Vision Research 49 (2009) 84–95

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Depth propagation and surface construction in 3-D vision

Mark A. Georgeson [a,*], Tim A. Yates [a], Andrew J. Schofield [b]

[a] School of Life & Health Sciences, Aston University, Birmingham B4 7ET, UK
[b] School of Psychology, Birmingham University, Birmingham B15 2TT, UK

ARTICLE INFO

Article history:
Received 29 February 2008
Received in revised form 26 September 2008

Keywords:
Depth perception
Depth propagation
Surface construction
Lateral interaction
Stereopsis

ABSTRACT

In stereo vision, regions with ambiguous or unspecified disparity can acquire perceived depth from unambiguous regions. This has been called stereo capture, depth interpolation or surface completion. We studied some striking induced depth effects suggesting that depth interpolation and surface completion are distinct stages of visual processing. An inducing texture (2-D Gaussian noise) had sinusoidal modulation of disparity, creating a smooth horizontal corrugation. The central region of this surface was replaced by various test patterns whose perceived corrugation was measured. When the test image was horizontal 1-D noise, shown to one eye or to both eyes without disparity, it appeared corrugated in much the same way as the disparity-modulated (DM) flanking regions. But when the test image was 2-D noise, or vertical 1-D noise, little or no depth was induced. This suggests that horizontal orientation was a key factor. For a horizontal sine-wave luminance grating, strong depth was induced, but for a square-wave grating, depth was induced only when its edges were aligned with the peaks and troughs of the DM flanking surface. These and related results suggest that disparity (or local depth) propagates along horizontal 1-D features, and then a 3-D surface is constructed from the depth samples acquired. The shape of the constructed surface can be different from the inducer, and so surface construction appears to operate on the results of a more local depth propagation process.

1. Introduction

In the perception of 3-D surfaces, image elements or regions with ambiguous or unspecified depth can acquire perceived depth from unambiguous regions. In various contexts, such co-operative perceptual interactions have been called stereo capture (Kham & Blake, 2000; Ramachandran, 1986; Wu, Zhou, Qi, & Wang, 1998) where disparity-matching is ambiguous; disparity interpolation (Mitchison & McKee, 1987b; Yang & Blake, 1995) where features are sparse or monocular; and surface integration (Yin, Kellman, & Shipley, 2000), surface interpolation (Wilcox, 1999; Wilcox & Duke, 2005) or surface completion (Rubin, 2001; Yin, Kellman, & Shipley, 1997) where parts of a surface are missing or occluded. The general theme that emerges from such studies is that the visual system constructs a representation of visible surfaces (Marr, 1982) often from sparsely sampled and/or ambiguous data, and in doing so must interpolate depth values and/or surface shape in regions where explicit information is absent.

Like the filling-in of brightness, colour and texture (Pessoa & De Weerd, 2003), the interpolation of depth values may involve propagation of information across visual space. We describe here some striking new perceptual effects, where depth corrugation (disparity modulation) in flanking regions can induce vivid perceived depth in a central test region (Fig. 1B). We used a monocular gauge figure to measure the perceived shape of a variety of test images, and to quantify the magnitude of the induction effect. The nature and specificity of these induction effects lead us to propose that depth propagation and surface construction are different, perhaps successive, stages of visual processing.

2. Methods

2.1. Image generation

Images of a random, Gaussian white noise texture (512 × 512 pixels) were created in *Matlab* 5.2 on a Macintosh G4 computer, and displayed on a Clinton fast-phosphor monitor via a CRS Bits++ interface in true 14-bit greyscale mode. *PsychToolbox* (Brainard, 1997) and in-house software were used to calibrate the display system and run the experiments. Stereo viewing was achieved by using frame-interleaving FE1 goggles (CRS Ltd.) to present separate images to the two eyes. The high frame rate (150 Hz; 75 Hz per eye) ensured that the alternating display appeared as a steady 3-D image with no visible flicker. Luminance measurements and calculations showed that physical crosstalk between the two eyes' views was very low and would be undetectable. We calculated that for sine-wave gratings up to 80% contrast less than 1% of that contrast would effectively 'leak' through to the other eye.

* Corresponding author.
E-mail address: m.a.georgeson@aston.ac.uk (M.A. Georgeson).

0042-6989/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2008.09.030

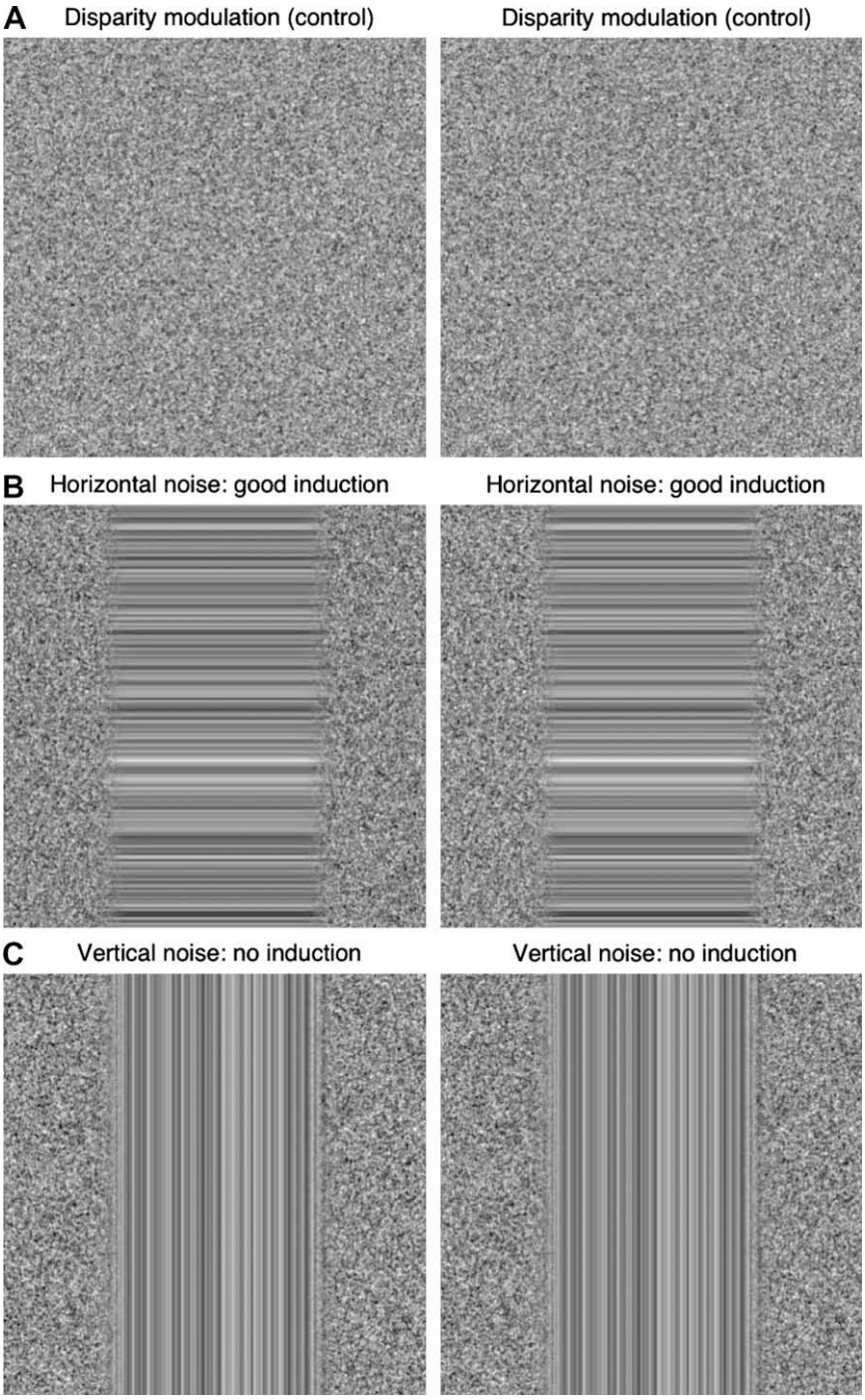

Fig. 1. (A) The slightly blurred Gaussian noise texture used in the experiments. Fusion of the left and right images should reveal a horizontal corrugated surface, produced by sinusoidal modulation of disparity. (B) Configuration of the central test region, flanked by the corrugated (disparity-modulated) inducing surface. Test texture is horizontal 1-D noise, and is the same in both eyes. Depth induction from the flanks was strong. (C) As B, but the test texture is vertical. Depth induction was absent. For more example images, see Supplementary Figures S1-S4.

