



# Disparity-evoked Vergence is Driven by Interocular Correlation

HANSPETER A. MALLOT,\*§ ANKE ROLL,† PETRA A. ARNDT‡

Received 5 July 1995; in revised form 2 November 1995

**Disparity-evoked vergence is studied in stereograms showing one or two depth planes which are defined by isolated dots of varying density and contrast. Vergence position immediately after stimulus presentation was measured using dichoptic nonius lines. Since the stimulus was not visible after the onset of the vergence movement, the experiment accesses the initiation of vergence rather than its eventual result. In the unequivocal stimuli (one depth plane), elicited vergence tends to reduce disparity. Disparities of 0.5–1 deg are most effective which is in accordance with earlier findings. If two depth planes are presented, elicited vergence lies between the two planes, approaching the plane with higher dot density and/or dot contrast. In quantitative measurements, we show that the depth-averaging mechanism uses signal power per depth plane as a weight. Therefore, the relative pulling strength of dot density compared with dot contrast follows a power law with exponent 2. We propose a population code for vergence control based on disparity-tuned pools of units. Copyright © 1996 Elsevier Science Ltd.**

Vergence Stereo Population coding Binocular correlation

## INTRODUCTION

Binocular depth perception has its highest resolution in the well-known Panum area, i.e. a rather small volume centred around the plane of fixation. In order to make optimal use of stereo vision in a given environment, it is crucial, therefore, to adjust the plane of fixation with respect to an “area of interest” in three-dimensional space. This is done by movements of the eyes, of which disjunctive, or vergence movements convey the depth adjustment of the plane of fixation (for review, see Collewijn & Erkelens, 1990; and Judge, 1991). Vergence movements occur in response to stereoscopic or accommodational stimuli (e.g. Semmlow & Wetzel, 1979). While the dynamics of disparity-evoked vergence in reaction to stimuli with unequivocal disparities is well studied (e.g. Rashbass & Westheimer, 1961; Schor, 1979; Erkelens, 1987; Pobuda & Erkelens, 1993; Semmlow *et al.*, 1994), little is known about what stimulus parameter

drives vergence response in stimuli comprising multiple disparity cues. In this paper, we study the information processing underlying disparity-evoked vergence in the presence of multiple disparity cues by means of computer-generated stimuli lacking accommodation cues.

### *Disparity-evoked vergence and stereo matching*

Since vergence adjustment is a precondition for proper stereo vision, it is not surprising that the mechanisms for disparity-evoked vergence are quite different from stereo mechanisms operating within Panum’s area. Some of these differences are as follows.

*Disparity information is obtained while the image is not properly fused.* Since it is normally assumed that image fusion includes solving the stereo correspondence problem, it appears that vergence control uses a simpler stereo mechanism. As a candidate, one could think of image difference mechanisms such as the “automap” scheme proposed by Julesz (see Julesz, 1971) for global stereopsis or the “vergence-displacement energy” proposed by Sperling (1970). Both of these schemes minimize the absolute value of image difference within some window subject to relative shifts of the half-images. Evidence for a squared difference mechanism in human stereoscopic depth perception has been presented, e.g. by Cormack *et al.* (1991), Weinshall (1991), Stevenson *et al.* (1992) and Arndt *et al.* (1995).

*Features inducing disparity-evoked vergence need not be fusible at all.* Westheimer and Mitchell (1969)

\*Max-Planck-Institut für biologische Kybernetik, Spemannstraße 38, 72076, Tübingen, Germany.

†Institut für Mathematik und Datenverarbeitung, Universität der Bundeswehr München, Werner-Heisenberg-Weg 39, 85577, Neubiberg, Germany.

‡Institut für Kognitionsforschung, Carl von Ossietzky Universität, 26111 Oldenburg, Germany.

§To whom all correspondence should be addressed [Tel: +49 7071 601603; Fax: 49 7071 601616; Email: ham@mpik-tueb.mpg.de].

||Parts of this work were presented at ARVO (Mallot & Arndt, 1992) and ECVF (Arndt & Mallot, 1992).

showed, for example, that a dark square on a grey background presented to one eye and a bright square on a grey background presented to the other eye will elicit vergence movement in the direction that brings the two squares into register. Proper fusion of these stimuli, however, is not possible. Other examples use a cross and a circle, etc. Again, it follows that the mechanism guiding vergence differs from standard stereo vision within Panum's area.

*Only the appropriate vergence angle (one number instead of a disparity map) has to be obtained from the stereogram.* This is a computational, rather than a psychophysical difference: the information processing performed for the guidance of vergence eye movements can be expected to be much simpler (and probably faster) than that underlying standard stereo processing.

*Interaction of vergence and correspondence stereo.* Marr and Poggio (1979) suggested that the stereo correspondence problem is solved in a coarse-to-fine hierarchy in which vergence movements are used to realign the image frames when moving to a finer scale. If this was true, the guidance of vergence movements would be simply a coarse version of standard stereo processing (see also Theimer & Mallot, 1994, for an alternative implementation of this idea). This view is supported by the fact that vergence position determines the outcome of the stereo matching mechanism in the wallpaper illusion (McKee & Mitchinson, 1988) and the double-nail illusion (Mallot & Bideau, 1990). However, recent work by Smallman and MacLeod (1994) and Mallot *et al.* (1996b) indicates that there is no strict coarse-to-fine hierarchy for the disambiguation of stereo correspondence. Furthermore, the involvement of vergence eye movements in a coarse-to-fine matching process has not been shown psychophysically (see Mowforth *et al.*, 1981).

In summary, there is good evidence indicating that the information-processing mechanisms guiding vergence eye movements are different from those matching the images and calculating a pointwise depth map.

#### *Disparity-evoked vergence in complex images*

In natural scenes, the guidance of vergence eye-movements must take into account the disparities of many image features occurring at different depth positions. If we assume that the stereograms or all their possible matches are represented in an array of disparity-tuned neurons (for review, see Poggio, 1995; Blake & Wilson, 1991), the question of how one disparity value is selected from this array arises. In this paper, we study the vergence evoked by stereograms depicting clouds of small dots located in two depth planes with disparities  $d_n$ ,  $d_f$ . As is shown in the Appendix, the probability of false-matches is rather low; it will thus be neglected in this initial discussion. In this situation, one of three different outcomes can be expected:

1. *Feature-selection hypothesis:* the vergence system could select one active cell (one match) at random and direct gaze to this feature pair. In this case, one would expect that (i) elicited vergence always

reduces one of the two disparities presented; and (ii) the frequency with which a given disparity plane is selected increases with the relative number of dots in that plane (and maybe with their contrast). Evidence for voluntary target selection for vergence movements in continuously presented stimuli was presented by Erkelens and Collewijn (1991).

2. *Maximal interocular correlation hypothesis:* the vergence system could select the largest peak of activity in the array of disparity-tuned neurons. For the stimuli used in this study, this mechanism predicts that (i) elicited vergence is always one of the two disparities presented; and (ii) the depth plane with the highest signal energy (depending on both dot number and contrast) is selected in all trials.
3. *Population coding hypothesis:* the vergence system could combine the excitations in the sense of a population code, i.e. determine global disparity as an average of the units' preferred disparity weighted with their current excitation (Richards, 1971; for a theoretical discussion, see Lehky & Sejnowski, 1990; Pouget & Sejnowski, 1994). In this case we expect that (i) vergence will take intermediate values in the empty space between the two depth planes presented; and (ii) the exact value will depend on the relative signal energy in the two depth planes.

As was shown by Arndt *et al.* [1995, equation (15)], the last two hypotheses make identical predictions in the case of continuous (low spatial frequency) stimuli. In the case of clouds of dots, however, they can be distinguished, since highly resolved depth information is available even with large correlation windows. Mathematical accounts of the latter two hypotheses and correlation histograms modelling the distribution of excitation in the array of disparity-tuned neurons will be given in the section entitled "Population coding model" and the Appendix. We will show in this paper that hypothesis (3) is closest to the data.

## METHODS

### *Subjects*

Six volunteers, aged 23–38 yr, participated in the experiments. All subjects passed a random dot stereogram test for binocular vision. If necessary, vision was corrected by eye-glasses during the experimental session.

Since it cannot be expected that all subjects are able to fixate the fixation point ideally, the vergence eye position was measured directly after the fixation point had been shown for 750 msec. All subjects exhibited an excess convergent eye position. The measurement was repeated with fixation points at disparities between  $-36$  and  $+9$  min arc. All subjects showed a constant offset indicating that they were affected by a latent convergent phoria (*strabismus concomitans convergens*) which is known to be the most common form of concomitant strabismus (Hruby, 1979).

