The role of a local reference in stereoscopic detection of depth relief

Yury Petrov and Andrew Glennerster

University Laboratory of Physiology, Parks Road, Oxford, OX1 3PT, UK

September 17, 2003

Abstract

Stereoacuity thresholds have been shown to depend on the disparity of a point with respect to a slanted reference plane through neighbouring points (Glennerster et al., 2002). Here we explored a wider range of conditions, including slanting the reference points about a horizontal axis and varying the spacing of the reference dots, allowing alternative hypotheses for the effect to be distinguished. The stimulus consisted of 3 dots; the outer two defined a line that was slanted in depth. Observers judged in which of two intervals the third, central dot was displaced from the location midway between the outer reference dots. The displacement consisted of both a disparity and a shift in the fronto-parallel plane. We compared performance for pairs of conditions in which the disparity was the same but the fronto-parallel shifts were in opposite directions. Models based purely on relative disparity predict that performance should be the same for these conditions. We found consistent differences: performance was always better when the target had a greater disparity with respect to the line joining the reference dots. The other stimulus parameters varied were: target disparity (concave/convex), stimulus size (large/small), slant sign (sky/ground) and axis (vertical/horizontal). The results suggest that either a) disparity with respect to the line drawn through the outer reference dots or b) difference in disparity gradients on either side of the target determines the depth discrimination threshold for these stimuli.

Keywords: psychophysics; stereopsis; binocular disparity; gradient; threshold

1 Introduction

It is well established that the perceived depth of a point is strongly dependent on the layout of features surrounding it (e.g. Anstis et al., 1978; Mitchison and Westheimer, 1984; Westheimer, 1986; Westheimer and Levi, 1987; Parker and Yang, 1989; Glennerster and McKee, 1999). Other aspects of stereo processing such as stereo correspondence, stereoacuity and binocular-fusion limits are influenced in similar ways. Thus, surrounding objects have been shown to determine which binocular matches were made in an ambiguous stereogram (Kontsevich, 1986; Mitchison and McKee, 1987; McKee and Mitchison, 1988). This surround effect increases as the separation between the ambiguous target and nearby objects becomes smaller (Petrov, 2002). Stereoacuity thresholds have also been shown to change with the separation between the target and reference (McKee et al., 1990). At the opposite end of the disparity scale, binocular fusion and correspondence limits have been shown to depend on the critical disparity gradient value, above which simultaneous stereo-processing of nearby features becomes impossible (Tyler, 1974; Burt and Julesz, 1980; Schor et al., 1989).

Evidence on stereo correspondence and stereoacuity are especially significant in understanding the early stages of stereo processing. This is clear in the case of the correspondence process and stereo matching limits, since binocular matching must necessarily precede the recovery of depth, but it is also likely to be the case for stereaoacuity. There is evidence, for example, that stereoacuity (expressed in angular terms) is largely independent of viewing distance (Ogle, 1958). Because perceived depth changes with viewing distance, this is evidence that stereoacuity is limited by processes before rather than after the recovery of metric depth.

Two of these methods - measures of correspondence (Mitchison and McKee, 1987) and stereoacuity (Glennerster and McKee, 1999; Glennerster et al., 2002) - suggest that the stereo system uses some measure of disparity relative to a locally-defined reference frame at an early stage in processing. For example, in Glennerster et al. (2002), a regular grid of dots slanted around the vertical axis was presented, and the subjects' sensitivity for detecting displacements of the central column of dots was measured. The results demonstrate that the sensitivity was determined by the distance of the column from the plane of the grid, rather than by the change of its relative disparity.

We have investigated whether the presence of a plane is necessary by reducing the stimulus to its most minimal form: two reference dots defining a reference line that is slanted in depth and a target dot whose position must be judged with respect to that line. Figure 1a illustrates the type of stimulus we used. Two reference points, A and C, define a slanted line. The subject's task was to detect, in a two-interval forced choice design, the interval in which the target point T had been displaced from its central location, T_0, to one of the two test locations, T_1 or T_2. We were interested in the cue that the visual system uses to detect this displacement.

We classified the possible cues in our task into four categories based on orders of spatial derivatives of the disparity signal:

- I: zeroth order, e.g. relative disparities between the target dot T and reference dots A and B.

- II: first order, e.g. disparity gradients between the target dot T and reference dots A and B.

- III: second order, e.g. change in disparity gradient at B normalised by the angular distance AB (which is a measure of disparity curvature).

- IV: other disparity cues: e.g. disparity gradient difference at the target point T, disparity relative to the reference line AB, or signals proportional to these in the current experiments.

If the critical cue is disparity of the target relative to the two reference points (category I), then results for position T_1 and T_2 should be the same (see experiment 1). If observers are using principally the disparity gradient between a pair of points, then removing the third point should have little effect (experiment 2). If observers use disparity curvature, then reducing the size of the stimulus should improve performance (experiment 2).

The last category includes two relevant cues that can not be represented as a spatial derivative of disparity. Thus, disparity gradient difference has to be normalized by the angular distance AB to give a second derivative of disparity. These two cues are not identical, but can not be distinguished using the experiments described here. In the Discussion we suggest alternative experiments to discriminate between the the two cues. Note that for each of the cues mentioned here the underlying measures used by the visual system could be related to them in an arbitrary monotonic way. The experiments described here can do no more than identify the category of cue most consistent with the data.

In addition to stereoscopic cues, monocular lateral displacement of the target dot could provide a cue. Earlier studies showed that monocular displacement thresholds are several times higher than stereoacuity thresholds (e.g. Westheimer and McKee, 1979). Here sensitivity for monocular displacement of the target dot was checked for two observers (YP and AG). In agreement with the earlier studies monocular detectability was found to be approximately three times lower than stereoscopic detectability. Thus, the thresholds we report are primarily a reflection of stereoscopic detection mechanisms.

Our results show that, for a series of different conditions, performance is consistent with the visual system using the disparity of the target with respect to a local reference line or some cue that is proportional to this value (category IV).

2 Method

2.1 Apparatus

Stimuli were generated on a Sun Ultra-10 Workstation and displayed on two high-resolution colour monitors (Flexscan T961, Eizo). Stereo images were viewed via a modified Wheatstone stereoscope at a viewing distance of 2.65 metres. The display was 1600x1280 pixels, and each pixel subtended 18 seconds of arc. Anti-aliasing of circular dot edges was used to generate sub-pixel resolution. Stimuli were viewed in a dark room. The background luminance was very low ($< 0.1\ cd/m^2$), and the stimuli were bright (55 cd/m^2).

