbook of sensory physiology* (Vol. 8). New York: Springer.

Kaneko, H. & Howard, I. P. (1994). Relative shear disparities and the perception of surface inclination. *Vision Research, 34*, 2505–2517.

Kaneko, H. & Howard, I. P. (1997). Spatial properties of shear disparity. *Vision Research, 37*, 315–323.

Marr, D. & Poggio, T. (1976). Cooperative computation of stereo disparity. *Science, 194*, 283–287.

Mitchell, D. E. (1970). Properties of stimuli eliciting vergence eye movements and stereopsis. *Vision Research, 10*, 145.

Morgan, M. J. & Castet, E. (1995). Stereoacuity for oblique gratings predicted by phase shifts, not disparities. *Investigative Ophthalmology and Visual Sciences Supplement, 36*, 231.

Nielsen, K. R. & Poggio, T. (1984). Vertical image registration in stereopsis. *Vision Research, 24*, 1133–1140.

Ogle, K. N. (1952). On the limits of stereoscopic vision. *Journal of Experimental Psychology, 44*, 253–259.

Ogle, K. N. (1955). Stereopsis and vertical disparity. *Archives of Ophthalmology, 53*, 495.

Poggio, G. F. & Poggio, T. (1984). The analysis of stereopsis. *Annual Review of Neuroscience, 7*, 379–412.

Prazdny, K. (1985). Vertical disparity tolerance in random-dot stereograms. *Bulletin of the Psychonomic Society, 23*, 413–414.

Rogers, B. J. & Bradshaw, M. F. (1996). Does the visual system use the epipolar constraint for matching binocular images? *Investigative Ophthalmology and Visual Sciences Supplement, 37*, 3125.

Tyler, C. W. (1991). Cyclopean vision. In Regan, D. (Ed.), *Vision and visual dysfunction: binocular vision* (pp. 38–74). London: Macmillan.

Acknowledgements—This work was supported by NEI grants EY08882 and EY10531.