The measured phoria is shown as a baseline in the plots

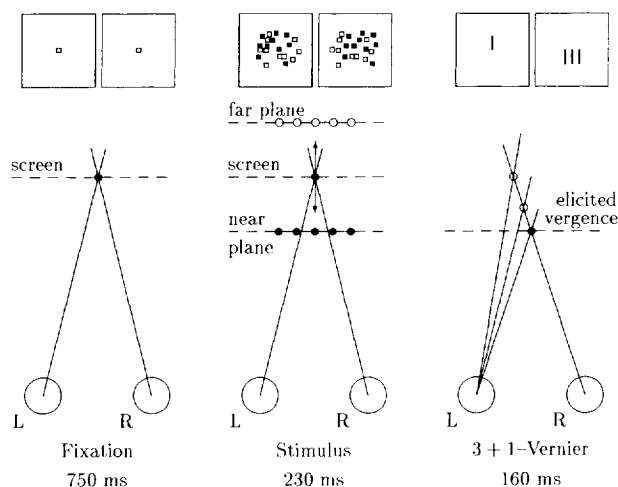


FIGURE 1. Procedure. After a fixation target (vergence angle 3.25 deg arc) is turned off, subjects are presented with a stimulus comprising dots in one or two depth planes (far: disparity = +18 min arc; near: disparity = -18 min arc). To access elicited vergence, a non-fusible "3 + 1" vernier target is presented and subjects are asked to report whether they perceive the upper (single) bar centrally, or shifted to the left or the right with respect to the lower three bars. If a shift is reported, the trial is repeated using a new random stimulus with the same parameters and a 3 + 1 vernier whose upper and lower parts have been offset in such a way as to compensate for the perceived shift. If no shift is perceived, the horizontal separation of the single line and the middle lower line of the vernier is taken as the elicited vergence in this adjustment.

of Figs 2, 3 and 7. The values for each subject (in min arc) are: AE:  $-3.3 \pm 1.5$ , BF:  $-5.5 \pm 1.2$ , HM:  $-0.2 \pm 1.1$ , MK:  $-9.1 \pm 3.6$ , SD:  $-5.9 \pm 2.4$ , WR:  $-2.3 \pm 1.5$ .

### Apparatus

Stimuli were calculated on a Sun-4 workstation. The stereograms were presented on a colour monitor (Mitsubishi Colour Display Model No. HL6905 STGR) in interlace mode, each stereogram with a frequency of 60 Hz. Subjects wore liquid crystal shutter-glasses (Stereographics Inc.) which open and close the left and right eye-glass alternately with a frequency of 60 Hz, in synchrony with the presentation of the half-images on the stereo monitor (see Hodges, 1992 for a review of time-multiplexed or haploscopic stereo computer graphics). The time-averaged transmission of the stereo glasses was 13%, the ratio of transmission in the open and closed phases 50:1. From this ratio, a constant crosstalk of 2% can be expected. In addition, the shutter needs about 1 msec of its 8.3 msec open and closed phases for switching, resulting in another 6% of crosstalk. A total of 8–10% of crosstalk was also found in a psychophysical compensation measurement, where some percentage of each half-image was subtracted from the other half-image prior to display. For the right percentage, the desired image would result after crosstalk. Subjects reported the vanishing of ghosts of high contrast edges at about that

value. For an analysis of the effect of crosstalk on our stimuli, see the Appendix.

The luminance of the screen was calibrated at 52 equally spaced slots out of the 256 slots of the colour map with an UDT photometer (S370 with CRT brightness sensor 265). The resulting calibration curve reflects the built-in  $\gamma$ -correction of this monitor. For a desired intensity value, the interpolated calibration curve was evaluated inversely and the closest colour-map slot was chosen.

Spatial resolution of the screen was  $1152 \times 900$  pixels each subtending an angle of 0.9 min arc in the horizontal direction. Disparities were therefore varied in multiples of 1.8 min arc such that symmetrical offsets of entire pixels would occur in both half-images.

### Procedure

All experiments were performed in a dark room. Subjects were seated comfortably in front of the stimulus monitor. Viewing distance was 115 cm and the head of the subject was fixed with a forehead and chin rest.

Elicited vergence was measured psychophysically using an adjustment task (cf. Fig. 1). For each stimulus condition, 30–40 adjustments were performed, each consisting of a varying number of *single decisions*. A single decision started with the presentation of a binocular fixation target (simulated distance 115 cm resulting in vergence angle of 3.25 deg arc) which was visible for 750 msec. When the fixation point was turned off, the stimulus was presented for 230 msec. This duration was long enough to elicit a vergence movement and short enough to exclude feedback by the beginning vergence movement. For each decision, a new version of the stimulus pattern was generated by randomly varying the positions of the dots within each plane. Immediately after stimulus presentation, a "3 + 1" vernier was shown, i.e. a diplopic image with a single nonius line in the upper half presented to the right eye and a three nonius lines in the lower half presented to the left eye. The spacing of the three lines was 5 deg arc. An offset  $d_i$  could be introduced symmetrically between the upper line and the lower triple. The 3 + 1 vernier was visible for 160 msec. Subjects were instructed to judge whether the single line presented to the right eye was aligned with the central one of the three lines presented to the left eye. If not, they also reported the direction of misalignment. If alignment was perceived, the current vernier offset  $d_i$  was recorded as the elicited vergence and the current adjustment was terminated. If misalignment was perceived, vernier offset was incremented (single line displaced to the right) or decremented (single line displaced to the left), i.e. changed so as to compensate for the perceived offset. A new decision was then carried out with the refined offset and a newly randomized stimulus pattern. The vernier offset was varied in steps of 1.8 min arc which determines the resolution of the method. Only for the large disparities used in Experiment 1b (Fig. 3) were coarser steps of 3.6 or 7.2 min arc used. Since the subjects' task was to report the presence and

TABLE 1. Summary of parameter settings used in the different experiments. In Experiments 3 and 4, one plane was presented with a contrast of 0.51 in all cases

| Expt | Plane No. | Disparity                       | Number of dots | Dot contrast           |
|------|-----------|---------------------------------|----------------|------------------------|
| 1a   | 1         | $-7.2^\circ, \dots, +7.2^\circ$ | 100            | 0.51                   |
| 1b   | 1         | $-18'$                          | 100            | 0, ..., 0.51           |
| 2    | 1         | $-18'$                          | 0, ..., 100    | 0.51                   |
|      | 2         | $+18'$                          | 100, ..., 0    | 0.51                   |
| 3    | 1         | $-18'$                          | 50             | 0, ..., 0.51           |
|      | 2         | $+18'$                          | 50             | 0.51, ..., 0           |
| 4    | 1         | $-18'$                          | Adaptive       | 0.13, 0.26, 0.38, 0.51 |
|      | 2         | $+18'$                          | Adaptive       | 0.13, 0.26, 0.38, 0.51 |

direction of the perceived vernier offset, their attention presumably was focused on the vernier rather than on the briefly flashed stimulus.

Experiments were performed in blocks of 20 adjustments. In each block, four or five different stimulus conditions were employed and each stimulus condition was tested four or five times. Conditions were presented in randomized order. Within each block, only one stimulus parameter (Experiment 1: disparity; Experiment 2: dot distribution; Experiment 3: dot contrast) was varied. In Experiment 4, reciprocal contrast combinations ( $c_1/c_2$  and  $c_2/c_1$ ) were combined in one block. In all cases, the conditions included in one block were selected such that stimuli with predominantly crossed and uncrossed disparities were about equally likely. Thus, expectations biasing the subjects to either convergent or divergent movements were excluded.

The adjustment task used in our experiments could be modified easily into a staircase procedure. However, since vergence started from the fixation spot in each single decision, the relevance of hysteresis effects is low in our procedure and the search for up-down runs in the adjustment steps appears to make little sense.

### Stimuli

Vergence movements were elicited with stereograms containing 100 light dots in front of a uniform dark background. Background luminance was  $29.6 \text{ cd/m}^2$ ; dot luminance was varied between  $38.8 \text{ cd/m}^2$  (13% contrast) and  $91.7 \text{ cd/m}^2$  (51% contrast). All luminance values are given as measured at the display screen. To obtain the average luminances seen through the shutter-glasses, they have to be multiplied by the transmission factor 0.13. Dot size was  $1.8 \times 1.8 \text{ min arc}$  ( $2 \times 2$  pixels). The dots appeared at randomized positions in a field of  $3.8 \times 5.1 \text{ deg}$  visual angle (width  $\times$  height). The random distribution was corrected such that no overlaps between dots occurred. The dots were distributed in one or two depth planes each with a constant disparity and dot contrast. Thus, the parameters describing the stimuli were: the number of depth planes, the disparity of each plane, dot contrast per plane and dot number per plane. In all cases, the dot numbers summed up to a total of 100 in

the entire image. At each single presentation, the random positions (visual directions) of the dots were calculated newly.