2.2 Subjects

All three observers had normal or corrected monocular visual acuity: two were experienced stereo-observers (YP, AG), while the third (GF) had not previously taken part in a psychophysical experiment, and was naive to the hypothesis tested. All subjects were allowed to train for 30 minutes before the beginning of the experiment.

2.3 Psychometric Procedure

We used a two-interval forced-choice paradigm. A stimulus, such as that shown in Fig. 1b, was displayed in one interval with the target dot T placed exactly half way between the two reference dots A and B (position 0, shown by the open circle in Fig. 1a). In the other interval, the target dot was displaced. The presentation order was randomised and the subject's task was to identify the interval in which the target dot had been displaced. For each subject the displacement magnitude was chosen approximately at his/her detection threshold, i.e. at 75% correct performance. Although the displacement was zero in one of the intervals, the stimulus normally appeared to have a small degree of depth curvature in both intervals, sometimes even in the direction opposite to the one actually shown. This is characteristic of sub-threshold and near-threshold perception. Subjects were instructed to indicate the interval in which the target dot appeared either closer or farther away, depending on the curvature of the stimulus in the particular set. In a run of 100 trials, two different target displacements were tested (position 1 and 2 in Figs. 1 to 4), with the trials testing each displacement randomly interleaved. The data in Figs. 2 to 4 shows the proportion of correct responses made over at least 200 trials for each condition. Error bars in these plots show the standard error of the mean computed from the binomial distribution.

2.4 Stimuli

In all of the experiments except one, the separation between the reference dots A and B was 86 arcmin as viewed from the cyclopean point. The relative disparity between them was 8.6 arcmin, i.e the disparity gradient was ± 0.1. In the half-scale experiment (Fig. 3a), all the distances between points in the stimulus were halved (hence also the disparities). The size of the dots, however, remained as 2.5 arcmin diameter. The dots in all experiments were blurred with a Gaussian kernel for the purposes of anti-aliasing. The disparity gradient between points was not affected by the reduction in scale.

The parameters determining the layout of the points are shown in Fig. 1d. Although only one stimulus type is shown (the one illustrated in Fig. 1a), the parameters apply to all the stimuli, including the horizontal stimulus in experiment 3. D is the cyclopean angular separation between the reference dots and the central dot in position 0. 2Δ is the corresponding relative disparity between the central dot and either of the reference dots. The displacement of the target dot T consisted of a displacement d in the fronto-parallel plane and a disparity change 2δ. The direction of the displacement d was vertical in experiments 1 and 2 and horizontal in experiment 3. The two different positions of the target (position 1 and 2) had the same disparity 2δ with respect to the undisplaced position (position 0) but displacement d had opposite signs. This means that relative disparities between target and reference points were the same for the two displaced positions and so were their cyclopean separations, providing in both cases that the ordering is ignored. The corresponding disparity gradients, on the other hand, were different since here it does matter in which order the disparities and cyclopean separations are paired (see section 2.5).

The different configurations we tested are illustrated in Fig. 1c, with the central dot shown in position 1 in each case. In experiments 1 and 2, we tested stimuli using a vertically aligned trio of dots, slanted about a horizontal axis so that the top reference dot A was nearer to the observer than the bottom reference dot B ('sky' slant) or vice versa ('ground' slant). The target dot T could be displaced either toward the observer (convex curvature) or away (concave curvature). In experiment 3, the stimulus orientation was horizontal and the slant was about a vertical axis.

Each of the two stimulus intervals lasted 1.5 seconds, during which the subject was free to move their eyes. The screen was blank in the 1 second inter-stimulus interval. Before the first stimulus interval, a fixation stimulus was presented for 1 second. It consisted of a central diamond outline (36 arcmin) and four bright dots forming a 4° square, also centred on the midpoint of the screen. This stimulus provided a visual reference for fronto-parallel. After the second interval the screen was blank until the subject gave their response which triggered the next stimulus to be displayed.

2.5 Disparity gradients

This section gives formulae defining the disparity gradients between the target and reference dots. We show later that the change in disparity gradient at the target dot is predictive of subjects' performance in the task.

Referring to Fig. 1d, the two disparity gradients between the reference dots and the target dot in position 1 are given by

$$\nabla_1^l = 2\,\frac{\Delta + \delta}{D - d} \quad \text{and} \quad \nabla_1^s = 2\,\frac{\Delta - \delta}{D + d} \tag{1}$$

where ∇^l and ∇^s stand for the larger and smaller of the two gradients, respectively. Definitions of the lengths D, Δ, d and δ are given in section 2.4. They apply equally to the horizontal stimuli used in experiment 3, giving rise to the same formulae for disparity gradients. Disparity gradients for position 2 are obtained by changing the sign of the vertical shift d:

$$\nabla_2^l = 2\,\frac{\Delta + \delta}{D + d} \quad \text{and} \quad \nabla_2^s = 2\,\frac{\Delta - \delta}{D - d}\ . \tag{2}$$

In this experiment, δ/d was always larger than Δ/D, i.e. positions 1 and 2 were both located either in front of or behind the reference line AB. It is easy to see that $\nabla_1^l \geq \nabla_2^l \geq \nabla_2^s \geq \nabla_1^s$, while for their differences:

$$\nabla_1^l - \nabla_1^s = 4\,\frac{\delta D + \Delta d}{D^2 - d^2} \quad \text{and} \quad \nabla_2^l - \nabla_2^s = 4\,\frac{\delta D - \Delta d}{D^2 - d^2}\ . \tag{3}$$

This shows that the stimulus with the target dot at position 1 has both the largest and the smallest disparity gradient, so that the difference of the disparity gradients is larger than when the target is at position 2.

Disparity curvature is closely related to disparity gradient difference. It is defined as the rate of change of disparity gradient over visual angle (i.e. it is the limit of $(\nabla^l - \nabla^s)/\phi$, where ϕ stands for the visual angle separating points A and B (Rogers and Cagenello, 1989)). For our stimulus, one measure of disparity curvature is disparity gradient difference divided by the visual angle $2D$.

3 Results and Discussion

3.1 Experiment 1

The purpose of this experiment was to test stereoacuity performance differed for target positions 1 and 2. If so, then cues other than relative disparity must be responsible for the difference (see Introduction). The stimulus was arranged in the four configurations shown in Fig. 1c: ground convex, ground concave, sky convex (YP and AG) and sky concave (GF only). The results are shown in Fig. 2 as the proportion of correct responses, i.e. when the displacement of the target dot was detected. One can see that for position 1, detection was always better for all observers and in all configurations studied. The difference was significant ($P < 0.05$) with the exception of GF ground convex results. This rules out the possibility that performance is determined by relative disparity between the target and reference dots, or type I cues in general. Instead, performance can be explained by cues in the remaining categories, e.g. the largest disparity gradient ∇^l, disparity curvature $(\nabla^l - \nabla^s)/2D$, disparity gradient difference $\nabla^l - \nabla^s$, or disparity relative to the reference line AB.