To create the appearance of horizontal corrugations (Fig. 1A), with sinusoidal modulations of disparity, each row of the texture had to be shifted to the left or right by a small (often sub-pixel) distance that was a sinusoidal function of the row's vertical position (y) in the image. These small shifts were achieved by blurring each row separately with a Gaussian kernel whose space constant (σ) was 1 pixel (1.17 min arc), and whose peak was displaced by $\pm\delta/2$ for the left and right eyes respectively, where δ is the desired disparity of the row. Thus the convolution (blur) kernel r was defined by:

$$r(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x \pm \delta/2)^2}{2\sigma^2}\right) \quad (1)$$

where x is sampled in 1-pixel steps. To ensure that the blurring was isotropic (circularly symmetric), a similar blurring was then imposed on each column of pixels, but with no spatial offset.

The adequacy of sub-pixel resolution is influenced both by the effects of quantization (number of grey levels) in the graphics display system, and by the effects of gamma-correction. To evaluate the rendering of disparity one would, ideally, record the CRT

output light distributions, and then compare the Fourier phase spectra for corresponding rows of images seen by the left and right eyes. For a given disparity, the phase difference should increase linearly with spatial frequency, with a predictable slope. In practice, it is very difficult to acquire real luminance data with the spatial resolution required, so we relied instead on a simulation. In gamma-correcting a visual display, the experimenter sets up a model of the CRT's voltage-to-luminance transformation, estimates parameters of the model from luminance measurements, and then inverts the model equation to construct the look-up table that linearizes the system overall. Thus, starting with the (double-precision) image array in *Matlab*, quantizing it to the known number of grey levels, passing it through the (software) look-up table and then through the CRT model equation we arrived at the (simulated) luminance output of the CRT. We then used *Matlab*'s *FFT* function to examine both the phase difference spectrum (as above) and the amplitude difference spectrum, and we compared them with their ideal counterparts. We concluded from many such analyses, simulating graphics systems with different greyscale resolutions, that with our 14-bit greyscale the precision of disparity rendering was very high, and was sufficient for disparities as low as 0.1–0.2 s.

At the viewing distance of 114 cm, the image subtended 10×10 deg square, set in a uniform background (16 deg wide - 12 deg high) of the same mean luminance (26 cd/m^2, as measured through the frame-interleaving stereo goggles, which reduce time-average luminance in each eye to 12.5% of its original value).

The display consisted of two distinct regions (Fig. 1B and C): the flanking regions that contained a disparity-modulated (DM) random texture that appeared horizontally corrugated when viewed stereoscopically, and a central region that could contain a variety of different binocular or monocular test images. Binocular test images were always the same in both eyes (zero disparity), while monocular test images were created by setting one eye's test image to zero contrast (mean luminance). The borders of these regions were defined by a smooth, 1 deg wide, rise or fall in contrast (Fig. 2B). Note that the DM flanks (with disparity) and the test regions (with no disparity) were generated and windowed separately, then added to form a single image for each eye (Fig. 1). Since the window functions (Fig. 2) were the same for both eyes, this ensured that no disparity was introduced into the test images.

The RMS contrast of the flanking texture (standard deviation of luminance divided by mean luminance) was 0.2, and the spatial frequency f of disparity modulation was 0.15 or 0.3 c/deg. These spatial frequencies are close to the peak of disparity modulation sensitivity (Bradshaw & Rogers, 1999; Schumer & Julesz, 1984; Tyler, 1974). We quantify the disparity modulation by its sinusoidal amplitude a, ie., half the difference between the peak and trough disparities. Thus the flanking surface was defined by its sinusoidal disparity profile:

$$\delta(y) = d + a.\cos(2\pi f y - \phi) \quad (2)$$

where δ, a, f and y are defined above, ϕ is the phase of the modulation relative to the screen centre ($y = 0$), and d is the mean disparity (0, except for experiment 4). For these experiments, $a = 3$ min arc, and $\phi = 0, 90, 180$ and 270°, chosen at random from trial to trial. Different random noise samples were used from trial to trial.

2.2. Slant matching, using a monocular gauge figure

To quantify perceived surface shape, we superimposed a gauge figure (Fig. 2A) at the vertical mid-line of the test region, in one eye only. The smoothed, anti-aliased lines of the gauge figure were dark, with contrast 0.5, and its background was transparent. The gauge figure was 42 pixels (49 min arc) wide, and was presented at one of eight positions that evenly sampled one period of the flanking modulation. Using the computer mouse, the observer could quickly adjust the apparent slant of the gauge figure, in 5 deg steps, to match the perceived slant of the underlying test surface. The slanted disc should appear to lie on the surface, while the stick is normal to the surface. There was no time limit, but with practice each setting could be made in about 2–3 s. All eight settings, in random order of gauge position, were completed before the computer chose another test condition. All conditions were repeated eight times per subject in random order, and across trials the gauge figure was presented equally often to the left and right eyes. In all experiments the test image was presented either to one eye or to both eyes. When the test image was monocular (left or right eye), the gauge figure was superimposed on the test image, while the other eye saw a uniform (mean luminance) test image. The DM flanks were of course always binocular.

2.2.1. Experiment 1

This first experiment sought to establish that depth induction could be quantified with the gauge figure method (above), and to determine whether depth induction depended on the form of the test texture (1-D *vs* 2-D) and its orientation (horizontal *vs* vertical). The test image was either 2-D static Gaussian noise, with RMS contrast 0.2, as described above, or it was the 1-D analogue formed by forcing all the rows (or columns) to contain the same noise sample (Fig. 1B and C). Spatial frequency of the flanking corrugations was 0.15 c/deg. Test and flanking regions were smoothly blended, using the 'medium' window type (Fig. 2B).

2.2.2. Experiment 2

Here we studied how depth induction was influenced by the type of transition and overlap between the test and inducing regions (Fig. 2B). The width of the test region (at half-height on the contrast envelope) was 4.5, 5, 5.5, and 10 deg for the four window types (narrow, medium, wide and full), while the width of the corresponding 'hole' between the flanks was 5.5, 5, 4.5 and 5 deg. In the narrow condition there was thus a 0.5 deg gap (at half-height) between test and flank boundaries. See Supplementary Figures S5-S8 for illustrations of all four window conditions. Other conditions were similar to experiment 1, but only the horizontal 1-D noise test images (Fig. 1B) were used.

2.2.3. Experiment 3

This experiment examined further the role of horizontal contours in depth induction, using periodic gratings rather than noise. The test image was a horizontal luminance grating, of contrast 0.4, with no noise texture. The grating (Fig. 3) was either a sine-wave, a square-wave, or thin dark lines (3.5 min arc wide). The spatial frequency of these gratings matched that of the DM flanks (0.15 or 0.3 c/deg), except that in the thin-line condition there were two lines per cycle of modulation. The phase of the test grating, relative to the DM flanks, was offset by 0, 90, 180 or 270°, chosen at random from trial to trial. Other conditions were similar to experiment 1.

2.2.4. Experiment 4

This was similar to experiment 1 in all respects, but was designed to explore the interaction of vertical (V) and horizontal (H) noise test images. The data were collected in the same sessions as experiment 1. The test image was a horizontal, 1-D noise image of contrast 0.2, added to a vertical 1-D noise image of contrast 0.02 (see Fig. 7A). To form the composite image – a noise plaid – the contrast values of these two images were linearly summed, with no change in mean luminance. The observer made slant judgements as usual, but did so for the H and V components separately. In pilot experiments, we explored a range of contrast levels for the vertical noise component and chose a low, but easily visible contrast (0.02) for the main experiment because it seemed to produce effects similar to the higher contrasts, while maximizing the

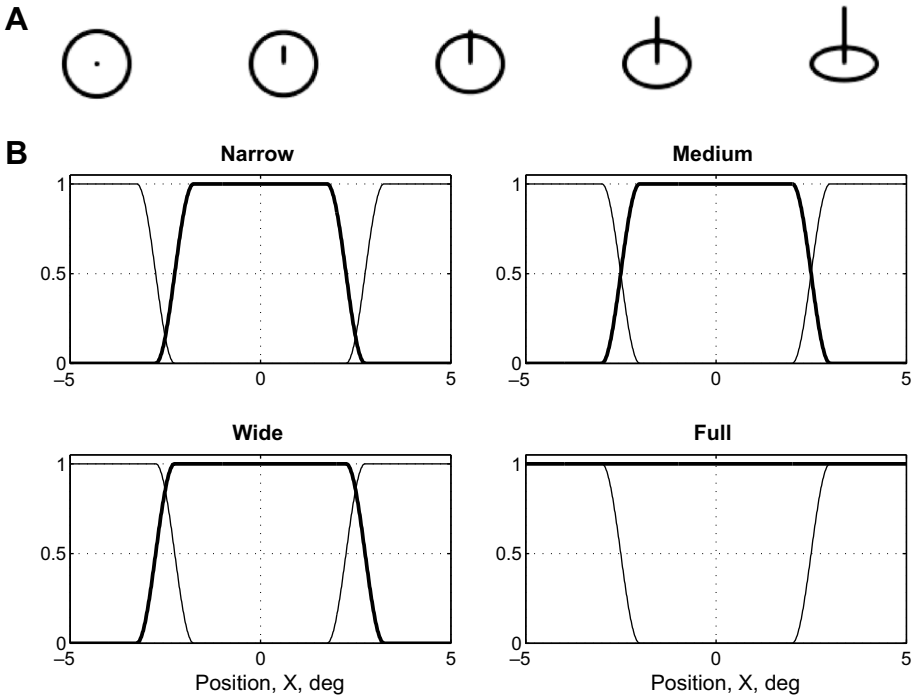

Fig. 2. (A) Examples of the monocular gauge figure used to measure perceived surface slant. We defined positive slants to be where the stick points upwards (as shown) and negative slants where the stick points down. (B) Contrast envelopes that defined the boundaries of the central test region (thick curve) and the inducing flanks (thin curve). The test region width was 'medium' in experiments 1 and 4, and 'wide' in experiment 3. All four types were compared in experiment 2; for example images, see Supplementary Figures S5-S8.