In the different experiments, the following parameter settings were chosen (see also Table 1).

*Experiment 1a: Unequivocal stimulus.* Random dot stereograms composed of 100 dots in one depth plane. The contrast of each dot was 51%. Disparity was varied between  $\pm 432 \text{ min arc}$  ( $\pm 7.2 \text{ deg}$ ) for two subjects and between  $\pm 36 \text{ min arc}$  for one additional subject. The disparity steps for the adaptive nonius task were chosen in dependence of disparity size between 1.8 and 7.2 min arc. There were 20 adjustments per parameter setting.

*Experiment 1b: Contrast threshold.* Random dot stereograms composed of 100 dots in one depth plane ( $-18'$  disparity). The contrast was varied between 0, 3.1, 6.3, 9.4, 12.5, 25 and 51%. There were 20 adjustments per parameter setting.

*Experiment 2: Distribution of dots.* Random dot stereograms showed two depth planes with disparities  $d_n = -18'$  and  $d_f = +18'$ , spaced symmetrically before and behind the fixation plane. Here and in the following, the indices  $n$  and  $f$  denote the near and far planes, respectively. The contrast of each dot was  $c_n = c_f = 0.51$ . The total number of dots was  $k_n + k_f = 100$ , the distribution over the two planes ranging from  $k_n/k_f = 0/100$  (all dots in far plane) through  $100/0$  (all dots in near plane) in 11 steps (0/100, 10/90, 20/80, 30/70, 40/60, 50/50, 60/40, 70/30, 80/20, 90/10 and 100/0). For each parameter setting, 40 adjustments were performed.

*Experiment 3: Distribution of contrast.* Each of the two depth planes with disparities  $d_n = -18'$  and  $d_f = +18'$  consisted of  $k_n = k_f = 50$  dots. In all conditions, the dots of one plane were presented with a contrast of  $c = 0.51$  while the dots of the other plane had one of the following nine contrast values: 0.0, 0.06, 0.13, 0.19, 0.26, 0.32, 0.38, 0.45 and 0.51. Thus, 17 contrast combinations were tested. For each parameter setting, 40 adjustments were performed.

*Experiment 4: Compensation.* For a contrast combination  $c_1/c_2$ , we define the *compensatory dot ratio*  $k_1/(100 - k_1)$  as the dot ratio where disparity inversion, i.e. the exchange of the two half-images, does not affect elicited vergence. Two depth planes at disparity  $d_n = -18'$  and  $d_f = +18'$  were presented in one of six contrast combinations (0.51/0.13, 0.51/0.26, 0.51/0.38, 0.38/0.51, 0.26/0.51, 0.13/0.51). For each combination, from three to five dot distributions were tested to determine the compensatory dot ratio.

## RESULTS

### Single layer of dots

*Experiment 1a: Unequivocal stimulus.* A random dot stereogram consisting of 100 dots in one plane was presented at different disparities between  $\pm 432 \text{ min arc}$  (7.2 deg) to examine the dependence of elicited vergence position on stimulus disparity. With this experiment the range of sensible stimulus disparities and the magnitudes

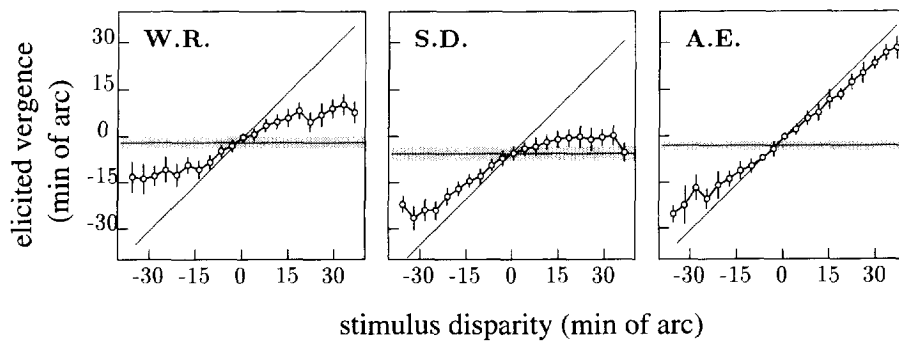


FIGURE 2. Results of Experiment 1a for small disparities. Dots and error bars are mean and SD of 20 adjustments. The horizontal line and shaded area indicate the subject's phoria (with SD) measured independently (see Methods). Vergence is roughly a monotonical function of disparity, but gain varies strongly between subjects.

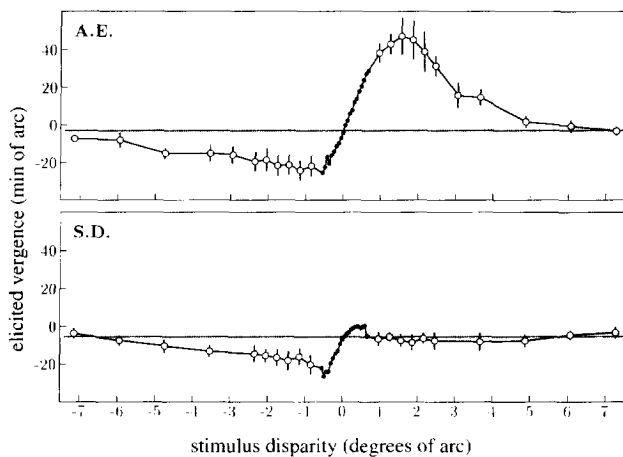


FIGURE 3. Results of Experiment 1a for larger disparities. Dots and error bars are mean and SD of 20 adjustments. For comparison, data plotted in Fig. 2 are replotted as small solid dots. The horizontal line and shaded area indicate the subject's phoria (with SD) measured independently (Methods). For large disparities, elicited vergence goes down to zero (i.e. phoria).

of vergence eye movements elicited by different, unambiguous disparities can be clarified.

Plots of elicited vergence as a function of stimulus disparity are shown in Figs 2 and 3. Within a central range of about  $\pm 40$  min arc, elicited vergence is roughly a monotonic function of the presented stimulus disparity with a gain factor weakly (subject AE) or substantially (subjects WR and SD) below unity. For larger disparities (Fig. 3), elicited vergence decreases to the subject's level of phoria which had been measured independently. Note the large differences between subjects in vergence gain, in the peak position (most effective disparity) and in the asymmetry for crossed and uncrossed disparities.

With the psychophysical method, we can measure only vergence position at some point in time. It is conceivable, therefore, that the differences between subjects or between divergent and convergent movements are due to differences in vergence velocity. In order to test this possibility, Experiment 1a was repeated with an additional time interval of 170 and 570 msec between

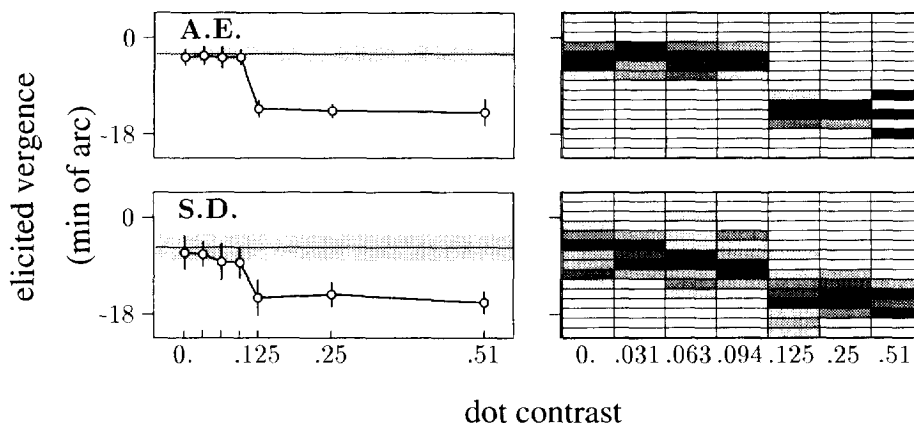


FIGURE 4. Results of Experiment 1b (contrast threshold). Left frames: mean and SD of elicited vergence from 20 adjustments per point. As before, the horizontal line and shaded area indicate the subject's phoria (with SD) measured independently (see Methods). Right frames: raw data shown as histograms. For each bin, the grey levels indicate the number of adjustments in which a given vergence was elicited by some contrast value. For subject AE, a clear step between  $c = 0.094$  and  $c = 0.125$  is found. For subject SD, a slightly blurred step occurs between the same contrast values. Note the different scales on the contrast axis for the plots and the histograms.