The target also has a component of displacement along the line AB, which could in principle provide an independent cue and affect performance. In order to check whether this component has an important effect in practice we ran a control experiment on one subject (AG) in which we repeated one of the conditions (vertical orientation, ground slant) with the target at position 1, position 2 or at two locations with the same values of vertical component as at positions 1 and 2, but lying on the line AB. Performance was close to chance (0.55) in this case. Thus, the displacement along the line does not appear to be a very useful cue in this type of experiment. One reason is likely to be that the vertical displacement could be either up or down. This means that comparing the vertical position of the target between the two intervals in a trial is not a useful strategy. The same is not true of the depth component, which was always in the same direction.

3.2 Experiment 2

The results of Experiment 1 indicate that relative disparity alone cannot determine stereoacuity thresholds. To further narrow the alternative explanations the stimulus was modified by reducing all distances and disparities by a factor of two. Disparity gradients for this scaled version of the stimulus are the same as before (and hence disparity gradient differences too), but both relative disparities and disparity curvatures are changed: disparities decrease by a factor of two, while curvatures increase by the same factor (see section 2.5). Therefore, assuming that the noise remains constant, variations in the magnitudes of these cues should determine detectability.

In fact, as Fig. 3a shows, performance was almost the same for the half-scale stimulus as it was for the full size stimulus (Fig. 2). The difference in performance between position 1 and 2 remains ($F(1, 1194) = 19.72$; $P < 10^{-5}$)[1]. Overall, there was a drop in detectability of the target displacement

[1]2-way ANOVA results: subjects and target position (1 or 2) were treated as two

for the half-scale compared to the full size stimulus (a ratio of 0.82 ± 0.07, which is significantly different from 1: $F(1, 2388) = 7.37$; $P < 0.0067$)[2].

This result appears to be incompatible with disparity curvature as a cue (category III), since it predicts an improvement in performance by a factor of two. It also appears to be incompatible with disparity with respect to the reference line (one of the category IV cues) since it predicts that detectability should halve for the half-scale stimulus. However, there is a possible 'get out clause' for the latter cue if the original assumption of constant noise is challenged. Thus, if both signal and noise are halved in the half-scale stimulus, performance should stay the same. (This is a familiar explanation of Weber's law).

There is evidence for a linear relationship between stimulus scale and stereoacuity (i.e. Weber's law behaviour), at least at large scales (McKee et al., 1990). Performance on the half-scale stimulus was not exactly the same as for the standard size (it dropped by 18%) which does not fit precisely the predictions of Weber's law. However, a careful look at previous stereoacuity

independent parameters. Variance within each 2-way ANOVA cell was calculated assuming a binomial distribution for a measurement mean in each cell (given 200 trials for each cell).

[2] 3-way ANOVA results: subjects, target position (1 or 2), and stimulus scale were treated as three independent parameters.

data shows a similar deviation from Weber's law at smaller scales. Stereo thresholds decrease approximately linearly as the separation between target and reference lines becomes smaller, but then reach a plateau (McKee et al., 1990) and finally rise at very narrow separations (Westheimer and McKee, 1979). For the half-scale stimulus, the separation between the central and flanking dots was decreased to 21.5 arcmin, which falls into the 'plateau' range. This means that the performance for the half-scale stimulus can be explained by a halving of the signal strength and a not-quite halving of the noise.

Taken together, then, Experiments 1 and 2 rule out cues in categories I and III, which leaves cues in categories II (e.g. disparity gradient) and IV (e.g. disparity gradient difference or disparity with respect to the reference line) as possible candidates. It has been shown that surface curvature and surface discontinuities can be more easily discriminated than surface slant (e.g. Gillam et al., 1984; Rogers and Cagenello, 1989). This suggests that type IV cues may be of more importance for stereoacuity than type II cues.

To check this hypothesis, Experiment 1 was repeated with the upper reference dot of the stimulus omitted (the remaining two dots were vertically oriented and had ground slant). Type IV cues were thus removed. The observer's task was to report in which of the two AFC intervals the target (upper dot here) was farther away from the observer. For this stimulus, performance dropped dramatically for AG and GF (Fig. 3b), which suggests that these two subjects indeed used type IV cues as the principal depth cues. YP's performance also deteriorated significantly after the upper reference dot was removed, but the change was not as large. This indicates that in some cases, type II cues can provide a significant degree of depth information.

3.3 Experiment 3

There is a well-known horizontal/vertical anisotropy in the perception of stereoscopic slant (Wallach and Bacon, 1976; Rogers and Graham, 1983). A vertical gradient of disparity (slant about a horizontal axis) is usually much more readily perceived than a horizontal gradient (slant about a vertical axis), although the magnitude of the anisotropy differs markedly between subjects (Mitchison and Westheimer, 1990; Mitchison and McKee, 1990). It was interesting, therefore, to determine if the effect of type IV cues on stereoacuity depended on stimulus orientation. To this end, the experiment was repeated with the stimuli oriented horizontally, as shown in the lower panel of Fig. 1c. The magnitude of disparity gradient characterising the horizontal slant was 0.1, as in experiments 1 and 2. As can be seen comparing the lower panel of Fig. 2 with Fig. 4, the results for the horizontal and vertical stimulus orientations are nearly the same for one of the subjects (YP). Performance of the other two subjects deteriorated dramatically with horizontal stimulus orientation. Nevertheless, considering the data of all three observers together, the difference in performance between target positions 1 and 2 still remained ($F(1, 1194) = 11.23$; $P < 10^{-3}$). Interestingly, the data correlates with subjects' perception of the horizontal stimulus. Subject YP perceived no difference in the stereoscopic slant between vertical and horizontal orientations, while AG and GF found both the slant much less obvious and the task more difficult for the horizontally oriented stimulus.