perceived transparency and separability of the H and V test components. The H and V test components were monocular, or binocular with zero disparity, as in experiment 1. Based on pilot observations we introduced a new variation into the DM flanks. The mean disparity d in Eq. (2) was set to 3, 0 or −3 min arc. This displaced the flanking corrugations backward ($d = 3$) or forward ($d = -3$) in depth, and so meant that the plane of the V test noise (with zero disparity) appeared at the same depth as the peaks, zero-crossings, or troughs of the DM flanks. We shall call these the 'front', 'middle' and 'back' test planes respectively. With cross-eyed fusion, Fig. 7A illustrates V noise in the 'front' plane: the V noise appears to lie in the same plane as the peaks of the flanking surface.

2.3. Sign conventions

Positive values of disparity are uncrossed, i.e., 'far'. For the DM cosine grating, phase 0 therefore places a trough ('far') at the screen centre. The y-axis was positive downwards (a *Matlab* convention for image arrays). Positive phase values in Eq. (2) shift the grating downwards. Phase 180 places a ridge ('near') at the screen centre. Gauge figure positions are also expressed as phase values relative to the DM grating: 0 would align the gauge figure with the far trough; positive values move it down; 90° aligns it with disparity zero-crossings; 180° aligns it with the near peak. For zero slant the gauge stick is horizontal, pointing at the observer, hence reduced to a dot (left of Fig. 2A). Positive values of slant make the gauge stick point upwards (Fig. 2A).

Four observers, including two of the authors, were tested. All had good stereo vision, with spectacle correction if necessary.

3. Results

Slant measures on the test surface were obtained at eight test points, sampling one period of the inducing corrugation. When the test surface appeared corrugated, these slant measures varied periodically with vertical position (Fig. 4A). Note that peaks and troughs in depth have slants of zero, and so yield settings near zero in the slant profile (Fig. 4A). We computed the Fourier transform (FFT) of the eight slants and took the amplitude of the fundamental component as the magnitude of induced depth. To allow for possible differences in response scaling across observers, this amplitude was then expressed as a proportion of the amplitude obtained in the control condition (Fig. 4A, dashed lines), where the test image had the same disparity-modulated texture as the flanks (Fig. 1A). Control conditions of this kind were routinely included in all four experiments. If observers saw the same depth variation in the induction and control conditions, they would make similar settings in both cases, leading to a relative amplitude of 1.0 in Fig. 4C and D.

3.1. Experiment 1

In general, when depth was induced in the test region, it appeared to be in-phase with the flanking surface. This is illustrated in Fig. 4A, where the induced slant profile for binocular, horizontal noise (filled circles) was similar to that for a truly stereoscopic (DM) test surface (dashed lines), but slightly reduced in amplitude. Thus the induced depth is some form of perceptual completion and not a contrast phenomenon. This induction is the opposite of 'grating induction' in the luminance domain where an *out-of-phase* illusory grating is seen in the blank (mean luminance) gap between two inducing gratings (Blakeslee & McCourt, 1997; McCourt, 1982).

Fig. 4 shows how the perceived depth induced by the flanking corrugation depended on the type of test texture used. First consider binocular test images. When the test image was horizontal noise (H), depth induction was strong (Fig. 4A and C) – about 70–75% of the control value. But with vertical noise (V) or 2-D noise induction was absent (Fig. 4A and C). An obvious interpretation is

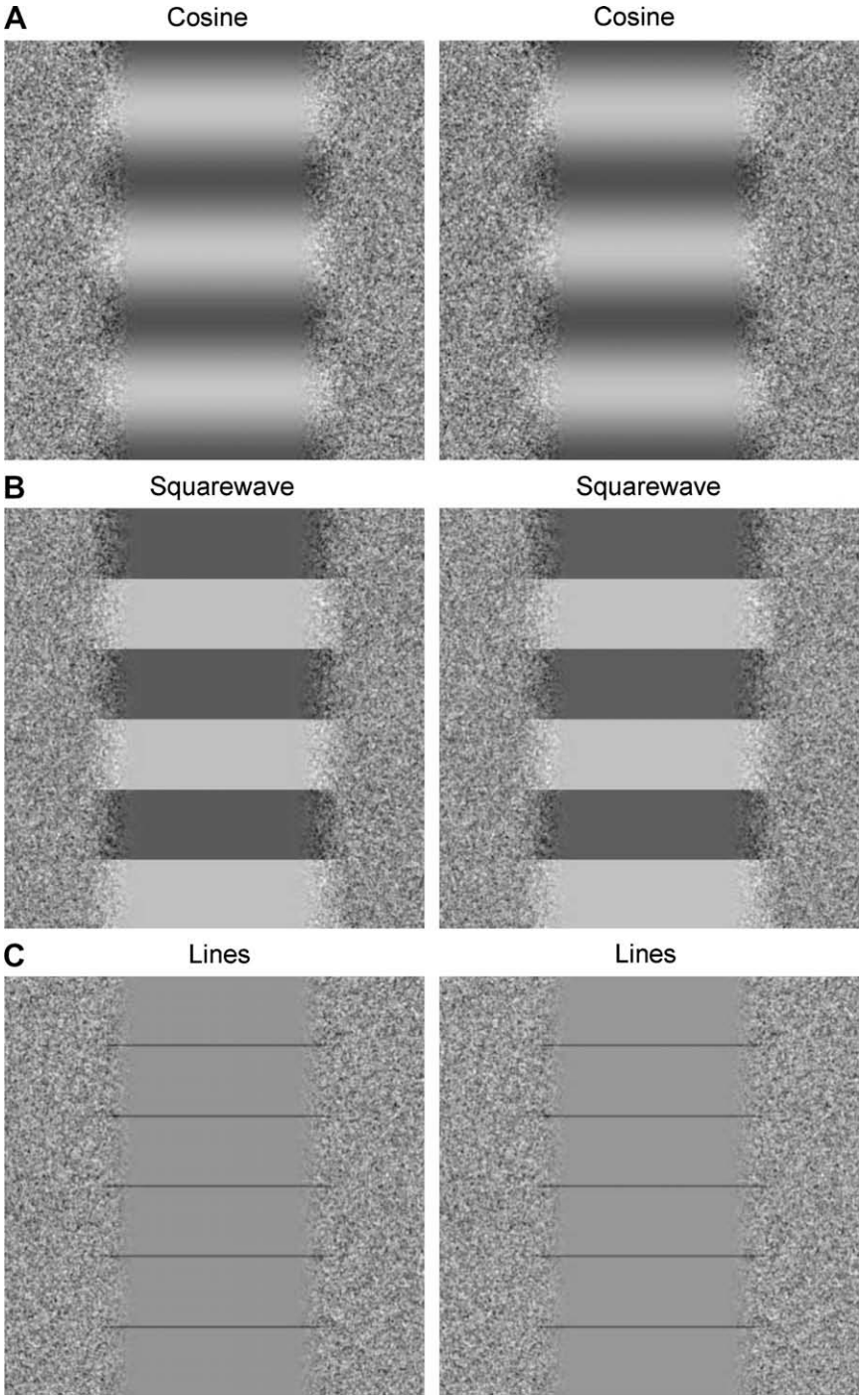

Fig. 3. Experiment 3 shows the three types of test image and their spatial relationship with the disparity-modulated (DM) flanks. The experiment used four different phase offsets between the test image and the DM flanks.

that induction is possible for the horizontal noise because its horizontal disparity is indeterminate, while the vertical and 2-D noise have a specific (zero) disparity that is readily detectable as a flat surface. This may be true, but is unlikely to be a complete answer because all three test types lacked disparity when presented monocularly (Fig. 4B and D) but again only the horizontal noise showed clear induction while the other two showed little or none. Thus it appears that horizontal orientation and the absence of measurable disparity information are both important requirements. A lack of specific disparity may be a necessary condition for induction, but it is clearly not sufficient.

Monocular pattern cues also play a role in determining the likelihood of different surface interpretations. Surface contours that run across (rather than along) the corrugations of a surface project as wavy lines in the image of the surface and play a key role in the perception of surface shape (Knill, 1992; Li & Zaidi, 2000, 2001; Stevens, 1981). Thus when the test image contains straight vertical lines, their lack of waviness makes it very unlikely that the surface is corrugated horizontally. This cue to flatness may well contribute to the lack of depth induction for vertical test lines. In contrast, when a low-pass filtered horizontal (H) noise or 2-D noise surface is corrugated horizontally, the projected texture is also visibly

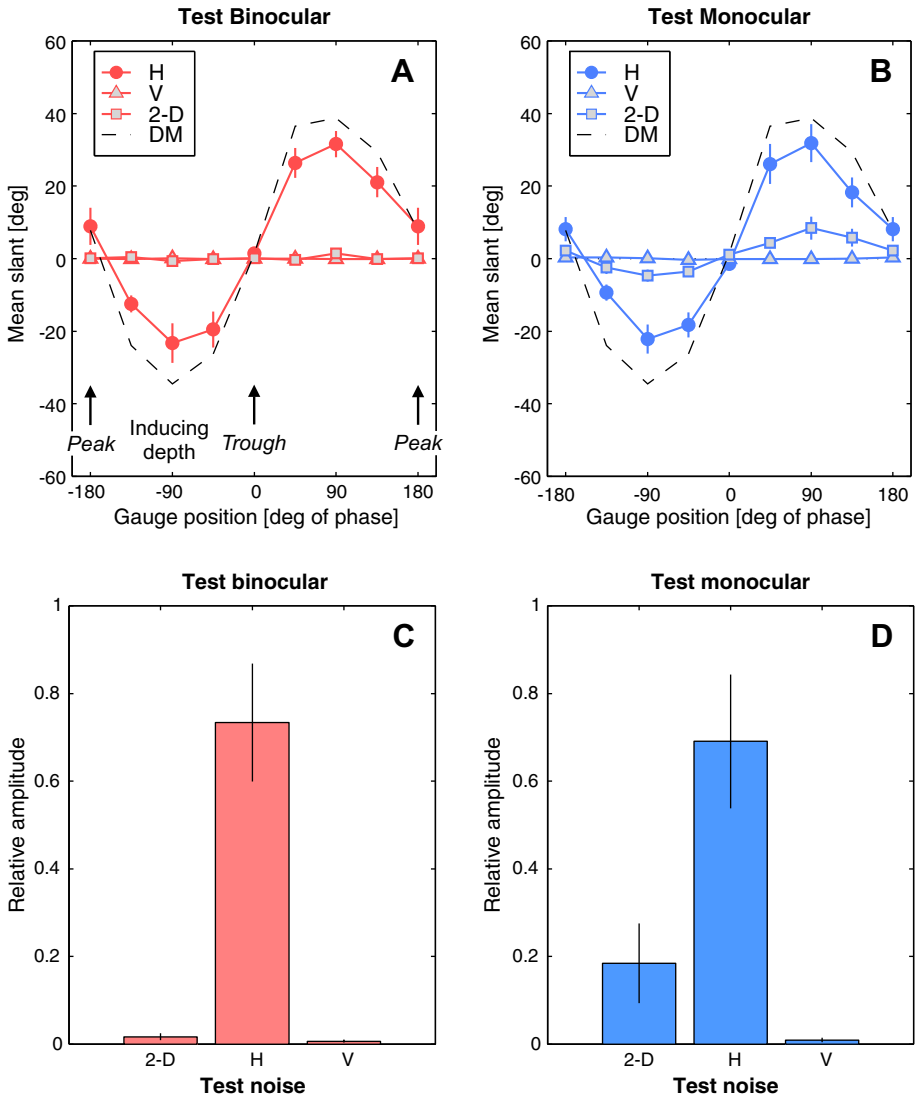

Fig. 4. Experiment 1, perceived depth induced by the flanking corrugations, for 2-D and 1-D (horizontal, H or vertical, V) noise test images that were binocular (left panels) or monocular (right panels). (A) Slant profile (mean of $n = 4$ observers, ±1 s.e.) for binocular test images. Dashed curve shows control data for a stereoscopic (DM) test surface. Abscissa plots gauge position relative to the inducing DM flanks; see *Sign conventions*. Depth peaks and troughs are points of zero slant (arrowed). (B) As A, but for monocular test noise. (C and D) Magnitude of induced depth was summarized by the amplitude of the slant profile. Bar height (±1 s.e.) shows the mean induced depth relative to the control measures obtained with true stereo DM texture. Only the horizontal noise exhibited strong induced depth.

modulated in both cases, but such modulations are much less effective as a veridical depth cue (Li & Zaidi, 2000). Nevertheless, if the absence of such modulations – the uniformity of texture – acted as a cue to flatness, we might expect depth induction to fail in both cases, but in fact it failed only for the 2-D noise and not the H noise. This therefore suggests some special status for horizontal test contours, where induction was strong both binocularly and monocularly.

3.2. Experiment 2

For three of the four window types (medium, wide and full; Fig. 2B), the borders of the test region blended smoothly with, or overlapped, the flanking regions and in these three cases depth induction was strong (Fig. 5), though perhaps a little less with the smallest overlap (medium window) that was also used in experiment 1. When a small (0.5 deg) gap was introduced between the flanks and the test region (using the narrow window), depth induction was almost completely blocked (Fig. 5). This suggests that depth values somehow 'flow' from the flanks into the test region along horizontal contours, and that this involves a process of local propagation that can be interrupted by a blank gap. The results were very similar with monocular and binocular test images.

3.3. Experiment 3

To study further the role of horizontal contours in depth induction, experiment 3 used simple, periodic, horizontal test contours (Fig. 3) rather than a dense texture. We found that the shape of the induced surface depended on the type of test grating used. Fig. 6A shows the average slant settings made on the cosine test grating, as a function of the gauge figure's vertical position relative to the inducing grating. As in experiment 1 (Fig. 4A), note the similarity between the experimental data (symbols) and the control data (black, dashed curve; no symbols). Thus for the cosine test grating, induced depth closely followed the perceived depth of the inducing grating itself, though at a somewhat lower amplitude (about 75% of the control amplitude).

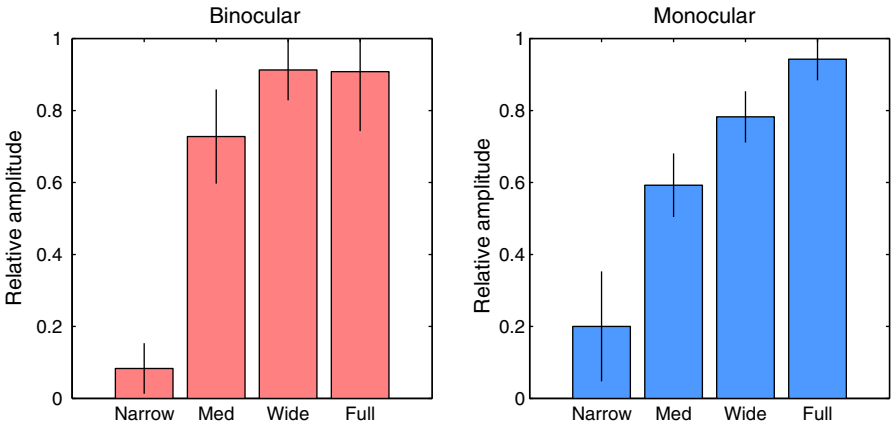

Fig. 5. Experiment 2, perceived depth induced by the flanking corrugations (0.15 c/deg), for horizontal 1-D noise test images that were binocular (left) or monocular (right). Means of four observers ± 1 s.e. Four test/flank configurations (Fig. 2B) were compared. A small (0.5 deg) gap between test and flanks (the 'narrow' window) blocked depth induction.

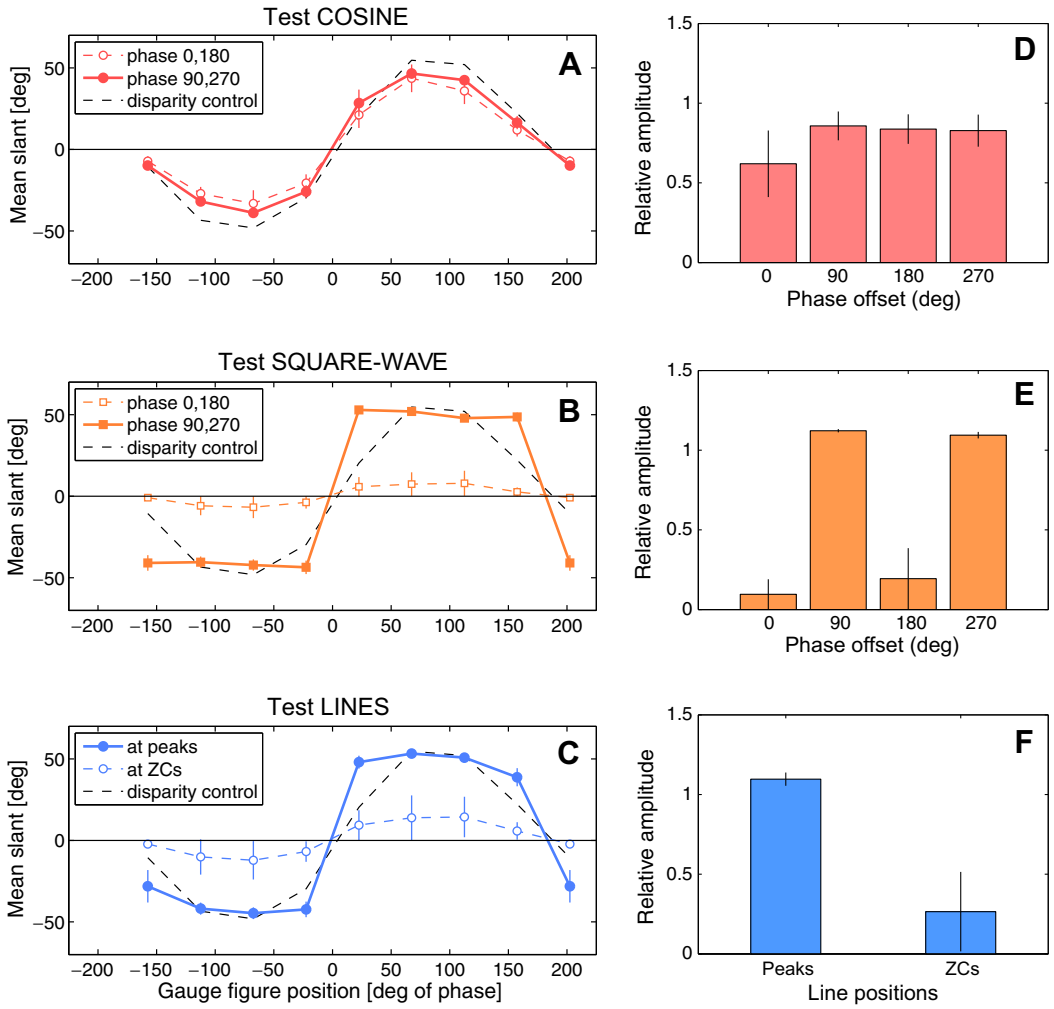

Fig. 6. Experiment 3, perceived depth induced by the flanking corrugations (0.3 c/deg), for binocular, horizontal grating test images. (A, B and C) Perceived slant profiles for the three types of test grating. Results for complementary test phases (0, 180) or (90, 270) were always very similar, and their averages are shown here. (D and E) Magnitude of depth induction (relative to control) for the four test phases. (F) As D and E, but for the line test images. Test lines were aligned either with the DM peaks and troughs, or with the DM zero-crossings (ZCs). Note: the derivative (gradient) of a triangle-wave is a square-wave. Hence the nearly square-wave slant profiles in B and C (solid symbols) imply a perceived triangle-wave surface, as reported by all observers.

Interestingly, for the cosine test pattern, the induced depth was not significantly influenced by its spatial phase. Four phases (relative to the DM grating) were tested, but Fig. 6A and D clearly show that induced depth was much the same – in both phase and amplitude – in all four cases. The induced depth thus depended almost entirely on the flanking pattern of disparity. We had anticipated

that induced depth might appear reduced at test phase 90, where the combination of disparity and luminance pattern (considered as shading) is consistent with lighting from below, and/or enhanced at phase 270 which is consistent with lighting from above. Fig. 6 D shows that there was no such difference.

A more complex, phase-specific pattern of results emerged from the square-wave test condition (Fig. 3B). When the square-wave test edges were aligned with the zero-crossings of disparity (at test phase 0 or 180) there was little or no induced depth (Fig. 6E; and open symbols, Fig. 6B). But when the test edges lined up with the inducing peaks and troughs (at test phase 90 or 270), the induced depth was strong, but appeared to be a *triangle-wave* depth profile, not a cosinusoidal one. A triangle-wave has alternating half-cycles of constant, opposite, slant: hence the measured slant profile (Fig. 6B) was a square-wave. Perceived slant of the induced triangle-wave surface was about equal to the peak control slant (Fig. 6B) and was the same for both test phases 90 and 270 (Fig. 6E).

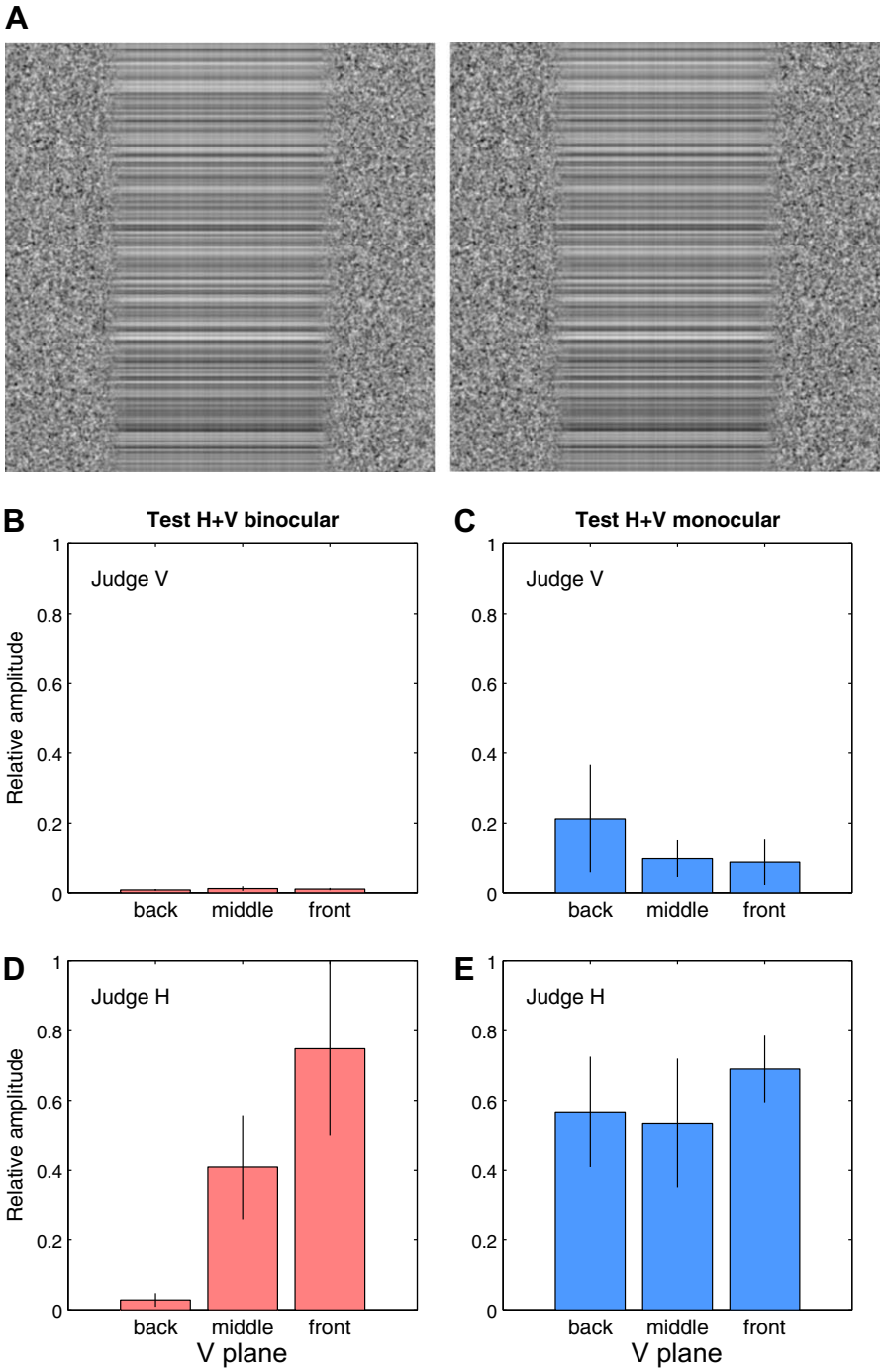

Fig. 7. Experiment 4, (A) stereoscopic DM flanks with spatial frequency 0.15 c/deg. Test images contained H and V noise components whose surface shapes were judged separately. The V component had very low contrast. In this example, with cross-eyed fusion, the V component is in the 'front' position. Middle row: judge the vertical component. Bottom row: judge the horizontal component. (B and D) binocular test images, as shown in A; the V noise plane had the same disparity as the troughs, zero-crossings, or peaks of the DM flanks ('back', 'middle' or 'front', respectively). (C and E) Monocular test images (with binocular DM flanks); the three monocular V plane conditions differed in name only.

It might appear so far that for periodic test gratings the induced slant profiles follow the test luminance profile, at least in shape if not in phase. The 3rd test condition (thin lines) shows that this is not generally correct. The results were similar to the square-wave test condition, even though their luminance profiles are obviously very different. A (nearly) square-wave slant profile was produced (Fig. 6C) when the test lines lay at the peaks and troughs of inducing disparity. Induced depth for thin test lines (two lines per inducing period) depended on the test *line* position in much the same way as for the *edge* position in the square-wave test condition.

We have described (in Fig. 6) the results at a spatial frequency of 0.3 c/deg. Results at 0.15 c/deg were similar in all respects (not shown), except that the mean depth induction was a little less than at 0.3 c/deg, largely due to the weak induction seen by one of the four observers at the lower SF.

We also noted informally that as the number (spatial frequency) of thin, dark, test lines was increased, so the induced depth shifted from a triangle-wave to a cosinusoidal appearance that was similar to that seen in the horizontal noise condition. An important factor, then, could be the position (and number) of key *features* in the test image, rather than the test luminance profile *per se*.

3.4. Experiment 4

Experiment 1 revealed that depth induction was strong for horizontal test noise, but absent for vertical noise. What would happen if the test image was a composite of horizontal and vertical components – a noise plaid (Fig. 7A)? We guessed that the vertical noise, which never suffered induction, might capture the horizontal contours and so eliminate induction by the flanks altogether. The actual outcome was much more interesting and elaborate.

For monocular test images the result (Fig. 7C and E) was simple, and similar to experiment 1. For H noise the induction effect was fairly strong (about 60–70%), but for V noise it was weak or absent. Not surprisingly, there was no difference between results for the back, middle and front V plane conditions monocularly, because they then differed in name only.

With binocular test images, however, the disparity of the V plane turned out to be crucial. The V noise itself always looked completely flat (Fig. 7B), but its disparity controlled depth induction in the H component. When the V plane was 'in front', the H noise appeared strongly corrugated behind it, but when the V plane was 'back' (at the same disparity as the DM troughs), the H noise appeared flat (Fig. 7D). When the V plane was in the middle depth position, induction magnitude was approximately halved, compared with the 'front' condition (Fig. 7D).

From subjective reports, we refer to this as the *shower curtain effect*: the V noise looked like a flat, transparent curtain, through which the H noise was seen (Fig. 7A). It seemed that depth induction could occur *behind* the 'curtain', but not in front of it. When the 'curtain' was in front, the H noise acquired depth corrugation from the DM flanks just as it did in experiment 1, but when the 'curtain' was at the back, the H noise remained behind the curtain, with no depth induction. In the middle position, troughs of induced depth could be seen behind the curtain, but with no induced peaks in front of it. It was as though depth induction could not penetrate the curtain. Such truncation (half-wave rectification) would predict a halving of depth magnitude, as observed. We now show that

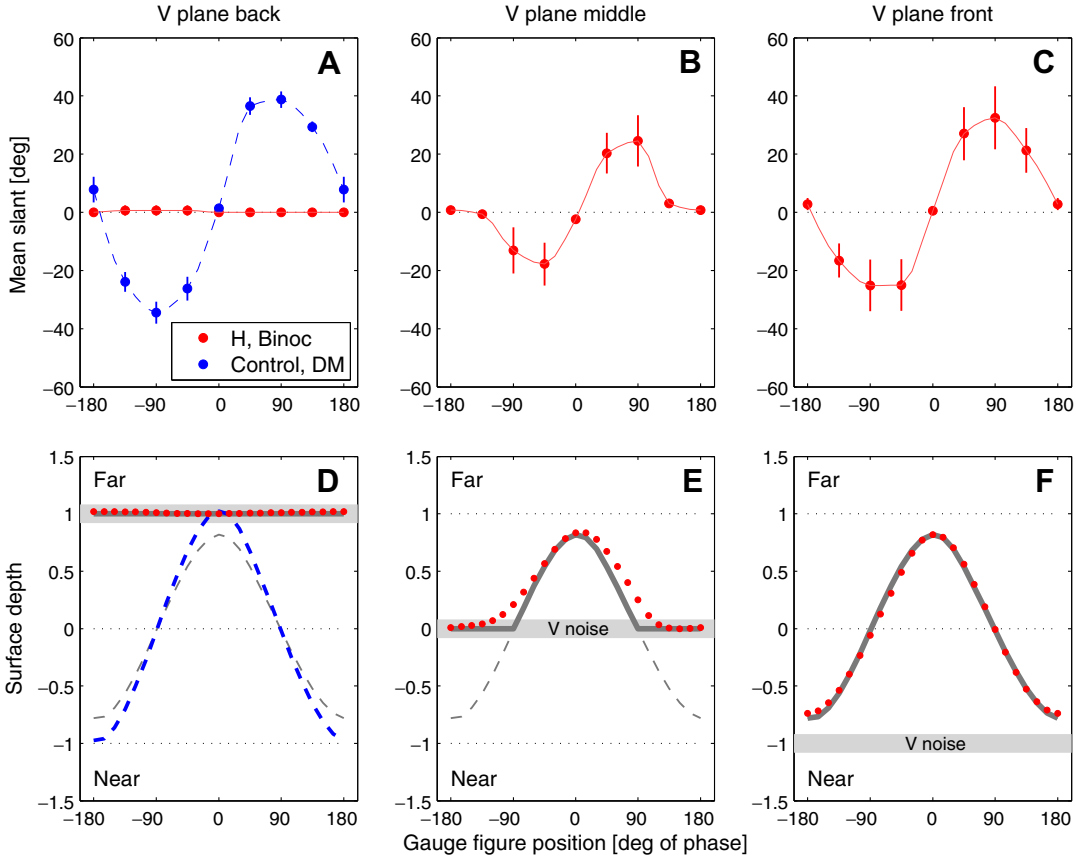

Fig. 8. Experiment 4, binocular test conditions only. (A, B and C) Induced slant profiles for the binocular H noise when the V plane was (A) 'back', (B) 'middle' or (C) 'front'. Dashed curve in A is for the control (DM) test condition. (D, E and F) Smooth curves fitted to the slant data (A, B and C) were integrated (see *Appendix*) to infer the perceived surface shapes (dots). Control surface (thick dashed curve, D) was scaled to a mean of 0, with amplitude ±1. Thin dashed curves show the induced surface expected as 0.8 × (control surface). Horizontal grey bar represents the depth of the V noise surface. Solid grey curves show how induced depth that would otherwise lie nearer than the V noise appeared to be captured by it (D and E), while induced depth beyond the V noise (E and F) was unaffected by it.

these subjective reports are supported, in detail, by further analysis of the experimental data.

Fig. 8A, B and C show the average slant profiles for the three cases where the V noise was at the back, middle and front, respectively. Fig. 8A also shows the slant profile for the DM control surface. We integrated these slant data to infer the perceived surface shapes, plotted as dots in Fig. 8D, E and F, along with the perceived control surface (thick dashed curve in Fig. 8D). See Appendix A for computational details. It is very clear that the disparity of the V noise controlled the amplitude and shape of the induced surface. These surface plots strongly suggest that *depth induction occurs behind the V plane, but not in front of it*. When the V noise was in front, the induced surface (Fig. 8F) was very close to 0.8 × (control surface), much as it was without V noise (experiment 1). But when the V noise was in the middle position (Fig. 8E) the perceived surface (dots) was close to a half-wave rectified version of the induced surface, as shown by the smooth grey curve. When the V noise was at the back (Fig. 8D), no induced depth was seen, and from subjective reports we conclude that both H and V were seen as a flat plane in the back position. In all three cases then, the following simple rule (solid grey curves in Fig. 8) predicts perception of the Horizontal noise surface:

$$\text{induced\_surface} = \max[k \cdot (\text{control\_surface}), \text{V plane}],$$

where $k = 0.8$ for this experiment, and V plane = 1, 0, −1 for the three V positions. In words, the rule proposes that depth values are passed horizontally from the flanks to the H test pattern, with some slight attenuation, but if the induced depth would be *nearer* than the V noise plane then that induced depth is vetoed and captured by the V plane. Evidently the flat vertical 'curtain' of noise with its explicit disparity does not always veto depth induction; it does so only when the inducer is closer than the 'curtain'. The source of this front/back asymmetry remains unclear to us.

4. Discussion

In previous work, depth interpolation has been described as a process in which "a depth has been assigned to locations where disparity is ambiguous, and ... this perceived depth is based, in part, on the depth values at locations where disparity is well-defined" (Wurger & Landy, 1989, p.39). Similarly, Collett (1985) (p. 55) suggests that "the binocular surfaces give a framework across which the monocular surface is stretched or a region from which depth information can spread to areas where depth is not defined explicitly." Mitchison and McKee (1987a, 1987b) provided experimental evidence that depth interpolation from unambiguous disparity matches is used to resolve the ambiguity of disparity matching in periodic ('wallpaper') stereograms. At brief durations (150 ms) and small dot spacings (<5′ arc), this interpolated depth is *perceived*, rather than any of the possible discrete disparity matches that the dots could make. At long durations and with larger dot spacings (eg 7′ arc) the interpolated depth plane is not perceived, but it *guides* the discrete matching of the ambiguous test dots. We can infer this, because the perceived depth of the test dots corresponds to disparity matches that are closest to the interpolated plane. Later experiments showed that as well as guiding stereo matching, the interpolated plane also determines psychophysical sensitivity to spatial displacement: stereoacuity for test dots depended on their disparity relative to the interpolated plane (Glennerster & McKee, 2004; Glennerster, McKee, & Birch, 2002).

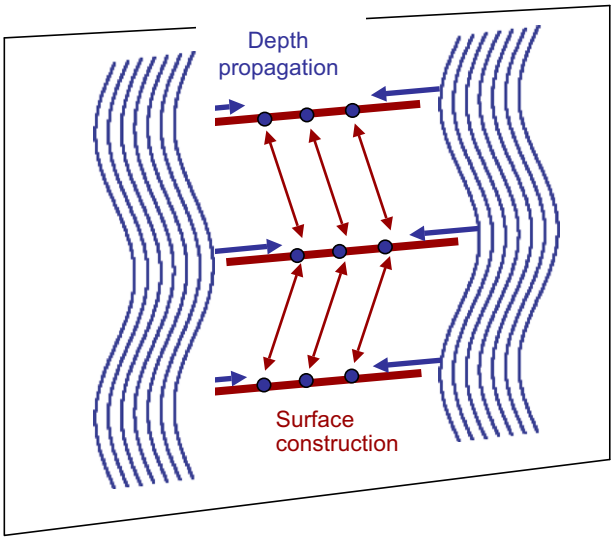

Fig. 9. The induced surface is not merely a replica or completion of the flanking surface. Our results suggest that depth values from the disparity in the flanks are propagated mainly along horizontal *features* in the test region, and then a visible surface is constructed to be consistent with these depth values and perhaps other available cues such as shading.

4.1. Depth propagation

Our results reveal some new characteristics of the depth interpolation process, and suggest that it may be best interpreted in at least two distinct stages: depth propagation and surface construction. The idea is sketched in Fig. 9. In experiment 1, we found that corrugated (disparity-modulated) flanking regions to the left and right of a central test region could induce a strong apparent corrugation in the test region. This induction was strong when the test region contained a dense texture of horizontal lines, monocularly or binocularly, but was weak or absent for vertical lines or a random 2-D texture, even when they were monocular and lacked any explicit disparity. It cannot be the strong texture boundary that blocked interpolation because that was also present for the horizontal lines where interpolation was strong (Fig. 1B) and absent for the 2-D test texture where interpolation was very weak. One might suppose that depth would be interpolated when the binocular and monocular portions of the images appeared to belong to the same object or surface. If so, we should expect best interpolation with the monocular 2-D test that matches the flanking texture, but this was not the case. Horizontal test lines were best (Fig. 4), despite the strong texture segmentation between the test and flanking regions.

We are therefore led to the view that depth values propagate selectively along horizontal contours. Extended horizontal contours are unique because they are the *only* image structure that cannot convey the horizontal binocular disparity information that is crucial for stereopsis. Thus the visual system might come to treat depth interpolation along horizontal contours as mandatory, while for most other image features it is not. A 2-D texture lacks disparity when it is monocular, or when the disparity is outside the range of stereopsis, but only horizontal contours are always disparity-free. This gives them a special need for depth interpolation. Whether depth would flow up and down vertical contours from induction regions above and below is not yet known, but on this view we should expect it to be less compelling than for horizontal.

Nishina, Okada, and Kawato (2003) also noted that the disparity of a horizontal bar is ambiguous except at its end-points, and discussed the idea, and the evidence, that to resolve the ambiguity depth information is propagated along the bar by a process that takes time to complete. From their psychophysical data, based on judgements of depth in the centre of a bar whose endpoint disparities were modulated over time, the transmission speed is of the order of

3–6 ms/deg. Nishina and Kawato (2004) also developed a successful computational model of depth propagation carried out by a local, iterative process based on short-range connections between neurons. Our results (experiment 2, Fig. 5) could be interpreted to mean that this functional chain of connections is broken by a small gap between the inducer and the 'carrier' contour.

It is also possible that depth propagation is involved in the detection of disparity itself. Tyler and Kontsevich (2001) studied the detectability of 'depth ripples' (disparity modulation) for horizontal and vertical ripples of several different extents. They found that the highest sensitivity was for long, horizontal ripples, and so argued for the existence of extended summation fields specialized for detecting horizontal 3-D structure. It is plausible, therefore, that the mechanisms for propagation of depth and for the extended spatial integration of disparity variation are functionally related.

Buckley, Frisby, and Mayhew (1989) provided a large set of demonstrations that extended Collett's (1985) work on surface interpolation. They used 'random-curve stereograms', depicting a V-shaped ridge with a sharp horizontal 'crease' in depth. We consider just two of their effects. (i) When the entire horizontal crease region was blanked out in one eye, the perceived shape of this region was fairly flat (as Collett also observed), and not sharply creased. (ii) When a central, vertical region was blanked out in one eye, leaving the stereo flanks intact (more like our displays), a sharp ridge was seen in the monocular region, much like the original stereo ridge that had no blanked regions. Both of these effects are consistent with the simple propagation of depth from the borders of the stereoscopic flanks into the monocular region. If we are correct about the role of contours in depth induction, then the success of the Buckley et al. demonstrations may lie in their choice of 'random-curve' patterns, rich in continuous overlapping contours, that were able to carry depth into the monocular region.

We acknowledge that there is other evidence suggesting that depth can propagate into monocular 2-D textures (Collett, 1985), and through empty regions (Wurger & Landy, 1989). We conjecture therefore that depth propagation may be relatively weak in these cases and much more effective when 'carried' by contours. Experiment 3 provides direct support for that idea. Depth was induced when lines or edges were aligned with the inducing peaks and troughs of disparity, but not when they lined up with the zero values of disparity. Hence depth carried by these features was strong, while any depth propagated through empty space must have been weak.

When people detect disparity modulation (DM), they could be using mechanisms sensitive to disparity *per se*, disparity gradient (slant), or disparity curvature. From measures of DM sensitivity at different spatial frequencies, Lunn and Morgan (1997) concluded that human vision is more sensitive to variation in disparity than to variation of slant or curvature. It is therefore interesting to ask, for depth induction, what kind of depth code is being propagated along the test contours? We think it is most likely to be depth values that propagate, rather than slant or curvature. If it were slant values, the phase effects (Fig. 6B and C) should be reversed: induction would be best when the test features were aligned with points of greatest inducing slant – namely the zero-crossings – the opposite of what we observed. If it were curvature, propagated from the peaks and troughs, this would be consistent with the observed phase dependence, but presumably the induced shape would always be curved, never the triangle-wave that was consistently observed (square-wave slant profile; Fig. 6B and C).

4.2. Surface construction

We suggest that, in the absence of explicit local disparity, the perceived test surface is constructed from depth samples acquired from the flanks, mainly by propagation at 1-D feature points. When those features were dense (the horizontal noise), a smoothly curved surface was constructed. In the line and square-wave test images there were only two features (lines *or* edges), hence two depth samples, per inducing period. When the line or edge features were aligned with flanking peaks and troughs of depth, the resulting triangle-wave surface might suggest that a simple join-the-dots procedure (linear interpolation) underlies the surface construction. Similarly, when those features were displaced by one-quarter period, to align with the DM zero-crossings, they would all acquire the same depth value, and linear interpolation would create the flat surface observed (Fig. 6B and E; 0/180 phase). When there are very many samples (the noise image), even linear interpolation would look smoothly sinusoidal, as observed. However, the smooth surface seen with the cosine test image (experiment 3) suggests that shading and stereo information are likely to be combined in the surface construction procedure. If the visual system has a prior for interpreting smooth luminance variation as shading, arising from smooth changes in surface orientation, rather than reflectance change (Schofield, Hesse, Rock, & Georgeson, 2006), then this would be incompatible with a triangle-wave surface, but would support the construction of a smoothly undulating one. By the same token, the uniform luminance regions of the square-wave grating, taken as shading, imply a constant surface gradient between the edge features, and so strongly support the observed triangle-wave surface. They are not compatible with an undulating surface, simply because the luminance gradients don't change smoothly.

The details of the surface construction process remain unknown, but the integration of local disparity, induced depth, luminance shading and other cues must surely be central to it. Likova and Tyler (2003) for example, found evidence that shading and disparity information interacted in a positional acuity task that required spatial interpolation within sparsely sampled targets. They found that the ability to interpolate the luminance shading cue (between narrow lines spaced 15 min arc apart) could be completely cancelled by a small but critical amount of disparity that signalled depth opposite to the shading; at this point the task became impossible. Thus they argued that interpolation – even for luminance cues – takes place at a level of surface depth construction that combines multiple cues to depth. See Nakayama and Shimojo (1996) for a broad discussion of the constraints and heuristics that may be involved.

5. Conclusion

Perceiving a 3-D surface can involve at least two kinds of interpolation process: depth propagation along horizontal 1-D features and surface construction from local depth estimates. In our experiments, the induced surface is not merely a replica or completion of the flanking surface. Instead, our results point to a scheme for surface perception like that sketched in Fig. 9. Explicit disparities encoded in the flanking regions propagate their depth values along 1-D features where disparity is unspecified. From these propagated depth samples, guided by local shading information (and no doubt other pictorial depth cues if present), a visible surface is constructed. These two processes could be successive stages of visual processing, with the second operating on the output of the first.

Acknowledgment

We thank the Engineering and Physical Sciences Research Council (U.K.) for grant support (GR/S07254/01, GR/S07261/01).

Appendix A

A.1. Integrating the slant data to estimate perceived surface shape

First we used the eight measured slant values (evaluated at intervals of 1/8 period) to interpolate intervening values (at 1/24 period intervals), using a smooth, shape-preserving piecewise cubic spline interpolation (*interp1* in *Matlab 7.4*). We assumed that the perceived surface did not appear to slant backward or forward overall, and so we corrected for any asymmetry of that kind in the data. To do this, we forced the mean slant to be 0, by subtracting the average slant value from all individual points. Then we integrated these gradient values (using *Matlab*'s *cumtrapz* function) to recover the perceived surface. Since the units of perceived depth are unknown, we scaled the control surface to have range ±1, with mean 0 (Fig. 8D, thick dashed curve). The amplitude of the recovered test surfaces was then scaled in units of the control surface, thus preserving relative amplitude. The unknown constant of integration (the absolute depth of the surface) was handled as follows:

(1) for V noise in front, assume that the induced surface has mean depth = 0, the same as the inducing surface;
(2) for V noise at the middle or back position, assume that the nearest point of the induced surface coincides with the V plane itself.

The surfaces so produced are plotted in Fig. 8D, E and F as solid dots at the measured and interpolated positions. The interpolation produced nice smooth surfaces, but we also confirmed that the surfaces were essentially the same even without interpolation.

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.visres.2008.09.030.

References

Blakeslee, B., & McCourt, M. E. (1997). Similar mechanisms underlie simultaneous brightness contrast and grating induction. *Vision Research, 37*(20), 2849–2869.

Bradshaw, M. F., & Rogers, B. J. (1999). Sensitivity to horizontal and vertical corrugations defined by binocular disparity. *Vision Research, 39*(18), 3049–3056.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436.

Buckley, D., Frisby, J. P., & Mayhew, J. E. W. (1989). Integration of stereo and texture cues in the formation of discontinuities during three-dimensional surface interpolation. *Perception, 18*(5), 563–588.

Collett, T. S. (1985). Extrapolating and interpolating surfaces in depth. *Proceedings of the Royal Society of London Series B-Biological Sciences, 224*(1234), 43–56.

Glennerster, A., & McKee, S. (2004). Sensitivity to depth relief on slanted surfaces. *Journal of Vision, 4*(5), 378–387.

Glennerster, A., McKee, S. P., & Birch, M. D. (2002). Evidence for surface-based processing of binocular disparity. *Current Biology, 12*(10), 825–828.

Kham, K., & Blake, R. (2000). Depth capture by kinetic depth and by stereopsis. *Perception, 29*(2), 211–220.

Knill, D. C. (1992). Perception of surface contours and surface shape: from computation to psychophysics. *Journal of the Optical Society of America A, 9*(9), 1449–1464.

Li, A., & Zaidi, Q. (2000). Perception of three-dimensional shape from texture is based on patterns of oriented energy. *Vision Research, 40*(2), 217–242.

Li, A., & Zaidi, Q. (2001). Information limitations in perception of shape from texture. *Vision Research, 41*(12), 1519–1534.

Likova, L. T., & Tyler, C. W. (2003). Peak localization of sparsely sampled luminance patterns is based on interpolated 3D surface representation. *Vision Research, 43*(25), 2649–2657.

Lunn, P. D., & Morgan, M. J. (1997). Discrimination of the spatial derivatives of horizontal binocular disparity. *Journal of the Optical Society of America A, 14*(2), 360–371.

Marr, D. (1982). *Vision*. San Francisco: Freeman.

McCourt, M. E. (1982). A spatial frequency dependent grating induction effect. *Vision Research, 22*(1), 119–134.

Mitchison, G. J., & McKee, S. P. (1987a). Interpolation and the detection of fine-structure in stereoscopic matching. *Vision Research, 27*(2), 295–302.

Mitchison, G. J., & McKee, S. P. (1987b). The resolution of ambiguous stereoscopic matches by interpolation. *Vision Research, 27*(2), 285–294.

Nakayama, K., & Shimojo, S. (1996). Experiencing and perceiving visual surfaces. In D. C. Knill & W. Richards (Eds.), *Perception as Bayesian Inference* (pp. 391–407). Cambridge: Cambridge University Press.

Nishina, S., & Kawato, M. (2004). A computational model of spatio-temporal dynamics in depth filling-in. *Neural Networks, 17*(2), 159–163.

Nishina, S., Okada, M., & Kawato, M. (2003). Spatio-temporal dynamics of depth propagation on uniform region. *Vision research, 43*(24), 2493–2503.

Pessoa, L., & De Weerd, P. (2003). *Filling-in: from perceptual completion to cortical reorganization*. New York: Oxford University Press.

Ramachandran, V. S. (1986). Capture of stereopsis and apparent motion by illusory contours. *Perception & Psychophysics, 39*(5), 361–373.

Rubin, N. (2001). The role of junctions in surface completion and contour matching. *Perception, 30*(3), 339–366.

Schofield, A. J., Hesse, G., Rock, P. B., & Georgeson, M. A. (2006). Local luminance amplitude modulates the interpretation of shape-from-shading in textured surfaces. *Vision Research, 46*(20), 3462–3482.

Schumer, R. A., & Julesz, B. (1984). Binocular disparity modulation sensitivity to disparities offset from the plane of fixation. *Vision Research, 24*(6), 533–542.

Stevens, K. A. (1981). The visual interpretation of surface contours. *Arificial Intelligence, 17*, 47–73.

Tyler, C. W. (1974). Depth perception in disparity gratings. *Nature, 251*(5471), 140–142.

Tyler, C. W., & Kontsevich, L. L. (2001). Stereoprocessing of cyclopean depth images: horizontally elongated summation fields. *Vision Research, 41*(17), 2235–2243.

Wilcox, L. M. (1999). First and second-order contributions to surface interpolation. *Vision Research, 39*(14), 2335–2347.

Wilcox, L. M., & Duke, P. A. (2005). Spatial and temporal properties of stereoscopic surface interpolation. *Perception, 34*(11), 1325–1338.

Wu, X. N., Zhou, Q., Qi, X. L., & Wang, Y. J. (1998). Stereo capture: local rematching driven by binocularly attended 3-D configuration rather than retinal images. *Vision Research, 38*(14), 2081–2085.

Wurger, S. M., & Landy, M. S. (1989). Depth interpolation with sparse disparity cues. *Perception, 18*(1), 39–54.

Yang, Y. D., & Blake, R. (1995). On the accuracy of surface reconstruction from disparity interpolation. *Vision Research, 35*(7), 949–960.

Yin, C., Kellman, P. J., & Shipley, T. F. (1997). Surface completion complements boundary interpolation in the visual integration of partly occluded objects. *Perception, 26*(11), 1459–1479.

Yin, C., Kellman, P. J., & Shipley, T. F. (2000). Surface integration influences depth discrimination. *Vision Research, 40*(15), 1969–1978.