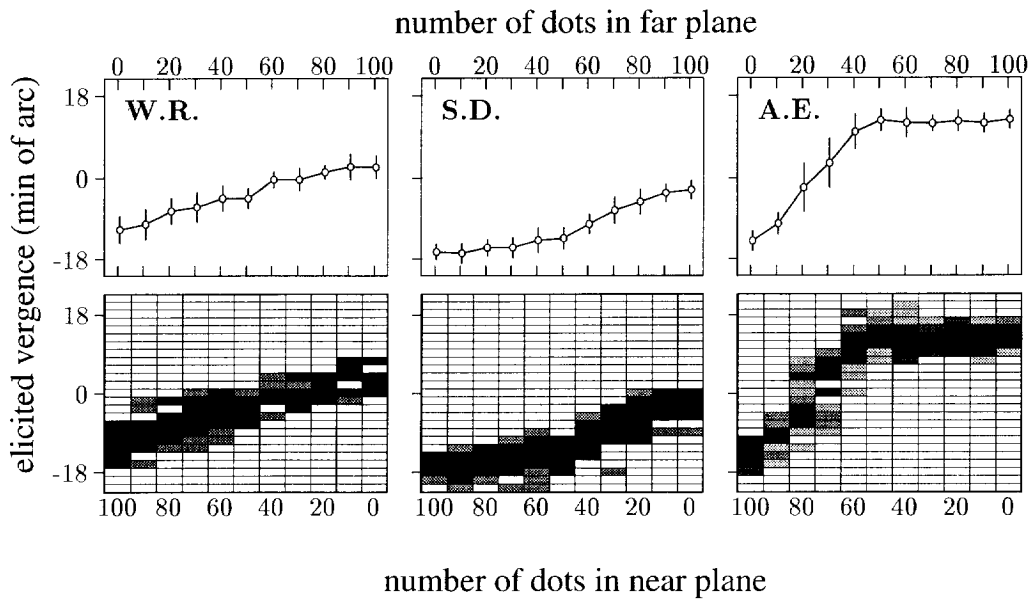


FIGURE 5. Results of Experiment 2 (distribution of dots). Top: mean and SD of elicited vergence from 20 adjustments per dot. Bottom: same data replotted as vergence histograms (bin width: 1.8 min arc). For intermediate distributions of dots between the two planes, intermediate vergence positions are found.

stimulus and nonius presentation. During this interval, the screen was blank. In a further test, stimulus presentation time was extended to 800 msec. In none of these cases was a significant difference found compared with the original condition (subjects PU and SD).

*Experiment 1b: Contrast threshold.* The effect of contrast in the presence of a single depth plane is illustrated in Fig. 4. For contrasts below about 0.1, elicited vergence is zero. Above this contrast level,

elicited vergence switches to a value independent of contrast. In the single plane situation, the vergence response is an all-or-nothing process.

*Competing depth information*

*Experiment 2: Distribution of dots.* Random dot stereograms composed of 100 dots in two depth planes with disparities of  $\pm 18$  min arc with respect to the fixation point were used as test stimuli. This arrangement

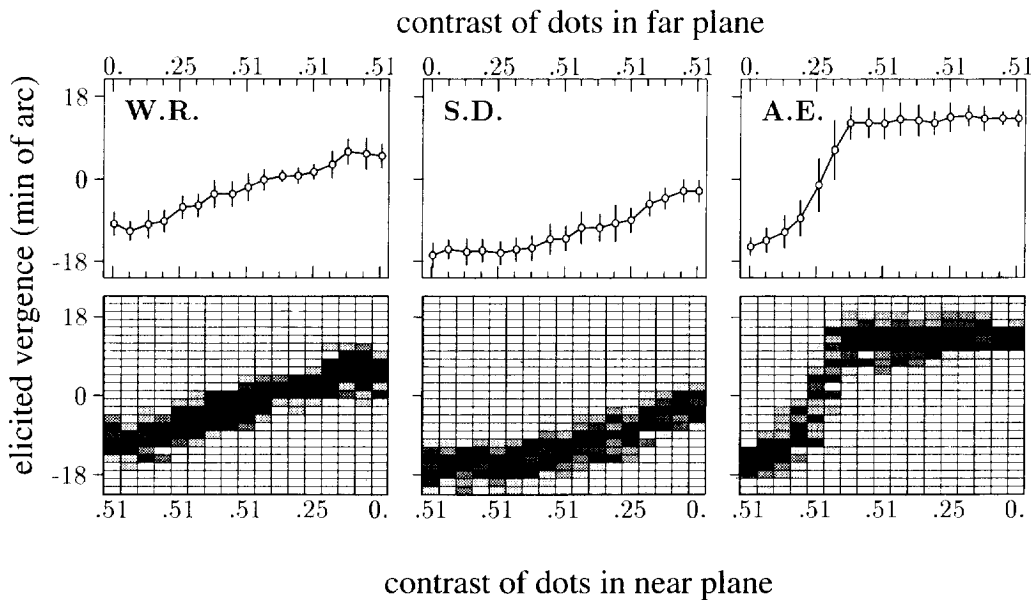


FIGURE 6. Results of Experiment 3 (distribution of contrast). Top: mean and SD of elicited vergence from 20 adjustments per dot. Bottom: same data replotted as vergence histograms (bin width 1.8 min arc). For intermediate distributions of contrast between the two planes, intermediate vergence positions are found.

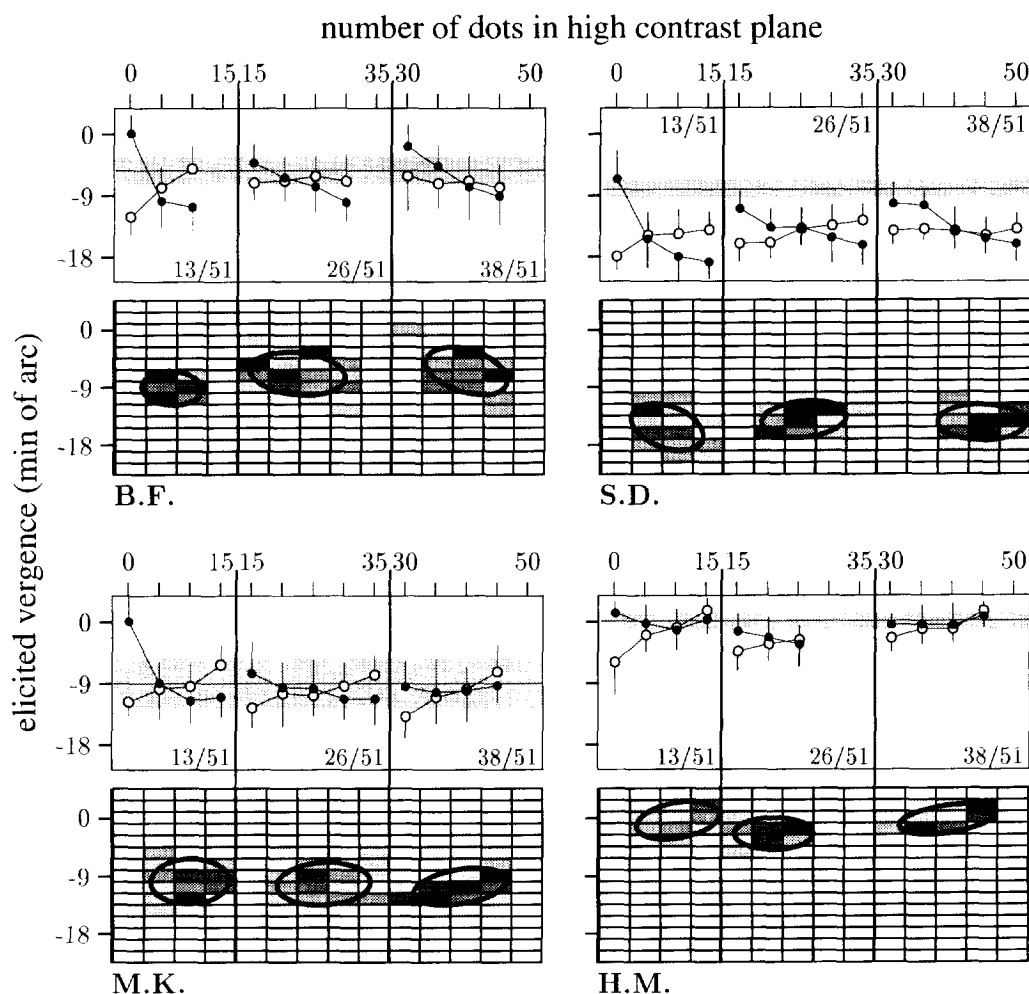


FIGURE 7. Results of Experiment 4 (Contrast vs density trade-off): for each subject, the top panel shows means and SD of elicited vergence from 20 adjustments per point; (●—●) high contrast plane (51%) in front; (○—○) low contrast plane (13, 26, or 38%) in front. In each subpanel, data for one contrast ratio (shown in the lower right-hand corner) are shown; stimulus conditions for open and solid symbols differ by an exchange of the half-images of the stereogram. In the histograms, the grey level is proportional to the product of the relative frequencies of elicited vergence, i.e. dark regions indicate that this vergence angle was obtained frequently with and without inversion of the two half-images. The ellipses mark the point where inversion of the two half-images has no effect on elicited vergence.

results in a conflicting vergence situation. The dot ratio in the two planes (near/far) was varied between 0/100 and 100/0 in 11 steps to investigate the role of relative dot frequency per depth plane.

As can be seen from Fig. 5, elicited vergence changes monotonically with dot distribution. The histograms (lower part of Fig. 5) show that these results are not due to alternating decisions for con- and divergent eye movements, but that vergence movements are directed to depth positions between the two transparent dot planes. Alternating decisions would lead to a bimodal distribution in the vergence histograms with peaks at the two depth positions of the dot planes. In contrast, unimodal distributions with roughly constant variance are found, the mean being shifted in depth with changing dot density in the two planes.

The results also show the inter-subject differences that can be expected from the findings of Experiment 1a. For

the conditions 0/100 (all dots in the far plane) and 100/0 (all dots in near plane), which were already used in Experiment 1a, the results are in line with the findings reported there. For the intermediate conditions, intermediate vergence angles are found. For subject SD, for example, uncrossed disparities were almost ineffective in Experiment 1a. It is not surprising, therefore, that the results of Experiment 2 are confined to the same depth range. The intermediate vergence angles, however, are found in this subject as well.

*Experiment 3: Distribution of contrast.* In order to test whether the intermediate vergence values found for intermediate distributions of dot density in Experiment 2 also can be produced by variation of the distribution of dot contrast, we displayed stimuli composed of two depth planes, each with 50 dots. Contrast of the dots in the two planes was varied between two extreme conditions, where the dots in one plane had contrast zero (i.e. were

TABLE 2. Equivalent dot frequencies for three contrast ratios (results of Experiment 4). As an example, the vergence elicited by a contrast ratio of 0.13–0.51 in subject BF can be cancelled by a dot distribution of 93.4–6.6. Errors correspond to the horizontal width of the error ellipses of Fig. 7

|    | Contrast ratio |            |            |
|----|----------------|------------|------------|
|    | 13/51          | 26/51      | 38/51      |
| BF | 6.6 ± 4.8      | 22.1 ± 7.8 | 39.8 ± 6.6 |
| HM | 10.3 ± 6.7     | 21.0 ± 6.4 | 39.2 ± 7.7 |
| MK | 10.0 ± 6.4     | 26.8 ± 7.6 | 39.1 ± 7.6 |
| SD | 8.2 ± 5.9      | 25.6 ± 6.9 | 44.6 ± 7.2 |

not presented at all). The results are shown in Fig. 6. They resemble strongly the curves obtained in the previous experiment, including the characteristics of the three subjects: intermediate contrast distribution elicits intermediate vergence movements.

*Experiment 4: Contrast vs density trade-off.* In order to quantify the relative strength of dot density and dot contrast in the attraction of the vergence movement, we performed a cancellation measurement. The effects were said to cancel each other, if an exchange of the two half-images of the stereogram did not affect elicited vergence. This definition is independent of the phoria angle of the subjects. The data are plotted in Fig. 7, solid symbols (●) correspond to stimuli with the high contrast dots in front (contrast near condition) while open symbols (○) correspond to stimuli with low contrast dots in front (contrast far condition). At the intersection point of the two curves, cancellation occurs.

A more reliable measure of the equivalence point is obtained from the histograms appearing in the lower half of each panel in Fig. 7. They show the relative frequency of each elicited vergence in the contrast far condition multiplied by its relative frequency in the contrast near condition. Thus, strong labelling (dark colour) in the histograms marks points close to the equivalence point. The ellipses in Fig. 7 are covariance ellipses of the grey level distribution. Their projection onto the horizontal (dot density) direction are taken as a measure of the error of the equivalence point (see Table 2).

## DISCUSSION

### *Unequivocal stimuli*

If only one depth plane is presented, vergence eye movements are performed in the expected direction. The data from Experiment 1a show that elicited vergence reduces the presented disparity best in a range of about ±30 min arc, which is in good agreement with the gain measurements performed by Rashbass and Westheimer (1961). Disparity reduction is not complete, however, which is probably due to the rather short presentation times (Erkelens, 1987).

Experiment 1 is reported in this paper to illustrate two important points: firstly, the psychophysical method used

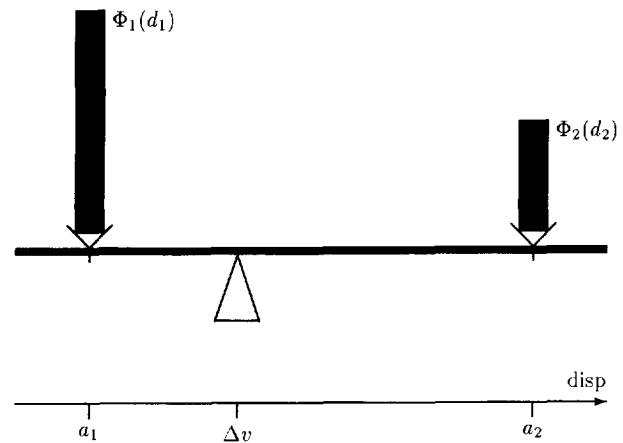


FIGURE 8. Illustration of the population coding scheme. For explanation see text and equation (2).

gives reliable and reproducible results. Of course, vergence velocity cannot be measured by this method. However, for the computational questions addressed in the present paper, measurements of vergence position suffice. Secondly, the curves presented in Fig. 2 can be used as reference for each subject in Experiments 2 and 3. In fact, they reveal considerable differences between the subjects.

The curves for elicited vergence as a function of disparity (Figs 2 and 3) show the same qualitative behaviour as the curves for perceived depth as a function of disparity reported by Richards (1971). This finding indicates that vergence movements are directed to the perceived depth position of the stimulus. Depth perception and vergence control thus might employ the same representation of disparity. Note that the data of Richards (1971) were obtained without eye movements. If vergence movements are allowed, veridical depth estimates are obtained over a much larger range (Foley & Richards, 1972).

### *Depth averaging*

The results of Experiments 2 and 3 (Figs 5 and 6) indicate that disparity-evoked vergence is directed towards average depth, even if matches appearing at this average depth position are rare (*cf.* Appendix). This result becomes clear from the histograms included in the figures: if the reported depth averaging is an artefact of averaging of the vergence data after the experiment, bimodal distributions should be expected in the histograms. Since this is not the case, we conclude that depth averaging occurs in the vergence system itself. We can therefore discard the first two hypotheses on vergence control formulated in the Introduction, i.e. the feature selection hypothesis and the maximal interocular correlation hypothesis, both of which predict bimodal distributions in the vergence histogram. In contrast, the population coding hypothesis is well in line with the data. We will show below that a straightforward

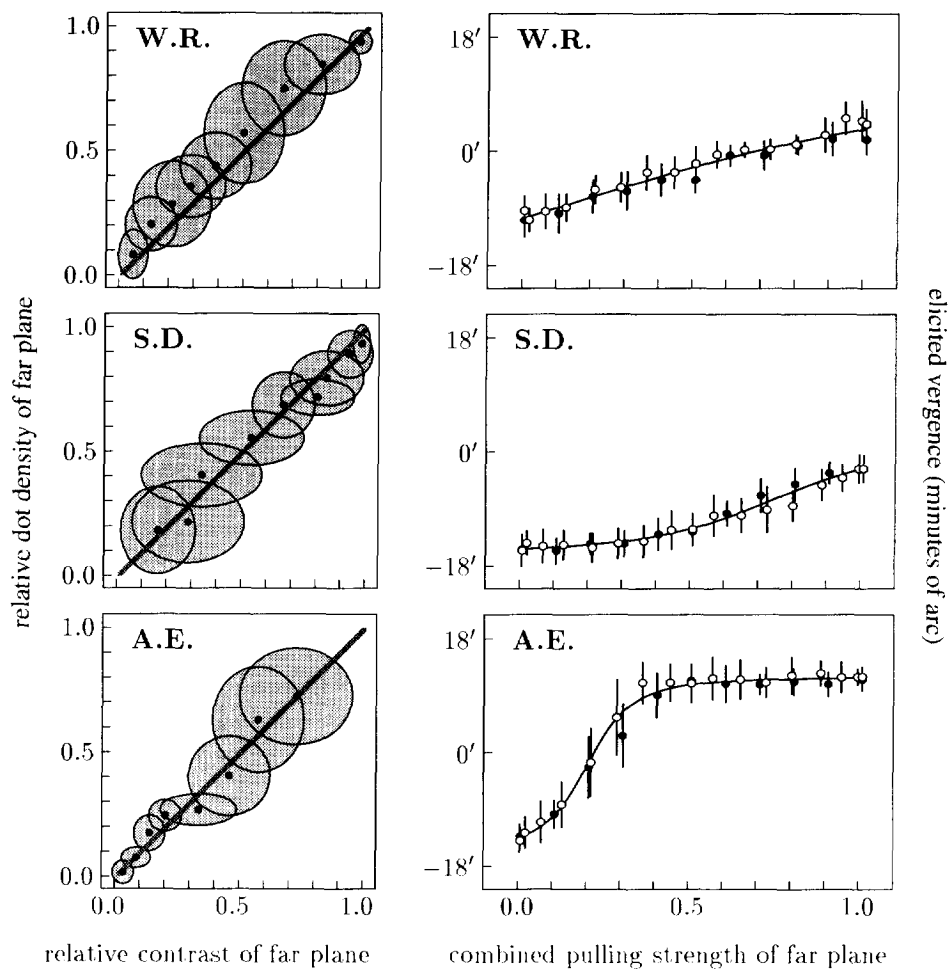


FIGURE 9. Comparison of the relative pulling strength of contrast and dot density from Experiments 2 and 3. Left column: for each vergence angle  $v$ , the distribution of contrast and dots eliciting this angle was determined by inversely reading the histograms of Figs 5 and 6. Each data point corresponds to a pair of contrast and frequency values leading to the same elicited vergence [see equations (6) and (7)]. The thick diagonal line marks the identity on which all data points should lie if equation (2) with  $\Phi(d) = \text{number of dots} \times \text{contrast}^2$  was ideally satisfied. Right column: elicited vergence of both experiments, replotted as a function of the total pulling strength of the far plane [equation (8)]. Solid symbols: variation of dot density (Experiment 2). Open symbols: variation of dot contrast (Experiment 3). The curves are least-squares fits of equation (2) with appropriate values of  $a_n$ ,  $a_l$  and  $\alpha$ .

mathematical formulation of this hypothesis allows for a quantitative description of the results.

The result is reminiscent of "binocular depth mixture" reported by Foley (1976), who studied double-nail stimuli with different contrasts in the crossed (left bar in right half-image, right bar in left half-image) and the uncrossed "nails". If contrast ratio was close to one, subjects would perceive both nails side by side at zero relative depth, fusing the left and right bars of the two half-images, i.e. bars with slightly different contrasts. For contrast ratios significantly different from unity, however, the perceived depth position approached the higher contrast bar. Krol and van de Grind (1986) have argued that the finding of Foley (1976) may be an artefact of vergence eye-movements which had not been controlled by Foley (1976). In fact, our results predict that Foley's stimulus with unequal luminances should elicit a

vergence movement in the direction of the brighter stimulus.

The finding of depth averaging in vergence eye-movements also is reminiscent of direction averaging in (versional) saccades to multiple targets as reported by Findlay (1982) (see also Findlay & Harris, 1993). The population coding scheme presented below includes the centre of gravity mechanism discussed by Findlay (1982).

#### Population coding model

The conditions used in the experiments reported here place different amounts of signal energy in different disparity planes. A simple model of the observed depth averaging in elicited vergence can be derived along the following lines (Fig. 8): Let  $d_i$ ,  $i = 1, \dots, n$  denote the disparities represented in the nervous system, e.g. by disparity-tuned neurons. Suppose now that for each disparity  $d_i$ , a weight  $\Phi(d_i)$  exists, pulling the vergence  $v$

towards this disparity. A possible interpretation of  $\Phi(d_i)$  is the total excitation of neurones tuned to disparity  $d_i$ . In this situation, the simplest assumption is that the resulting vergence movement  $\Delta v$  is the centre of gravity of the distribution of weights  $\Phi(d_i)$ :

$$\Delta v = \frac{\sum_{i=1}^n d_i \Phi(d_i)}{\sum_{i=1}^n \Phi(d_i)}. \quad (1)$$

In order to explicitly fit this model to our data, we need two small modifications: if, for some disparity  $d_o$ ,  $\Phi(d_o) > 0$  and  $\Phi(d_i) = 0$  for all  $d_i \neq d_o$ ,  $\Delta v$  in equation (1) always will approach the sole presented disparity  $d_o$ . Since this is not the case (Experiment 1a), we assume that averaging is not performed over the true disparities  $d_i$ , but over unknown effective disparities  $a_i$ . Also, we include an exponent  $\alpha$  for probability summation to account for the sigmoid curves of subject AE. The modified model then reads:

$$\sigma(\Delta v) = \frac{\sum_{i=1}^n \sigma(a_i) \Phi(d_i)}{\sum_{i=1}^n \Phi(d_i)} \quad (2)$$

where

$$\sigma(x) := \text{signum}(x)|x|^\alpha \quad (3)$$

for some exponent  $\alpha > 0$ .

The effective disparities  $a_i$  are subject-specific constants reflecting the subject's sensitivity to this particular disparity. In the light of Richard's (1971) "pool hypothesis", they can be interpreted as the density of units tuned to the corresponding disparity. If this density is low, i.e. if the disparity value is represented poorly in the population code, stimuli at this disparity will be less effective. In the following, we will develop relative measures for the effects of dot contrast and density in which the subject-specific parameters cancel out. In these relative measures, the inter-subject differences are considerably reduced (*cf.* left column of Figs 9 and 10).

#### Comparing the relative influence of dot density and contrast

In the above equations (1) and (2), we have included a "pulling strength"  $\Phi$  for each disparity plane. It can be considered a measure of the excitation of units tuned to this disparity in a population coding scheme. It is clear from the data that  $\Phi$  depends on both dot density and dot contrast. In this subsection, we study the quantitative relation between the influences of the two stimulus parameters.

Let  $k_n$  and  $k_f$  denote the number of dots in the near (negative disparity) and far (positive disparity) planes, respectively. In our experiments, we always had  $k_n + k_f = 100$ . Let further  $c_n$  and  $c_f$  denote the corresponding contrasts. We assume now that the pulling strength is proportional to  $kc^p$  and estimate the exponent  $p$  from the data. Note that for  $p = 2$ ,  $\Phi$  becomes the ordinary signal power or interocular correlation.

*Comparison of Experiments 2 and 3.* Consider the histograms in Figs 5 and 6. Instead of calculating the average vergence elicited by each stimulus condition,

likewise we can evaluate the histograms inversely, i.e. pick a vergence value  $v_0$  and find all stimulus conditions which elicited this particular vergence. Comparing the result for the histograms from Experiments 2 and 3, we obtain for each elicited vergence a dot density ratio and a contrast ratio which had the same effect on vergence. In terms of the above model, equation (2), we have:

$$\text{Expt 2 : } \frac{\sigma(a_n)k_n + \sigma(a_f)k_f}{k_n + k_f} + \tilde{v}_D = \sigma(\Delta v_0) \quad (4)$$

$$\text{Expt 3 : } \frac{\sigma(a_n)c_n^p + \sigma(a_f)c_f^p}{c_n^p + c_f^p} + \tilde{v}_C = \sigma(\Delta v_0). \quad (5)$$

The terms  $\tilde{v}_D$  and  $\tilde{v}_C$  summarize the effects of false-matching energy at other disparities for the dot experiment and the contrast experiment, respectively. A calculation of expected false-matching energy is given in the Appendix. Note that the contrast used in Experiment 2 and the number of dots per plane used in Experiment 3 were constant. They do not, therefore, show up in the above equations. Solving equation (4) for  $k_f/(k_n + k_f) = k_f/100$  and equation (5) for  $c_f^p/(c_n^p + c_f^p)$ , i.e. for the relative pulling strength of the far plane, we obtain:

$$\text{Expt 2 : } \frac{k_f}{k_n + k_f} = \frac{\sigma(\Delta v_0) - \tilde{v}_D - \sigma(a_f)}{\sigma(a_n) - \sigma(a_f)} \quad (6)$$

$$\text{Expt 3 : } \frac{c_f^p}{c_n^p + c_f^p} = \frac{\sigma(\Delta v_0) - \tilde{v}_C - \sigma(a_f)}{\sigma(a_n) - \sigma(a_f)}. \quad (7)$$

The function  $\sigma$  as well as the constants  $a_n, a_f$  depend only on the subject. The expected false-matching energy for both experiments is approximately the same. Thus, the right-hand sides of the above equations are equal for any given vergence  $\Delta v_0$ . The range of elicited vergence occurring in Experiments 2 and 3 was divided into bins with at least 25 data points from either experiment per bin. For each vergence bin, the mean and SD of  $k_f/100$  was determined from the data of Experiment 2, while mean and SD of  $c_f^2/(c_n^2 + c_f^2)$  (i.e. choosing  $p = 2$ ) were determined from Experiment 3. The results are shown in the left-hand column of Fig. 9. The heavy diagonal marks the expected curve if  $\Phi(k, c) \propto kc^2$ , i.e. if pulling strength is proportional to signal energy (interocular correlation) per depth plane. The grey ellipses show the SDs. The data are well in line with the signal energy assumption.

The right-hand column of Fig. 9 shows the data from both Experiments 2 and 3 replotted as a function of:

$$\frac{\Phi(k_f, c_f)}{\Phi(k_n, c_n) + \Phi(k_f, c_f)} = \frac{k_f c_f^2}{k_n c_n^2 + k_f c_f^2}, \quad (8)$$

i.e. the relative combined pulling strength of the far plane derived from the signal energy assumption for both experiments (solid symbols: Experiment 2, open symbols: Experiment 3). The data from both experiments coincide on the same curve. The continuous lines are least-squares fits to the combined data set with optimized

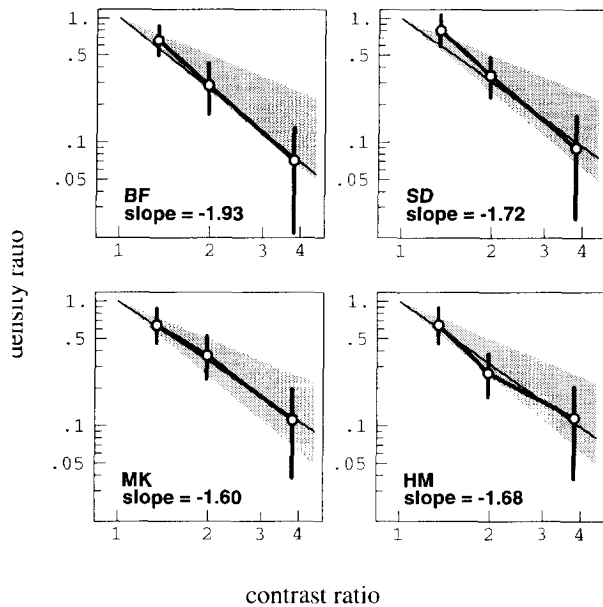


FIGURE 10. Equivalence points of contrast and density ratio as determined in Experiment 4 [cf. equation (10)]. Note that in this display, the curve of equivalence points must pass through the point (1, 1). The thin line is the best fitting straight line passing through (1, 1); its slope (estimate of exponent  $p$ ) is displayed in the lower left-hand corner of each panel. The grey triangles mark slopes between  $-1$  (upper edge) and  $-2$  (lower edge). Average slope from four subjects is  $-1.73$ ; the line for slope  $-2$  passes through all error bars.

exponent  $\alpha$  of the assumed nonlinearity and subject-specific constants  $a_f$ ,  $a_n$  [equation (2)].

*Equivalence points (Experiment 4).* A direct estimation of exponent  $p$  describing the relative contributions of contrast and dot density to the pulling strength can be derived from the data of Experiment 4. Equivalence points were defined as the distribution of signal energy with the property that exchange of the half-images did not affect elicited vergence. In terms of the population coding model [equation (2)]:

$$\sigma(a_n)k_n c_n^p + \sigma(a_f)k_f c_f^p + \tilde{v}_- \\ \sigma(a_n)k_f c_f^p + \sigma(a_f)k_n c_n^p + \tilde{v}_+ \quad (9)$$

In this case, the expected false-matching energies  $\tilde{v}_+$  and  $\tilde{v}_-$  are exactly the same and cancel out. Easy calculation yields:

$$\frac{k_n}{k_f} = \left(\frac{c_n}{c_f}\right)^{-p} \quad (10)$$

(provided that  $\sigma(a_n) \neq \sigma(a_f)$ , i.e. that the two planes pull towards different depth locations). Figure 10 shows the results of Experiment 4 (Table 2) in a double logarithmic plot allowing an estimate of the exponent  $p$ . Each data point corresponds to one subpanel (contrast ratio) in Fig. 7. The point satisfying contrast ratio = density ratio = 1 corresponds to a stimulus, where both depth planes have the same dot density and contrast. In this case, exchange of the half-images does not change the stimulus and

cannot, therefore, have any effect. Thus the curves of Fig. 10 must pass through the point (1, 1). The average exponent  $p$  estimated from four subjects is  $-1.73$ . The curve for  $p = -2$  passes through all error bars and is therefore consistent with the data, while the curve for  $p = -1$  is not. This result indicates that the pulling strength for elicited vergence is in fact proportional to interocular correlation  $kc^2$ .

## CONCLUSION

1. In stimuli with multiple disparities, elicited vergence is directed initially to the centre of gravity of the cues at different disparities. The centre of gravity calculation takes into account both dot density and dot contrast.
2. The weight of each disparity plane for the centre of gravity calculation is proportional to (number of dots)  $\times$  (dot contrast)<sup>2</sup>, i.e. it is proportional to the signal energy or interocular correlation of the depth plane.
3. The results are consistent with the notion of a population code for disparity and vergence control. The mechanism is thus close to the one proposed for intensity-based stereo by Arndt *et al.* (1995) and Mallot *et al.* (1996a) as well as to the mechanism proposed for saccadic target selection by Findlay (1982).

## REFERENCES

- Arndt, P. A. & Mallot, H. A. (1992). Disparity-evoked vergence is directed towards average depth. *Perception*, 21, Suppl. 86 (Abstr. 15th ECVF, Pisa, Italy).
- Arndt, P. A., Mallot, H. A. & Bülthoff, H. H. (1995). Human stereovision without localized image-features. *Biological Cybernetics*, 72, 279–293.
- Blake, R. & Wilson, H. R. (1991). Neural models of stereoscopic vision. *Trends in Neurosciences*, 14, 445–452.
- Collewijn, H. & Erkelens, C. J. (1990). Binocular eye movements and the perception of depth. In Kowler, E. (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 213–261). Oxford: Elsevier Science.
- Cormack, L. K., Stevenson, S. B. & Schor, C. M. (1991). Interocular correlation, luminance contrast and cyclopean processing. *Vision Research*, 31, 2195–2207.
- Erkelens, C. J. (1987). Adaptation of ocular vergence to stimulation with large disparities. *Vision Research*, 66, 507–516.
- Erkelens, C. J. & Collewijn, H. (1991). Control of vergence: Gating among disparity inputs by voluntary target selection. *Experimental Brain Research*, 87, 671–678.
- Findlay, J. M. (1982). Global processing for saccadic eye movements. *Vision Research*, 22, 1033–1045.
- Findlay, J. M. & Harris, L. R. (1993). Horizontal saccades to dichoptically presented targets of differing disparities. *Vision Research*, 33, 1001–1010.
- Foley, J. M. (1976). Binocular depth mixture. *Vision Research*, 16, 1263–1267.
- Foley, J. M. & Richards, W. (1972). Effects of voluntary eye movement and convergence on the binocular appreciation of depth. *Perception and Psychophysics*, 11, 423–427.
- Hodges, L. F. (1992). Time-multiplexed stereoscopic computer graphics. *IEEE Computer Graphics and Applications*, 122, 20–30.
- Hruby, K. (1979). *Kurze Augenheilkunde*, 5th edn. München: Urban & Schwarzenbeck.

- Judge, S. J. (1991). *Vergence*. In Carpenter, R. H. S. (Ed.), *Eye movements*, Vol. 8 of *Vision and visual dysfunction*, Cronly-Dillon, J. (Gen. Ed.). Boca Raton: CRC Press.
- Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago: Chicago University Press.
- Krol, J. D. & van de Grind, W. A. (1986). Binocular depth mixture: An artefact of eye vergence? *Vision Research*, *26*, 1289–1298.
- Lehky, S. R. & Sejnowski, T. J. (1990). Neural model of stereoacuity and depth interpolation based on a distributed representation of stereo disparity. *Journal of Neuroscience*, *10*, 2281–2299.
- Mallot, H. A. & Arndt, P. A. (1992). Disparity-evoked vergence is directed towards average depth. *Investigative Ophthalmology and Visual Science, Suppl.*, *33*, 707.
- Mallot, H. A., Arndt, P. A. & Bülthoff, H. H. (1996a). A psychophysical and computational analysis of intensity-based stereo. *Biological Cybernetics*, in press.
- Mallot, H. A. & Bideau, H. (1990). Binocular vergence influences the assignment of stereo correspondences. *Vision Research*, *30*, 1521–1523.
- Mallot, H. A., Gillner, S. & Arndt, P. A. (1996b). Is correspondence search in human stereo vision a coarse-to-fine process? *Biological Cybernetics*, *74*, 95–106.
- Marr, D. & Poggio, T. (1979). A computational theory of human stereo vision. *Proceedings of the Royal Society (London) B*, *204*, 301–328.
- McKee, S. P. & Mitchison, G. J. (1988). The role of retinal correspondence in stereoscopic matching. *Vision Research*, *28*, 1001–1012.
- Mowforth, P., Mayhew, J. E. W. & Frisby, J. P. (1981). Vergence eye movements made in response to spatial-frequency-filtered random-dot stereograms. *Perception*, *10*, 299–304.
- Pobuda, M. & Erkelens, C. J. (1993). The relationship between absolute disparity and ocular vergence. *Biological Cybernetics*, *68*, 221–228.
- Poggio, G. F. (1995). Mechanisms of stereopsis in monkey visual cortex. *Cerebral Cortex*, *5*, 193–204.
- Pouget, A. & Sejnowski, T. J. (1994). A neural model of the cortical representation of egocentric distance. *Cerebral Cortex*, *4*, 314–329.
- Rashbass, C. & Westheimer, G. (1961). Disjunctive eye movements. *Journal of Physiology (London)*, *159*, 339–360.
- Richards, W. A. (1971). Anomalous stereoscopic depth perception. *Journal of the Optical Society of America*, *61*, 410–414.
- Schor, C. M. (1979). The relationship between fusional vergence eye movements and fixation disparity. *Vision Research*, *19*, 1359–1367.
- Semmlow, J. & Wetzell, P. (1979). Dynamic contributions of the components of binocular vergence. *Journal of the Optical Society of America*, *69*, 639–645.
- Semmlow, J. L., Hung, G. K., Horng, J.-L. & Ciuffreda, K. J. (1994). Disparity vergence eye movements exhibit preprogrammed motor control. *Vision Research*, *34*, 1335–1343.
- 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 a neural theory. *American Journal of Psychology*, *83*, 461–534.
- Stevenson, S. B., Cormack, L. K., Schor, C. M. & Tyler, C. W. (1992). Disparity tuning in mechanisms of human stereopsis. *Vision Research*, *32*, 1685–1694.
- Theimer, W. & Mallot, H. A. (1994). Phase-based binocular vergence control and depth reconstruction using active vision. *Computer Vision Graphics and Image Processing: Image Understanding*, *60*, 343–358.
- Weinshall, D. (1991). Seeing ‘ghost’ planes in stereo vision. *Vision Research*, *31*, 1731–1748.
- Westheimer, G. & Mitchell, D. E. (1969). The sensory stimulus for disjunctive eye movements. *Vision Research*, *9*, 749–755.

*Acknowledgements*—This work was supported by the Deutsche Forschungsgemeinschaft. We are grateful to Heinrich H. Bülthoff,

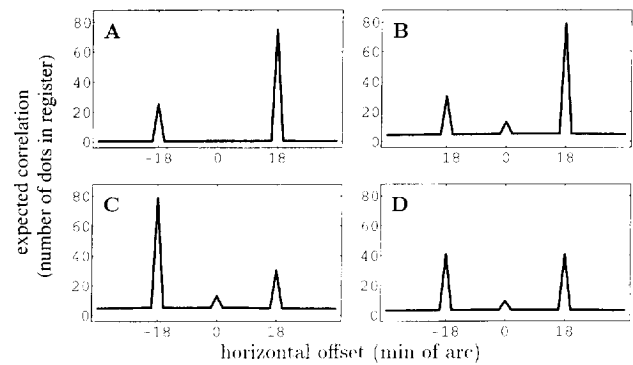


FIGURE A1. Expected correlation as a function of horizontal offset between the two half-images. Correlation is measured by the number of coincident dots (multiplied by their respective contrasts). As in the experiments, a total of 100 dots was simulated at disparities  $\pm 18'$ . (A) Ideal calculation for zero vertical disparity. Dot distribution:  $k_n = 25$ ,  $k_f = 75$ , contrasts  $c_n = c_f = 0.5$ . The expected correlations are  $\Phi(-18) = 25.4$ ,  $\Phi(0) = 0.44$  and  $\Phi(18) = 75.3$ . (B) Dots with vertical disparities of up to  $\pm 10'$  are counted also as matches and a crosstalk of 8% is included. The number of false matches increases;  $\Phi(-18) = 30.2$ ,  $\Phi(0) = 13.2$ ,  $\Phi(18) = 79.0$ . (C) Same as (B), with dot distribution inverted. The main correlation peaks are exchanged while the false-matches baseline and the crosstalk peak are unchanged. (D) Stimulus with original dot distribution ( $k_n = 25$ ,  $k_f = 75$ ) but unequal contrasts ( $c_n = 0.612$ ,  $c_f = 0.354$ ) chosen to cancel the effect of dot distribution. The false-matches baseline and the crosstalk peak are roughly as before;  $\Phi(-18) = 29.8$ ,  $\Phi(0) = 12.9$ ,  $\Phi(18) = 79.0$ .

Sabine Gillner, Eric Miska, and Thomas Vetter for discussions and help with the experiments.

## APPENDIX

### *Correlation Histograms for the Used Stimuli*

In the previous sections of this paper, it was assumed that the correlation histograms of the stimuli used had one or two sharp peaks at the offsets corresponding to the simulated depth positions of the planes and is independent of the stimulus condition otherwise. This assumption is not ideally satisfied. Here, we present a detailed analysis of the actual correlation functions, indicating that the agreement is, however, satisfactory.

Since all experiments were performed with a new random pattern of dots each time the adaptive procedure was repeated, it is sufficient to analyse the expected correlation between the left and right patterns. The images were composed of  $J = 170$  independent lines with  $I = 252$  pixels in the horizontal direction. Let  $N = IJ = 42,840$  denote the number of possible positions of a dot in each half-image. If we place  $K = 100$  dots in this raster without any two dots coinciding, the probability that a given position is hit is:

$$p_{\text{hit}} = K/N \approx 0.0023. \quad (\text{A1})$$

Strictly speaking, false-matches can occur only if two or more dots fall on the same line. The probability of this event occurring in a given line is:

$$P(m > 1) = \sum_{m=2}^I \binom{I}{m} p_{\text{hit}}^m (1 - p_{\text{hit}})^{I-m} \approx 0.118, \quad (\text{A2})$$

where  $m$  denotes the number of dots per line. The expected number of dots per line is  $E(m) = Ip_{\text{hit}}(1 - p_{\text{hit}}) \approx 0.587$ . The disparities of the false matches clearly depend on the distance of these dots on the line. For a given shift between the left and the right half-image, the expected number of false-matches is proportional to the overlap of the shifted

half-images. Figure A1(A) shows the expected correlation for a stimulus with 25 dots simulated at disparity  $-18$  min arc and 75 dots simulated at disparity  $+18$  min arc. The dot size ( $1.8 \times 1.8$  min arc) is accounted for by convolving the original histogram with the dot profile; it determines the width of the peaks and increases the baseline.

Figure A1(B) shows a more realistic calculation. If we assume that the image is blurred vertically or that matches at slight vertical disparities are also effective, the expected number of false matches increases. For the simulation shown in the figure, we have assumed that matches with vertical disparities in the range of  $\pm 10$  min arc are fusible and exert the same pulling strength as dots at zero vertical disparity. Since the probability of false matches is proportional to the overlap of the two half-images, they form a very shallow peak which in Fig. A1(B) looks like a baseline.

Crosstalk between the two half-images results in a narrow peak at zero disparity. Based on a psychophysical compensation experiment as well as on the characteristics of the shutter-glasses, the crosstalk was estimated to be in the order of 8%. In this case, each half-image seen through the appropriate eye can be matched to a weaker but otherwise identical "ghost image" seen through the other eye. As a result, a peak

of expected correlation at offset 0 is obtained which corresponds mainly to the number of matchable dots (100) multiplied by the amount of crosstalk. Also, additional false matches between each half-image and its "ghost image" become possible. Both effects, crosstalk and false-matches, with non-zero vertical disparities are shown together in Fig. A1(B).

It is important to note that the effects of crosstalk and false-matches are almost completely independent of the stimulus conditions used. Figure A1(C) shows the expected correlation histogram for a stimulus with reversed numbers of dots in the two planes. The stimuli of Fig. A1(B) and (C) lead to different results for elicited vergence as was shown in Experiment 2. Since the expected false-matching correlation is the same in both cases, and since it is rather small, we conclude that the measured effects are due to the reversed peaks at disparities  $\pm 18$  min arc. Figure A1(D) shows a stimulus condition from Experiment 4 where the contrast values were chosen to cancel the dot distribution in the correlation histogram. Again, the false-matches baseline and the crosstalk peak are about the same as in the other conditions shown.