4 General Discussion

4.1 Main Results

The results reported in this paper extend previous findings showing that sensitivity to stereoscopic depth is dependent on the disparity of a target point with respect to a local reference plane even when the points that define the plane are some distance away. This disparity with respect to the interpolated plane is a better predictor of stereoacuity thresholds than the relative disparity of the target with respect to neighbouring points (Glennerster et al., 2002). Here we have shown that it is not necessary to use a slanted reference plane to demonstrate the effect: the minimal stimulus, consisting of two reference points and a target point, will suffice. Second, we have shown that the effect of slant on stereo detection performance, which had previously been demonstrated for stimuli slanted about a vertical axis (Glennerster et al., 2002), applies also to slants about a horizontal axis (vertical stimuli, Fig. 2). In fact, we found that the effect tended to be stronger in this case. Third, we have shown that performance on this task is approximately scale invariant, varying very little when the entire stimulus, including the target displacements, was scaled by a factor of two. All the experiments suggest that, whatever the measure is that is important to the visual system in detecting depth displacements, lateral separation and disparity must be inextricably linked.

A parsimonious account of the data in this paper is that performance is limited by disparity relative to the reference plane or some cue that is a monotonic function of this quantity in the conditions we have tested (category

IV disparity cues). For example, difference in disparity gradient between the target and each of the two reference points falls into this category. On the other hand, disparity relative to either of the reference points will not do, as it cannot account for the differences in performance between target locations 1 and 2 (Figures 2, 3 and 4), which differ only in position on a fronto-parallel plane. Disparity gradient, also, does not seem to be a viable candidate. One example that demonstrates this is the poor performance of subjects GF and AG using the two-dot stimulus (Fig. 3b). Disparity curvature is another possible cue (Rogers and Cagenello, 1989). Disparity curvature is doubled when the stimulus size is halved whereas, when we halved stimulus size, subjects' performance slightly deteriorated (Fig. 3a). This is incompatible with disparity curvature being the limiting factor.

We have presented our data as evidence that these latter disparity cues (categories I, II, and III) can not explain the results. Instead, two category IV cues were suggested as the primary cues: disparity gradient difference or disparity with respect to the reference line AB. In fact, for the stimuli used in our study the two cues are proportional to each other. For the vertically oriented stimulus the disparity cue amounts to calculating horizontal displacement s of the target point T from the AB line in each eye, as shown in Fig. 1d. It is easy to show that $s = \delta + d\Delta/D$. Comparing this formula to equation (3) one can see that s equals the disparity gradient difference, $\nabla^l - \nabla^s$ times $(D^2 - d^2)/4D$. The factor $(D^2 - d^2)/4D$ was the same for our two critical conditions, because position T_1 and T_2 differ only in the sign of d. In order to study the two cues independently, performance for stimuli for which this factor is significantly different could be compared (for example when $d = 0$ versus $d \simeq D$). There are also differences in predictions for these cues for stimuli with more than two reference points. A paper reporting the study of these stimuli is under preparation.

There is one case in which disparity with respect to the reference line AB appears to be a poor predictor of our data. When the scale of the stimulus was reduced by 50% (Figure 3a), the disparity of the target dot with respect to the line AB was also halved, yet performance remained about the same. However, as discussed in section 3.2, this would be the case if the noise limiting performance was reduced when the separation of the reference features was smaller.

Other evidence supports the conclusion that disparity with respect to the reference line AB or disparity gradient difference cannot be the only factor limiting stereoacuity. For example, at spatial frequencies above 0.4

c/deg stereoacuity thresholds rise rapidly, despite increased disparity gradients (e.g. Tyler, 1974; Rogers and Graham, 1982). A very similar trend was found for stereoacuity thresholds tested using a single target line and two flanking reference lines (Westheimer and McKee, 1979, see section 3.2). Decreasing the separation between target and reference below 10 minutes of arc caused a rapid rise in thresholds. Both examples show that where modulations of disparity occur at a very small spatial scale, other parameters limit stereoacuity.

4.2 Psychometric function fit

If category IV cues are critical in our experiment, then the data should all coincide when plotted against the magnitude of these cues in each stimulus. In Fig. 5, we have plotted our data against the disparity gradient difference at the target but we could equally have used the disparity of the target with respect to the reference line AB, normalized by the length AB. Positions 1 and 2 of the target have different disparity gradient differences as shown in Fig. 1 and equation (3). Fig. 5 shows the data from Figures 2, 3, and 4 re-plotted in this way. The two-dot data (Fig. 3b) are plotted on the same axis for comparison even though, of course, there is no disparity gradient difference in this condition.

We have fitted the data with empirical Weibull functions, $1-(1-\gamma)e^{-(x/\alpha)^\beta}$ (Pelli, 1985), where the false alarm rate γ was fixed at 0.5 for the 2AFC experiment, and x is the disparity gradient difference. In doing so, we are making the assumption that disparity gradient difference, or some cue proportional to it, is the sole determinant of performance. This is not wholly justified, as discussed in section 3.1, so the fits should be treated with some caution. In fitting Weibull functions to the data, we separated the two-dot and the horizontal conditions from the rest because two of the observers performed markedly worse in these cases. The relative insensitivity of many observers to horizontal changes in disparity is well documented (Mitchison and McKee, 1990; Cagenello and Rogers, 1993). The poor performance in the two-dot case shows that two of the observers, and to a lesser extent the third subject also, are relatively insensitive to gradient changes alone, compared to their sensitivity for detecting 'bumps'. The result fits with previous observations (e.g. Gillam et al., 1984; Rogers and Cagenello, 1989; Lappin and Craft, 2000).

Although there are only two data points on these psychometric functions (and a third constraining point at zero disparity gradient difference), it is still possible to conclude something about their shape. The fit for observer AG follows the classical cumulative Gaussian shape, as predicted by the simplest form of signal detection theory. However, for observers YP and GF, the drop in performance as the disparity gradient difference approaches zero occurs faster than predicted by the cumulative Gaussian fit. This sigmoid shape is characterized by Weibull parameter β being larger than $\sqrt{2}$. The fast drop is also typical for psychometric functions obtained in contrast discrimination experiments (Nachmias and Sansbury, 1974; Legge, 1978). One possible explanation of this behaviour is that the actual signal processed by a signal detector is a nonlinear (accelerating) function of the cue assumed in the psychometric task (disparity gradient difference here) (Foley and Legge, 1981). Alternatively, it can be explained by the probability summation over one or more signal detectors with noise introduced by N detectors, which are irrelevant to the given task (Pelli, 1985). Because of the inherent uncertainty over which one of the detectors is relevant, there is a threshold for the relevant signal below which it is overshadowed by the irrelevant noise, and below which subjects' performance quickly reduces to chance. For subjects YP and GF, the fitted β values were 2.35 ± 0.75 and 2.28 ± 0.98 (corresponding to $N \simeq 25$ (Pelli, 1985)), which is close to the β values reported for the contrast discrimination experiments.

4.3 Ecological plausibility of the proposed cues

One advantage of measures like disparity gradient difference or disparity relative to reference line AB is that they provide a more compressed (less redundant) description of object shape than disparity. This is because they vary less when the observer views the surface from different angles. For example, relative disparities and disparity gradients depend to a far greater extent on the surface orientation with respect to the fronto-parallel plane. Given all surfaces viewed from all angles, relative disparities and disparity gradients will have a broader distribution than disparity gradient difference or disparity relative to reference line AB. As a result, the former distribution is characterised by a larger measure of Shannon entropy (Shannon, 1948) and requires a larger computational capacity to be processed.

The most compressed representation of an object is one that is invariant to the observer's viewing position. Such object constancy is clearly a desirable goal, aside from the issue of storage efficiency. Measuring disparity relative to reference line AB is a useful step towards computing a representation of this sort. Given that the projection of a surface in each eye is locally approximately orthographic, such a relief measure can be used to construct a 3D surface representation that is invariant under rotations of the surface or movements of the observer relative to the surface. For example, if there are several points at different depths relative to the line AB (e.g. T_0, T_1 and T_2 in Figure 1a), then the ratios of their disparities with respect to the line AB will remain invariant under rotations of the 'surface' (set of points) relative to the observer. A more general scheme is as follows.

From an observer's perspective a point in space is described by two coordinates of cyclopean visual direction and one of disparity. This is a familiar representation which we call a 'cyclopean space' here. Assuming that the retinal projections are orthographic, an arbitrary rotation of the surface as well as a change in observer's viewpoint is described by an affine (linear) transformation in the cyclopean space. This means that the coordinates of all surface points are transformed in exactly the same way. Four non-coplanar surface points can be chosen as a surface-centred reference frame. One point serves as an origin, while the others form three basis vectors which span the cyclopean space. The advantage of the surface-centred frame is that rotations and viewpoint transformations affect the points defining the reference frame in exactly the same way as all the other surface points. Therefore, described in terms of these basis vectors, the surface representation is invariant to viewing transformations.

A similar surfaced-centred representation was described by Koenderink and van Doorn (1991) in their structure-from-motion algorithm. Given a sequence of orthographic views, the algorithm divides the task of surface reconstruction into two stages: an initial stage using two views, which can recover surface shape modulo a linear relief transformation (depth-scaling and shear), and a calibration stage, at which a metric reconstruction is obtained by using a third view. An analogous two-stage process can be applied to the combination of two binocular images, where the second stage uses viewing geometry parameters (such as vergence angle, ϑ, and azimuth, υ) rather than a third view.

The invariance of this type of surface-based representation to viewing transformations holds despite the fact that it is not a full, metric reconstruction (which is by definition invariant to these transformations). Metric, or Euclidean, reconstruction requires the knowledge of viewing geometry parameters ϑ and υ and is often inaccurate (e.g. Johnston, 1991; Glennerster et al., 1996, 1998). Depending on the task, it may not be necessary to recover the metric structure of a surface - a simpler representation may be adequate (Koenderink and van Doorn, 1991; Tittle et al., 1995; Glennerster et al., 1996). For example, the metric structure is not required for the task subjects performed in our experiments.

A crucial stage in describing a surface in the surface-centred frame is finding the relative disparity between a given surface point and its orthographic projection onto a surface-centred coordinate plane. The disparity with respect to the reference line AB ($2s$ in our stimuli) is an example of such a disparity measure calculated for the case of a 2D depth-profile discussed here.

4.4 Conclusion

The stereoscopic system appears to be good at detecting 'bumps', i.e. deviations from a plane or line defined by surrounding points. Performance in this task cannot be predicted simply on the basis of the relative disparities of points or the disparity gradient between pairs of points. Nor can a measure like disparity curvature explain the data. Candidates for the signal include the difference in disparity gradients on either side of the target and disparity with respect to the reference line or plane. If the second type of signal is to explain the data, then the noise limiting performance must diminish when the separation of the reference features is reduced.

Acknowledgements

We would like to thank Andrew Parker and Suzanne McKee for helpful discussions, and three anonymous reviewers for their comments on the earlier variant of this paper. This work was supported by the Wellcome Trust (Grant 056657). AG is a Royal Society University Research Fellow.

References

Anstis, S. M., Howard, I. P., and Rogers, B. J. (1978). A Craik-O'Brien-Cornsweet illusion for visual depth. *Vision Research*, 18:213–217.

Burt, P. and Julesz, B. (1980). Disparity gradient limit for binocular fusion. *Science*, 208:615–617.

Cagenello, R. and Rogers, B. J. (1993). Anisotropies in the perception of stereoscopic surfaces - the role of orientation disparity. *Vision Research*, 33:2189–2201.

Foley, J. M. and Legge, G. E. (1981). Contrast detection and near-threshold discrimination in human vision. *Vision Research*, 21:1041–1053.

Gillam, B., Flagg, T., and Finlay, D. (1984). Evidence for disparity change as the primary stimulus for stereoscopic processing. *Perception and Psychophysics*, 36:559–564.

Glennerster, A. and McKee, S. P. (1999). Bias and sensitivity of stereo judgements in the presenece of a slanted reference plane. *Vision Research*, 39:3057–3069.

Glennerster, A., McKee, S. P., and Birch, M. D. (2002). Evidence of surface-based processing of binocular disparity. *Current Biology*, 12:825–828.

Glennerster, A., Rogers, B. J., and Bradshaw, M. F. (1996). Stereoscopic depth constancy depends on the subject's task. *Vision Research*, 36:3441–3456.

Glennerster, A., Rogers, B. J., and Bradshaw, M. F. (1998). Cues to viewing distance for stereoscopic depth constancy. *Perception*, 27:1357–1365.

Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. *Vision Research*, 31:1351–1360.

Koenderink, J. J. and van Doorn, A. J. (1991). Affine structure from motion. *J. Opt. Soc. Am. A*, 8:377–385.

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

Lappin, J. S. and Craft, W. D. (2000). Foundations of spatial vision: from retinal images to perceived shapes. *Psychological Review*, 107:6–38.

Legge, G. E. (1978). Sustained and transient mechanisms in human vision: temporal and spatial properties. *Vision Research*, 18:69–81.

McKee, S. P. and Mitchison, G. J. (1988). The role of retinal correspondence in stereoscopic matching. *Vision Research*, 28:1001–1012.

McKee, S. P., Welch, L., Taylor, D. G., and Bowne, S. F. (1990). Finding the common bond: stereoacuity and the other hyperacuities. *Vision Research*, 30:879–891.

Mitchison, G. J. and McKee, S. P. (1987). Interpolation in stereoscopic matching. *Vision Research*, 27:285–294.

Mitchison, G. J. and McKee, S. P. (1990). Mechanisms underlying the anisotropy of stereoscopic tilt perception. *Vision Research*, 30:1781–1791.

Mitchison, G. J. and Westheimer, G. (1984). The perception of depth in simple figures. *Vision Research*, 24:1063–1070.

Mitchison, G. J. and Westheimer, G. (1990). Viewing geometry and gradients of horizontal disparity. In Blakemore, C., editor, *Vision: coding and efficiency*. Cambridge University Press, Cambridge.

Nachmias, J. and Sansbury, R. V. (1974). Grating contrast: discrimination may be better than detection. *Vision Research*, 14:1039–1042.

Ogle, K. N. (1958). Note on stereoscopic acuity and observation distance. *J. Opt. Soc. Am.*, 48:794–798.

Parker, A. J. and Yang, Y. (1989). Spatial properties of disparity pooling in human stereo vision. *Vision Research*, 29:1525–1538.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *J. Opt. Soc. Am.*, 2:1508–1531.

Petrov, Y. (2002). Disparity capture by flanking stimuli: A measure for the cooperative mechanism of stereopsis. *Vision Research*, 42:809–813.

Rogers, B. J. and Cagenello, R. (1989). Disparity curvature and the perception of three-dimensional surfaces. *Nature*, 339:135–137.

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

Rogers, B. J. and Graham, M. (1983). Anisotropies in the perception of three-dimensional surfaces. *Science*, 221:1409–1411.

Schor, C., Heckmann, T., and Tyler, W. (1989). Binocular fusion limits are independent of contrast, luminance gradient and component phases. *Vision Research*, 29:821–835.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379–423.

Tittle, J. S., Todd, J. T., Perotti, V. J., and Norman, J. F. (1995). A hierarchical analysis of alternative representations in the perception of 3-D structure from motion and stereopsis. *Journal of Experimental Psychology: Human Perception and Performance*, 21:663–678.

Tyler, C. W. (1974). Depth perception in disparity gratings. *Nature*, 251:140–142.

Wallach, H. and Bacon, J. (1976). Two forms of retinal disparity. *Perception and Psychophysics*, 19:375–382.

Westheimer, G. (1986). Spatial interaction in the domain of disparity signals in human stereoscopic vision. *Journal of Physiology*, 370:619–629.

Westheimer, G. and Levi, D. M. (1987). Depth attraction and repulsion of disparate foveal stimuli. *Vision Research*, 27:1361–1368.

Westheimer, G. and McKee, S. P. (1979). What prior uniocular processing is necessary for stereopsis? *Invest. Ophtalmol. Visual Sci.*, 18:614–621.

Fig. 1. (a) A schematic diagram of the vertical three-dot stimulus used in experiment 1 (side view). Observers judged in which of two intervals the central dot, T, was shifted away from position 0 to positions 1 or position 2. (b) Stereo pairs of the stimulus (ground concave). The left and right eye images in the first two columns are arranged for uncrossed fusion (for crossed fusion use the last two columns). Stimuli with the central dot in positions 0, 1, and 2 are shown in the first, second and third row respectively. (c) Different stimulus configurations tested in experiments 1, 2 and 3. (d) The geometry of the stimulus for the left eye with the target dot, B, in position 1. Δ, D, d, and δ indicate the relative positions of the dots in monocular images (see text). Two disparity gradients, ∇_s (small) and ∇_l (large), between the target and the reference dots are illustrated by angles θ_s and θ_l, and are given by $2\tan\theta_s$ and $2\tan\theta_l$ respectively.

Fig. 2. Results of experiment 1. The proportion of correct responses is shown on the y-axis for conditions in which the target dot, T, was displaced to location 1 or 2, as shown in the cartoons on the right. The open circle marks the location of T in the non-signal interval of each trial. Results are shown for three observers and for three conditions: 'ground convex', 'ground

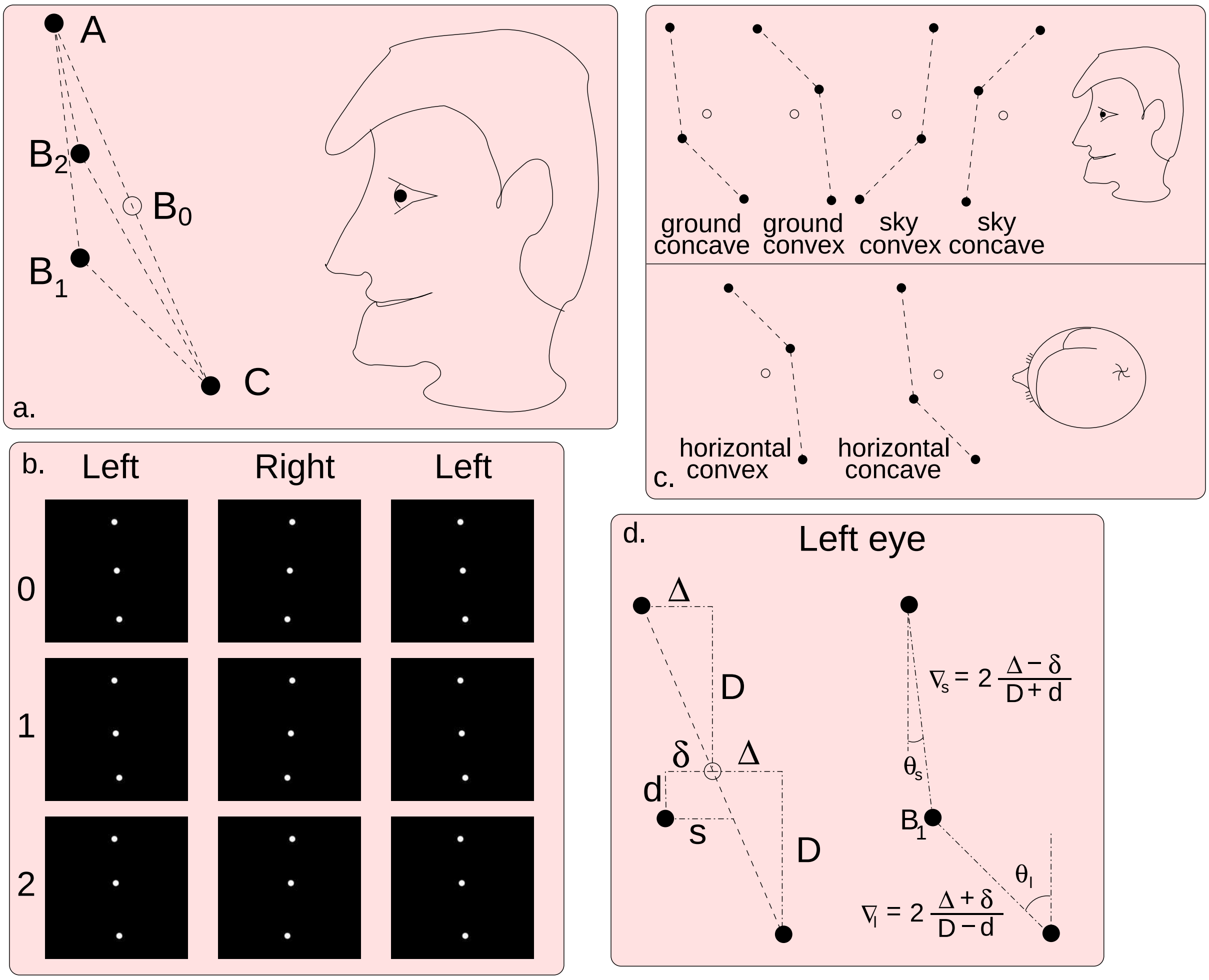
concave', and, for observers YP and AG, 'sky convex' ('sky concave' for GF).

Fig. 3. Results of experiment 2. (a) The stimulus was scaled down two-fold; (b) the upper reference dot was omitted. As in Fig. 2, the proportion of correct responses is shown for target position 1 and 2, as illustrated on the right. In (a) the concave stimulus illustrated on the right was used for YP and GF; a convex stimulus was used for AG.

Fig. 4. Results of experiment 3. The dots were arranged horizontally and slanted in depth (about a vertical axis). Again, the proportion of correct responses is plotted for target position 1 and 2. These positions are shown on the right in a top view of the stimulus.

Fig. 5. Data from all the conditions (Figs. 2, 3, and 4) are re-plotted here against the disparity gradient difference in each stimulus. Disparity gradient difference is the difference in disparity gradient on either side of the target dot T (see Fig. 1 and text). The curves are Weibull fits: $1-(1-\gamma)e^{-(x/\alpha)^{\beta}}$, where the false alarm rate γ was fixed at 0.5 for the 2AFC experiment. Separate fits are shown for the horizontal and two-dot conditions since for two ob-

servers performance was markedly different in these cases. The two-dot data was plotted on the same axes by calculating the disparity gradient difference that would have been present had the second reference dot been presented.



A
B_2
B_0
B_1
C
a.
ground concave
ground convex
sky convex
sky concave
horizontal convex
horizontal concave
c.
b.
Left
Right
Left
0
1
2
d.
Left eye
Δ
D
δ
Δ
d
s
D
$\nabla_s = 2\,\frac{\Delta - \delta}{D + d}$
θ_s
B_1
θ_l
$\nabla_l = 2\,\frac{\Delta + \delta}{D - d}$

Figure 1:

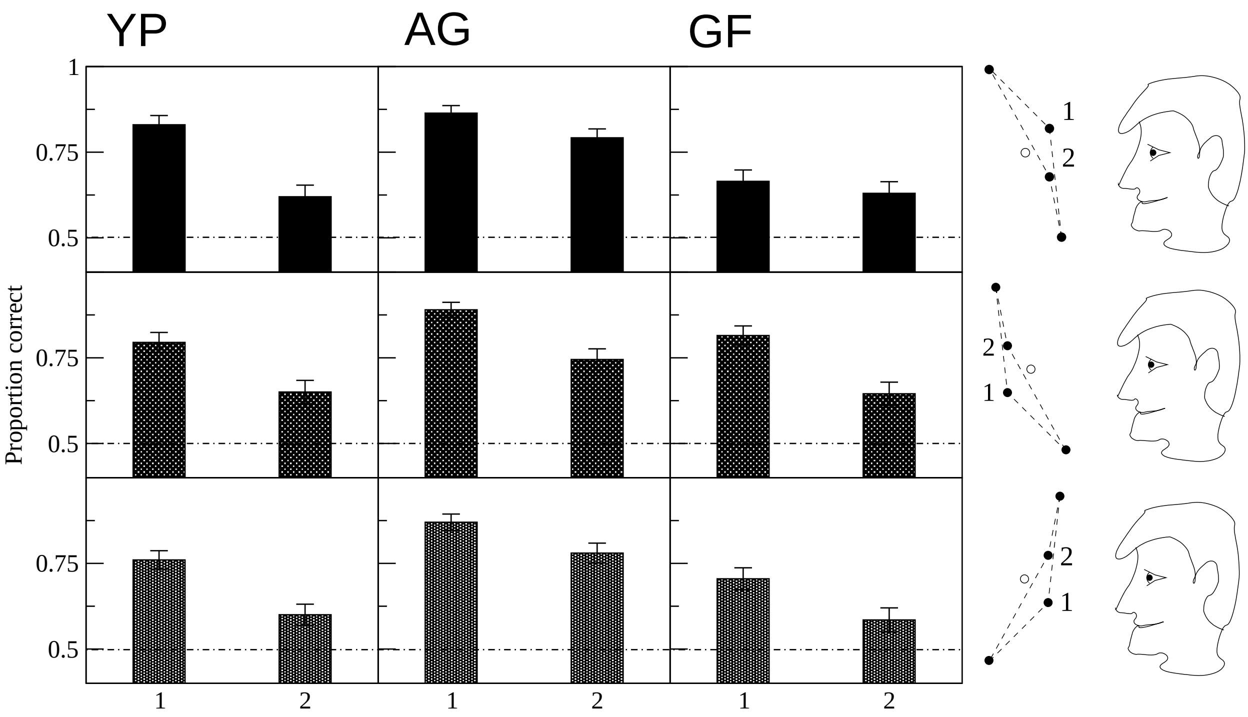


Figure 2:

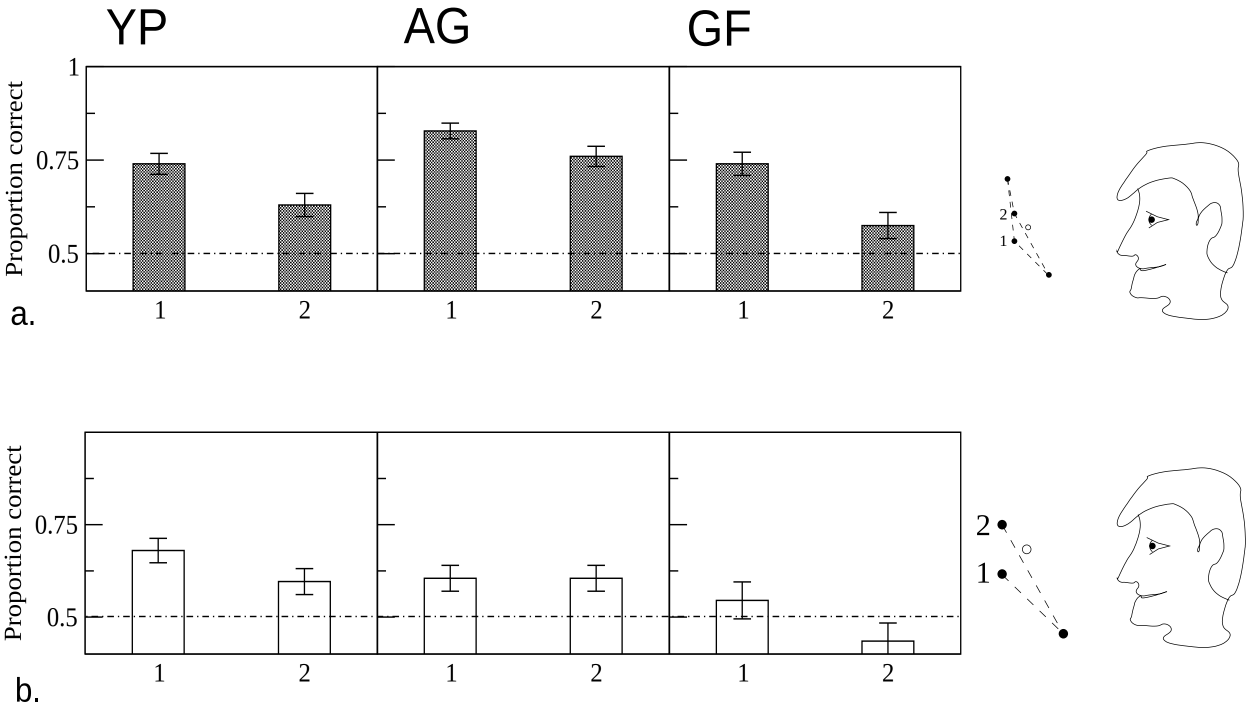


Figure 3:

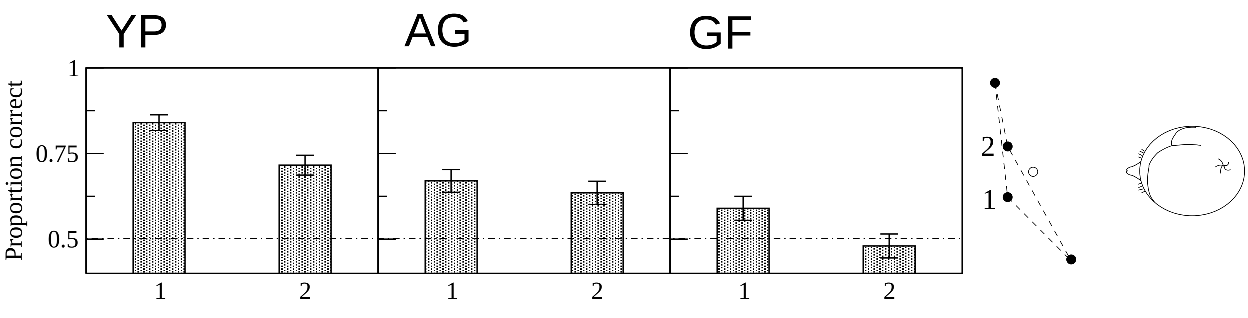
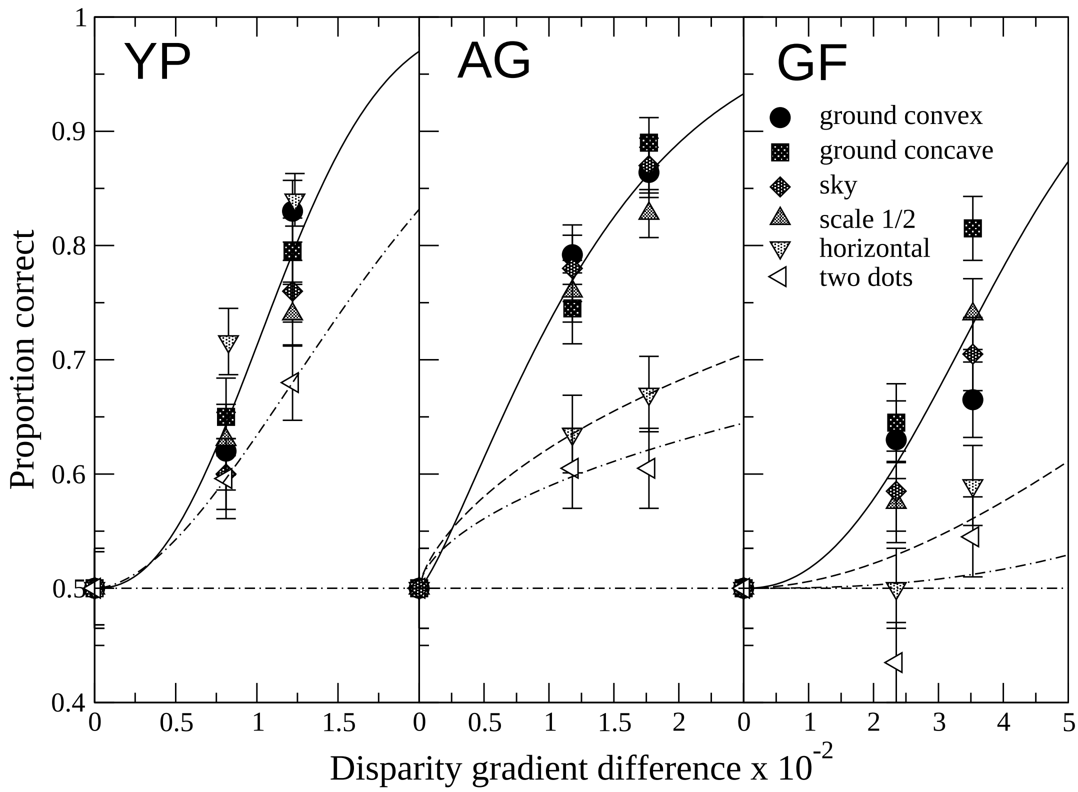


Figure 4:



YP
AG
GF
ground convex
ground concave
sky
scale 1/2
horizontal
two dots
Proportion correct
Disparity gradient difference x 10^-2
1
0.9
0.8
0.7
0.6
0.5
0.4
0
0.5
1
1.5
2
3
4
5

Figure 5: