Perception, 2000, volume 29, pages 437 – 452

DOI:10.1068/p2955

When texture takes precedence over motion in depth perception

Justin O'Brien, Alan Johnston
Department of Psychology, University College London, Gower Street, London WC1E 6BT, UK;
e-mail: justin.o'brien@ucl.ac.uk
Received 21 July 1999, in revised form 3 February 2000

Abstract. Both texture and motion can be strong cues to depth, and estimating slant from texture cues can be considered analogous to calculating slant from motion parallax (Malik and Rosenholtz 1994, report UCB/CSD 93/775, University of California, Berkeley, CA). A series of experiments was conducted to determine the relative weight of texture and motion cues in the perception of planar-surface slant when both texture and motion convey similar information. Stimuli were monocularly viewed images of planar surfaces slanted in depth, defined by texture and motion information that could be varied independently. Slant discrimination biases and thresholds were measured by a method of single-stimuli binary-choice procedure. When the motion and texture cues depicted surfaces of identical slants, it was found that the depth-from-motion information neither reduced slant discrimination thresholds, nor altered slant discrimination bias, compared to texture cues presented alone. When there was a difference in the slant depicted by motion and by texture, perceived slant was determined almost entirely by the texture cue. The regularity of the texture pattern did not affect this weighting. Results are discussed in terms of models of cue combination and previous results with different types of texture and motion information.

1 Introduction

Both texture and motion can be strong cues to depth. Experimental analysis of the properties of textured objects and surfaces that allow the visual system to estimate depth have been investigated many times since Gibson (1950) specified the texture density gradient as an explicit cue to depth. Size-constancy illusions from texture gradients are strong and robust, while a variety of algorithms have been developed which demonstrate the manner in which depth properties such as slant can be calculated from texture scenes (for review, see O'Brien 1997).

Motion can provide similarly strong cues to depth. Rogers and Graham (1979) demonstrated effects akin to stereoscopic viewing with motion parallax produced by random-dot fields yoked to an observer's head movement. Farber and McConkie (1979) questioned whether depth could be perceived without this head movement, but subsequent studies have shown observers' sensitivity to velocity gradients in a number of paradigms (Braunstein and Andersen 1981; Braunstein et al 1993).

A combination of texture and motion cues in depth-perception studies has usually led to a result indicating that motion is assigned a higher weighting by the visual system when there is a disparity between the depth indicated by the two cues (Braunstein 1968; Young et al 1993). Braunstein (1968) reports texture having little influence on the perception of slants of textured planar surfaces which were moving. Young et al (1993) found a general higher motion weight in a perturbation analysis of texture and motion cues with an experiment involving the assessment of depth relations with a stimulus consisting of a hemicylinder rotating back and forth in depth. While these reports provide strong evidence for the predominance of motion over texture within particular methodological and stimulus paradigms, it cannot be concluded from these studies that motion cues in general are stronger than texture cues in depth-perception tasks.

In common with a number of studies, Braunstein (1968) used a texture stimulus in which pictorial cues to depth conflicted. Since the dots of the display employed in the study were of invariant size, only the density of the dot spacing and the distances between dots graded in a manner appropriate for the depicted slant. It is not therefore possible to judge the extent to which the strength of the motion cue was intrinsic, or influenced by the cue-conflicts of the texture cues (Stone 1993).

In the study of Young et al (1993) the volumetrically defined texture was irregular, where the motion cue was not, and the type of motion used was a rotation about a horizontal axis. In this case, motion and texture cues were providing very different information about the depth (radius) of the rotating hemicylinder stimulus. The velocity field and the texture density gradient were very different. It is possible, however, for texture and motion cues to be used to provide identical information about depth, in cases where either cue alone can accurately define, say, the slant of a surface.

Koenderink and van Doorn (1976) showed that, in a region of an optic flow field, the flow could be completely described as the sum of a translation (trans) and three differential invariants—div, curl, and def. Following this, Freeman et al (1996) have demonstrated that slant can be calculated from the def component of an optic flow field, and that human observers' estimates of slant from moving random-dot displays are consistent with an estimate of this component.

Calculating slant in depth using motion information and using texture information can be very similar processes. Models that calculate slant from texture can be considered to fall into one of three categories. First, gradient-based 'invariant-seeking' approaches, usually implemented under perspective projection (eg Blostein and Ahuja 1989) rely on the extraction of texels from an image. Second, 'value-seeking' approaches (eg Witkin 1981; Stone 1993) rely on a consistent property of surface statistics. Witkin (1981), for example, assumes an isotropic distribution of edge segments, that is segments at every orientation in every part of the surface; while Stone (1993) assumes only a homotropic distribution, that is the same orientation distribution in different parts of a surface.

More recently, Malik and Rosenholtz (1994) have put forward a third method based explicitly on an analogy with depth from motion parallax. In a moving scene, relative depth can be calculated from the change in spatial relations over a number of frames, while with texture, these can be calculated by the change in spatial relations within a single image. Whereas Freeman et al (1996) demonstrated that slant could be calculated from the deformation component of an optic flow field, Malik and Rosenholtz show that slant can be calculated from texture deformation in a single image.

The pattern of variation in texture and motion information can provide similar information about, for example, the slant of a planar surface. Most models of depth cue combination, however, are concerned with different forms of representation, such as depth, slant, or curvature. Consequently they have as a primary concern how different depth cues are 'promoted' so that they provide equivalent or comparable measures of depth (Clarke and Yuille 1990). Kinetic depth effect, for example, provides an absolute estimate of depth only after fixation distance and perceived rotation direction have been established, while depth from stereo requires that interocular separation and gaze angles are known.

It is a difficult task, however, to analyse the weights assigned to different visual depth cues in conditions in which the cues can be inconsistent and impoverished. Young et al (1993) conclude a study of depth cue combination with the statement that a full description of the combination rules for such stimuli seems likely to resemble a microcosm of cognitive processing: elements of memory, learning, reasoning, and heuristic strategy may dominate (page 2695).

With texture and motion cues in slant, however, there are circumstances in which the same information about slant is provided in different modalities. The concern here is which information is more important to, or is more readily processed, by the visual system. Young et al were testing the modified weak fusion model using texture and motion cues, though it has also been applied to combinations including stereo information amongst others (Maloney and Landy 1989; Landy et al 1991, 1995). The modified weak fusion model is weak because it relies on a combination of independent estimates, but is modified because cues have to interact at an early stage in order that they be promoted to the common format required for integration. Cue promotion is necessary because normally different cues provide different types of information. The model relies on a system of weighted linear combination. The model furthermore assumes that any depth combination rule is dynamic, that is it changes in accordance with the reliability of the information from each type of cue; and the rule of combination is robust: lower weights are assigned to a cue whose depth estimate differs significantly from those of other cues.

There are, however, a number of naturally occurring scenarios in which different cues can provide very similar and very robust depth information. An investigation of the way in which cues are combined in these circumstances provides a benchmark from which to expand a model to include cues of varying reliability and conflicting depth information. Figure 1 is an example of a size-constancy illusion, in which identical globes appear to differ in size because of the depth indicated by the texture pattern on which they are superimposed. The texture patterns used are (a) a horizontally oriented luminance sine-wave grating at a slant angle of approximately 75°, (b) a vertically oriented grating at the same slant angle, and (c) a plaid comprising a summation of the previous two. The images are best viewed monocularly through an aperture. Observers normally report that the strength of the illusion is similar for both 1-D cases, and stronger for the 2-D case; and that it is not readily apparent that plaid (c) is a sum of the gratings in (a) and (b). It is easy to envisage the motion cues that could generate the same depth information as the texture here, since a drifting of the grating on the texture plane would lead to a motion parallax in the image plane.

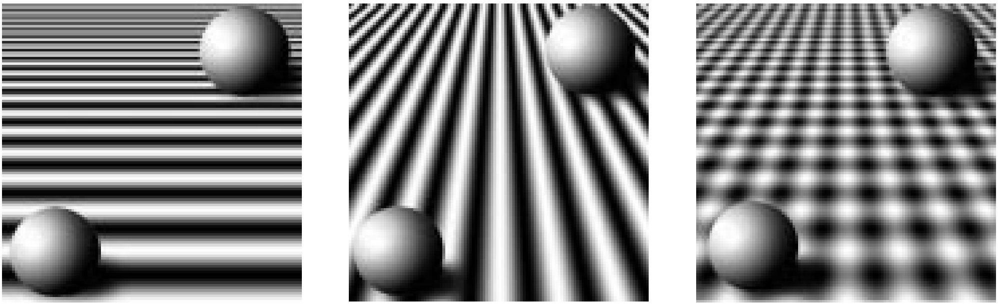

(a) (b) (c)

Figure 1. Size-constancy illusion. Identical globes appear to differ in size because of the depth indicated by the texture pattern on which they are superimposed. The texture patterns used are (a) a horizontally oriented luminance sine-wave grating at a slant angle of approximately 75°, (b) a vertically oriented grating at the same slant angle, and (c) a plaid comprising a summation of the previous two. The images are best viewed monocularly through an aperture.

A series of four experiments was therefore conducted for the purpose of the investigation of the relative weighting of motion and texture cues in slant perception, where both cues provided the same pattern of information about the slant of a surface. In the first, identical motion and texture cues were used, to test how the presence of motion in a texture stimulus affected an observer's slant discrimination threshold and bias. In the second, stimuli were used in which there was a disparity between the slant depicted by texture and that by motion. In the third, an irregular texture pattern was used in order to determine the effect of texture regularity on texture weighting. In the fourth, thresholds for detecting changes in spatial-frequency gradients and in velocity gradients were measured, in order to establish the relationship between cue strength and cue weighting.

2 General methods

It has usually been found, when measuring perception of planar surface slant, that observers underestimate the true slant of a surface. These underestimates can be very large indeed (Braunstein 1968). This is despite the fact that a wide variety of methods have been applied to the task. For example, observers' estimates of the slant of a surface have been determined with a kinaesthetic palm board (Gibson 1950), with a pivoting board set visually (Gruber and Clark 1956), or an unmarked rotating dial (Smith 1956). Other methods involved simply asking subjects which stimulus looked most convincing as a surface receding in depth (Vickers 1971; Cutting and Millard 1984). More conventional procedures include adjusting a test stimulus to a control (eg Epstein et al 1962). The underestimation finding is repeated, whether subjects view real surfaces binocularly (Gruber and Clark 1956) or 2-D images monocularly, and with or without occluding apertures. There is clearly a conceptual difficulty in measuring absolute slant, where a judgment is made by comparing the test stimulus to a control whose slant must also be judged by perceptual means. This problem is obviated in the present study by measuring only relative slant—whether one surface appears more or less slanted than another.

In the experiments reported here, the method of constants is adopted, in which an observer makes a forced choice between test and standard stimuli, the test stimulus being selected at random from a predetermined set of stimulus levels. A psychometric function is fitted as a cumulative normal sigmoid curve to the data obtained from each experimental run. An observer's slant discrimination threshold is calculated as the standard deviation of the underlying error distribution, and his/her slant discrimination bias as its mean (Watt 1991).

In order that an observer base a discrimination on perceived slant, rather than spurious pictorial artifacts, the method of single stimuli (MOSS) was adopted. MOSS involves temporal as opposed to spatial separation of comparison stimuli in a 2AFC task. In the first phase of an experiment, a standard stimulus was presented five times, in order that an observer might reliably remember the slant of the surface. In the second phase of an experiment, test stimuli of different slants were displayed. The observer compared the current stimulus with an internal representation of the standard stimulus. In order that the observer maintain an accurate model of the standard stimulus, it was presented on 20% of trials. Slant angles were chosen at random from the stimulus set during the experiment, with feedback on the observer's response. Data for these conservation trials were not analysed. Also, to ensure that an observer makes a judgment based on comparison of stimulus slant, and not spurious pictorial similarities, the spatial frequency of the test slant grating was varied randomly in pilot trials. Since this had no effect on slant discrimination thresholds in comparison with conditions in which test grating spatial frequency was constant, this control was omitted from the experiments reported here.

In experiments 1–3, the standard stimulus was an image of a planar surface oriented at a slant angle of 45°. This was textured with a plaid, composed of two orthogonal luminance sine-wave gratings. It was displayed five times, for 1.0 s, with an interstimulus interval of 1.5 s. The test stimuli were surfaces oriented in depth, textured with luminance sine-wave gratings that were horizontal, vertical, or the plaid summation of the two (with the exception of the irregular texture used in experiment 3). These were similar in kind to those illustrated in figure 1. They could be animated by drifting the gratings over time. Within an experiment, each test stimulus was presented eight times at each of nine levels (25° to 65°, at 5° intervals). These levels were determined in preliminary trials. In order to prevent perseveration and stimulus bias, different conditions were mixed within one block of trials.

2.1 *Apparatus*

The stimuli were pre-generated on a Sun Sparc 10 (the grating stimuli) or a Silicon Graphics Indigo (the irregular cloud stimulus of experiment 3). The stimuli were displayed by the Sun on a 19-inch Trinitron monitor, which had a dot pitch of 0.30 mm and a non-interlaced refresh rate of 66 Hz. Stimuli were displayed in a window measuring 512×512 pixels, updated at a double-buffered (flicker-free) 33 Hz. Stimuli were viewed monocularly from a distance of 50 cm through a circular aperture that occluded the edges of the window, positioned at 25 cm from the display (to ensure that the display appeared separated in depth from its surround). The display measured 15.4 cm × 15.4 cm, and subtended an angle of 17.5 deg at the eye.

2.2 *Texture components*

The spatial frequency for the sine-wave gratings was set at 0.75 cycle deg^{-1}. For a grating of 0° slant, this spatial frequency would hold throughout the image. For a surface slanted at 45°, the spatial frequency of the horizontal grating varied from 4.36 cycles deg^{-1} at the top of the display to 1.09 cycles deg^{-1} at the bottom. For vertical gratings at 45°, the corresponding frequencies were 1.11 and 0.45 cycles deg^{-1}. For a grating at 65°, the horizontal component maximum and minimum frequencies were 5.1 cycles deg^{-1} and 1.1 cycles deg^{-1}; and for the vertical component 1.1 cycles deg^{-1} and 0.56 cycle deg^{-1}. For the 25° grating these figures were 1.07 and 0.71, and 0.81 and 0.71 cycles deg^{-1}, respectively.

2.3 *Animation*

The stimuli were animated by generating 33 frames of the slanted grating texture in which the phase of the grating was shifted by $\frac{1}{33}$ of a cycle (0.19 radians) in each frame. These were displayed over a period of 1.0 s. The velocities in the motion stimuli are inversely proportional to the texture spatial frequencies. Since the moving stimuli were generated with a temporal frequency of 1 cycle s^{-1}, inverse velocity measured in seconds per degree has the same numerical value as spatial frequency measured in cycles per degree.

For a horizontal grating at 45°, velocities varied from 0.62 deg s^{-1} at the top of the display to 1.2 deg s^{-1} at the bottom, and at the centre of the display 0.86 deg s^{-1}, with corresponding velocities of 1.09, 1.56, and 1.33 deg s^{-1} for a vertical grating. For a horizontal grating at 65°, the corresponding velocities were 0.19 deg s^{-1} and 0.94 deg s^{-1}, with a centre-display value of 0.47 deg s^{-1} (vertical equivalents: 0.92, 1.8, and 1.3 deg s^{-1}). At the slant of 25°, the minimum and maximum velocities were 0.94 and 1.4 deg s^{-1} for the horizontal grating, and 1.2 and 1.4 deg s^{-1} for the vertical grating.

To illustrate that the stimuli used were within the range normally experienced in natural environments, we can consider the example of a person 6 ft (1.83 m) tall jogging at about 6 miles h^{-1} (2.4 m s^{-1}) along a level surface looking not directly towards the horizon, but instead to a point on the ground about 8 eye heights or 48 ft (14.64 m) ahead of the subject's feet. The ground here will flow at a rate of about 1 deg s^{-1}.

3 Experiment 1: Consistent motion and texture cues

There is something of a consensus in the slant-from-texture literature, with most studies concluding that linear perspective cues are stronger than compression cues in depicting slanted planar surfaces (Attneave and Olson 1966; Gillam 1970; Vickers 1971; Rosinski 1974; Cutting and Millard 1984). This would suggest that the vertical grating in figure 1 would appear more slanted than the horizontal grating. Occasionally, results contradict this (eg Saidpour et al 1997). It is evident that both the nature of the stimulus used and the method of measurement can influence the measured values (O'Brien 1997).

It is necessary, before measuring perceived slant where there is a disparity between motion and texture cues, to determine how motion and texture properties affect slant discrimination thresholds or perceived slant. This should allow a comparison of these types of 1-D and 2-D grating stimuli with other stimuli, and the methodology described here with others mentioned above.

It should be noted that there are similarities between the stimuli used in the following experiment and those employed by Flach et al (1992) in an altitude regulation task. Their task involved measuring sensitivity to spatial frequency and velocity gradients, although their variable of interest was not surface slant, but rate of expansion, and it has been established that div components in optic flow fields do not affect slant perception (O'Brien 1997). Flach et al (1992) used stimuli composed of parallel line elements in a simplified flight simulation task to measure observers' sensitivity to 1-D and 2-D texture and speed gradients. Their results indicated that texture parallel to the direction of motion (ie like the vertical grating texture in the following experiment) had a clear advantage over texture orthogonal to it (ie like the horizontal grating texture), as measured by the accuracy of simulated altitude regulation. Texture and motion cues were presented together, however. If the visual system is sensitive to both types of cues, then it would follow that the more of these cues are present in a stimulus, the more reliable and the more veridical a slant discrimination judgment based on them will be.

3.1 *Method*

3.1.1 *Design.* The stimuli in experiment 1 consisted of three texture types: a horizontal sine-wave grating, a vertical grating, and the plaid resulting from their summation. Each of these grating components could be animated, that is its phase could shift by one cycle during the presentation of the stimulus, with the direction of the phase shift randomised. There were thus two component textures and two types of component motion. These gave eight different combinations of texture and motion. Since motion and texture cues are not independent (a motion component cannot be present without the corresponding texture component also being present), the design can be considered to consist of a single factor—slanted stimulus type—with eight levels.

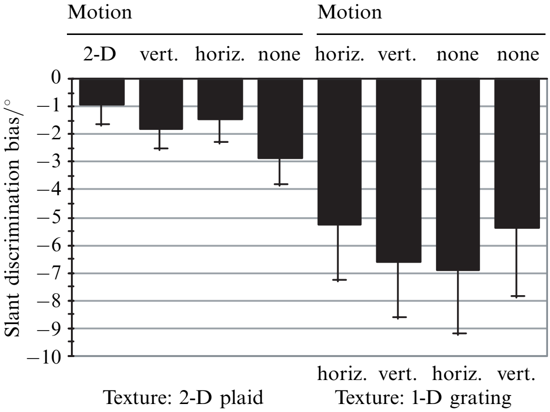
3.1.2 *Subjects.* The experiment was completed by eight observers, including one of the authors (JO).

3.1.3 *Procedure.* The standard stimulus included all four cues, that is both types of texture and both types of motion, giving a diagonally moving plaid. An experimental test block consisted of the initial standard stimulus presentations followed in random order by the eight stimulus types, each displayed eight times at each of their nine levels (mean 45°, at 5° intervals). The slant discrimination bias and slant discrimination threshold for each stimulus condition were estimated from these 72 measures. An additional 20% of trials were feedback trials. Stimulus duration was 1.0 s, with observer response times of up to 2.0 s. The presentation block was repeated twice to give three runs from which to take mean biases and thresholds.

3.2 *Results*

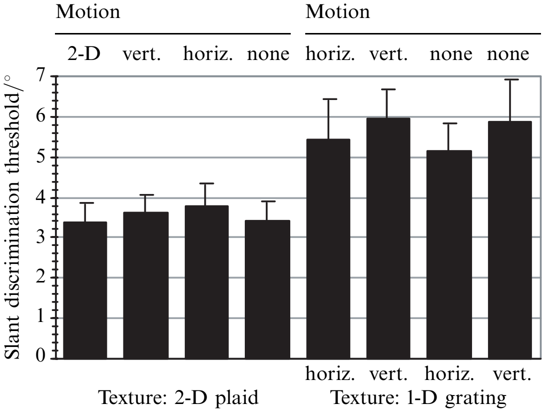
Figures 2 and 3 show the slant discrimination bias and threshold data respectively. For statistical purposes (Wilcox 1995), the median bias and median threshold were taken from the psychometric functions estimated for each of an observer's three runs, and it is the mean of the eight observers' medians which is plotted here. A clear trend is evident for both measures. Since the texture and motion cues are not independent, they were treated as a single factor in the experimental design. In testing the effects of motion and texture, it was therefore appropriate to use a *t*-test for the relevant comparisons (eg moving versus static, or 1-D versus 2-D), and adjust the critical p-values according to the number of comparisons made.

The stimuli in which the texture is 1-D generate slant discrimination thresholds which are higher than those for 2-D textures (paired samples *t*-test: $t_7 = 4.06$, $p = 0.005$). Surfaces with 1-D textures also appear much less slanted (larger negative bias) than those



vert., vertical
horiz., horizontal

Figure 2. The slant discrimination bias in each of the eight stimulus conditions indicates how much less slanted each stimulus type appeared compared to the standard stimulus with all four cues. The conditions in which the texture pattern is 1-D (the four on the right) show a larger slant underestimation than those with 2-D textures. Note that the bias for the full cues condition (extreme left) is less than 1°, where the test and standard stimuli were identical.



vert., vertical
horiz., horizontal

Figure 3. The mean of the slant discrimination thresholds for eight observers, across the eight motion/texture conditions. The difference in thresholds between 2-D textures (left half) and 1-D textures (right half) is evident.

with 2-D textures, as is evident from figure 1. This effect just fails, however, to reach significance at the 0.05 criterion level ($t_7 = 2.36$, $p = 0.051$). This does not, of course, mean that the perception of slant with 2-D textures is veridical. The standard stimulus was the 2-D texture with both types of motion, so all slant biases are relative to this. Thus, the expected bias in the full cues condition (2-D texture and 2-D motion) is 0°. In fact, it is 0.94°, indicating that observers can match surface slants to within 1°.

With respect to the threshold data, there is no consistent or large effect of adding a motion cue to a given texture pattern (a series of pairwise comparisons showed no significant effects). There is no mean difference between the 1-D textures, or between types of motion. Similar effects are found in the bias data. The bias for the 2-D motion/2-D texture is slightly lower than that for the static 2-D texture (paired samples *t*-test, $t_7 = 5.31$, $p = 0.001$) but the difference is small (less than 2°) compared to that between 1-D and 2-D cases. Otherwise, there is no significant effect of adding a motion cue to a static texture pattern.

3.3 *Discussion*

There is a considerable effect of texture type, yet very little of motion. Thresholds and biases are high for 1-D gratings, and low for 2-D plaids. The results do not indicate that one of the grating stimuli is a stronger cue to slant than the other. Nor is there any indication that the addition of a motion cue reduces slant discrimination thresholds or changes perceived slant.

The modest influence of motion cues on slant perception in these stimuli suggests that the visual system is less sensitive to velocity gradients (over the range of speeds tested) than it is to spatial-frequency gradients, despite that fact they contained the same information. As mentioned above, here the spatial frequency was inversely proportional to image velocity at every point in the image. It should also be noted that motion can be an effective cue to slant, as Freeman et al (1996) have demonstrated. This result is not consistent with the findings of Young et al (1993), since in this case, when there is a disparity between motion and texture slants, the texture cue predominates.

The largest effect was between 1-D and 2-D texture types. It is therefore clearly not the case that slant estimates are generated from both components of a 2-D texture independently, and the two are in some way combined (eg by linear weighted combination). This is not supported by the data, since the change in perceived slant for the plaid texture is less than that for either 1-D grating. This can be interpreted in a number of ways in relation to the models of slant estimation considered above. First, it could be that 2-D textures allow a disambiguation of the compression and scaling information in the horizontal component by using the scaling information exclusively to the vertical component. Second, in terms of gradient-based models, it is only with 2-D texels that density or area gradients can be calculated. Third, invariant-seeking models require properties of surface statistics present only in 2-D gradients.

The data for 1-D textures conflict with previous work cited above, since no dominance of any type of 1-D texture cue has been found above another. Previously, however, most studies (for example, Attneave and Olson 1966) have found that a horizontal (compression) texture generates only very poor impressions of slant compared to a vertical (perspective scaling or convergence) texture. This would suggest that there should be greater negative biases for horizontal than for vertical textures. More importantly, it might suggest that, since slant discrimination thresholds for the horizontal textures are as low as those for the vertical (at about 6°), observers are not basing their discrimination judgments on perceived slant, but on some other spurious cue which enables accurate responses to be made, but which is irrelevant to slant.

There are four grounds on which the assertion that observers' judgments were based on perceived slant can be justified. First, when the spatial-frequency texture pattern in the test stimulus is chosen at random from preset levels, results are identical to those when spatial frequency is fixed (O'Brien 1997), so it cannot be the case that observers responded to cues such as bar width or the number of bars present in the aperture. Second, preliminary trials on grating twist—the orientation of a texture pattern on a surface—(O'Brien 1997) indicate that, if the standard stimulus has a horizontal grating texture and the test stimuli have vertical grating textures, no bias effects are found, which would not be the case were different slants perceived. Third, the experimental procedure involved long runs in which all stimulus conditions were randomised, so it would not be possible for biases to be reduced by observers avoiding prolonged response in a single direction. Fourth, the possibility of the aperture through which the surface was viewed interfering or causing a 'frontal tendency' was reduced by using an intermediate aperture that had the effect of blurring the border of the stimulus, and appeared to all observers to be separated in depth.

4 Experiment 2: Motion – texture disparity—regular texture

Results from experiment 1 suggest that the motion cue is weak in comparison to the texture cue in slant perception, and consequently when there is a disparity between motion slant and texture slant, perceived slant would tend towards the slant depicted by the texture cue. Results from experiments using different paradigms suggest, however, that motion can be a stronger cue than indicated in experiment 1. It was shown that adding a 2-D-motion gradient to a 2-D-texture gradient increased perceived slant by nearly 2°, with no difference in slant discrimination threshold. Young et al (1993), however, showed that texture and motion cues could have roughly equal weights in a depth-discrimination task, when the motion cue results from rotation in depth.

It is therefore appropriate to determine the weights of texture and motion cues by perturbation analysis (Young et al 1993) for the stimulus generation and display paradigm used in experiment 1. Perturbation analysis involves the examination of changes in depth perception from two consonant cues as a function of a small difference between the cues.

A stimulus in which texture cues and motion cues indicate different slant angles can be constructed in the following way. Consider a texture pattern on a moving slanted surface which is itself an image of a slanted surface: the resulting texture pattern in the image plane will indicate a slant different to the motion slant. If the motion slant is y° and the texture slant is x°, the stimulus can be generated in three stages. First, an image is generated of a surface rotated in depth to a slant angle of $(x-y)^{\circ}$. Second, for each frame of the animation, this image is shifted slightly across the image plane. Third, this shifted image is itself rotated in depth to a slant angle of y°. The stimulus is the image of this final rotation. This was implemented in the stimulus generation algorithm as follows. Ray casting (Glassner 1989) was employed to calculate the intensity of a given pixel on the screen (image plane) by tracing the path of a ray from the point of observation, through the image plane to an interception point in the viewing frustrum. The point of intersection was first calculated between the ray from the observer in a direction D and the motion plane. A motion offset was applied to this for each frame of the animation. The ray then continued from this offset point on the motion plane in the original direction D, until it intercepted the texture plane. In other words, a ray was cast from the observation point O, in direction D, until it intercepted the texture plane. When texture and motion slants were different, the ray travelling in direction D was displaced in space to an extent determined by the orientation of the motion plane, before continuing in direction D to the texture plane.

4.1 *Method*

4.1.1 *Design.* There were two perturbation conditions. In the first, texture slant was held constant at 45°, while the motion slant was varied as a function of the texture slant at four set levels, explained below. In the second, motion slant was held constant at 45°, while texture slant was varied relative to this at four set levels. These levels were established during pilot trials. Clearly, a very high disparity between motion slant and texture slant produces an unsatisfactory stimulus, because it is important that, while the motion slants and texture slants are different, the observer nevertheless perceives the stimulus as coherent. It was established that a motion – texture slant disparity of 12° appeared coherent, even at very high slant angles (up to 77°). The four levels or relative slant were set at -12°, -5°, $+5^{\circ}$, and $+12^{\circ}$. The texture used was the 2-D plaid grating. The three types of motion described in experiment 1 were also employed.

4.1.2 *Subjects.* The experiment was completed by two observers: one of the authors (JO) and a naive observer who had not participated in experiment 1 (CB). The use of just two observers permitted a greater number of conditions in the experimental design, and is consistent with the implementation of perturbation analysis reported by Young et al (1993), especially since the intersubject variance in experiment 1 was low.

4.1.3 *Procedure.* The general procedure was as described for experiment 1. The standard stimulus was a plaid grating, for which texture slant and motion slant were both 45°. Test stimuli were presented eight times at each of nine levels (from 25° to 65°, at intervals of 5°). The slant discrimination bias and slant discrimination threshold for each stimulus condition were estimated from a psychometric function based on 72 measures. An additional 20% of trials were feedback trials. Stimulus duration was 1.0 s, with observer response times of up to 2.0 s. The presentation block was repeated twice to give three separate estimates of biases and thresholds for each condition.

4.2 *Results*

Figures 4a and 4b show the slant discrimination biases for observers JO and CB, respectively. The dashed lines indicate the bias when texture slant is held constant and motion slant is varied. For both observers, perceived slant changes very little with changes in slant depicted by any of the motion cues. The unbroken lines show the bias for the conditions in which motion slant is held constant and texture slant is varied. The trend here is for bias to follow very closely the change in the texture slant.

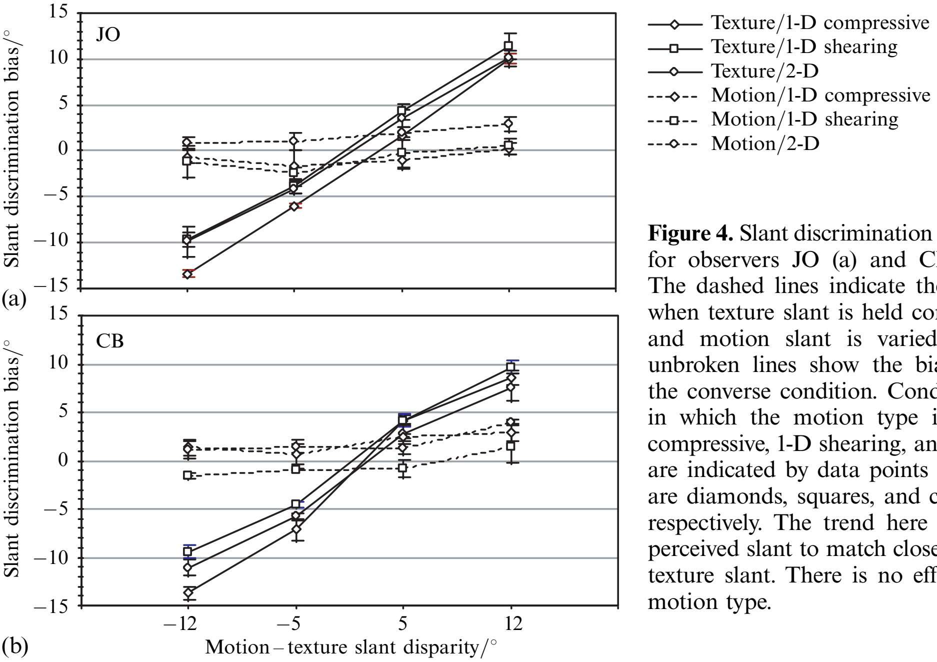


Figure 4. Slant discrimination biases for observers JO (a) and CB (b). The dashed lines indicate the bias when texture slant is held constant and motion slant is varied. The unbroken lines show the bias for the converse condition. Conditions in which the motion type is 1-D compressive, 1-D shearing, and 2-D are indicated by data points which are diamonds, squares, and circles, respectively. The trend here is for perceived slant to match closely the texture slant. There is no effect of motion type.

A repeated-measures ANOVA revealed no significant effect of motion type in either depth cue disparity condition for either observer. The data for the three motion types were therefore combined in a linear regression analysis to calculate the slope of the bias function in each disparity condition. When motion slant is constant and texture slant varies, slant discrimination bias $= 0.85 \times$ disparity $- 0.46$ (for disparity, $t = 30.187$, $p < 0.0001$; and for the intercept constant, $t = 1.740$, $p = 0.09$; $r^2_{\mathrm{adj}} = 0.97535$). When texture slant is constant and motion slant varies, slant discrimination bias $= 0.08 \times$ disparity $+ 1.57$ (for disparity, $t = 2.614$, $p = 0.0158$; and for the intercept constant, $t = 1.913$, $p = 0.07$; $r^2_{\mathrm{adj}} = 0.20235$). There was no significant effect of depth cue disparity on slant discrimination threshold.

4.3 *Discussion*

That texture is so much stronger a cue to slant than motion conflicts with previously discussed research, but is consistent with the result of experiment 1. With respect to the data of Young et al (1993), it is important to determine if the difference found is due to the regularity of the texture cue, since, in order to ensure that motion and texture information were quantitatively identical, experiments 1 and 2 used regular gratings, while Young et al used a volumetrically defined irregular texture. Thus, in experiment 3 the effect of using a highly irregular texture is considered.

If texture is assigned a higher weighting in slant discrimination tasks because discrimination thresholds for spatial-frequency gradients are much lower than those for velocity gradients, it might be expected that if the results were recalibrated in terms of perceptual units, weightings might equalise. In order to recalibrate the results, appropriate discrimination thresholds for changes in spatial-frequency gradients and in velocity gradients need to be established. This is the purpose of experiment 4.

5 Experiment 3: Motion – texture disparity—irregular texture

It was demonstrated in experiment 2 that when texture and motion cues provide similar information about the slant of a planar surface, the visual system assigns a much higher weighting to the texture information than to the motion information.

The more irregular a texture, the more difficult it is to calculate surface slant. If a surface is regular then a measure between any two texels will define the surface slant. The more irregular the surface, however, the more measures are needed to ensure that the estimated slant approximates the actual slant. There is no obvious metric by which to index texture irregularity, however, since the three different methods discussed above of calculating slant from texture would each suggest a different metric.

In order to investigate the effect of texture regularity, a number of texture images were rated on a subjective scale of irregularity by one of the authors. The texture that was ranked as least regular was a cloud-like pattern (Adobe Systems Inc. 1994). A texture that was ranked at an intermediate grade (halfway between the regular plaid of previous experiments and the cloud texture) was also used in pilot trials. These pilot trials showed no evidence, however, for a change in the weight of the texture cue with the intermediate texture, so only the cloud texture was employed in the experiment. The experiment was essentially identical to experiment 2, with the exception of the texture pattern of the test grating, which was a cloud image instead of a grating.

5.1 *Method*

5.1.1 *Design.* Experimental design was similar to that of experiment 2. Two disparity conditions were employed: one in which texture slant was held constant while motion slant varied, and another in which the reverse was the case. The same four levels of slant disparity as those used in experiment 2 were used here.

5.1.2 *Subjects.* Two observers completed the experiment: one of the authors (JO), and a volunteer who had participated in neither of the previous experiments (JS). Both had corrected vision.

5.1.3 *Apparatus.* Experimental apparatus was the same as for the previous experiments. Stimuli with the irregular texture pattern were generated in a slightly different manner, however, to the gratings used previously. Silicon Graphics OpenGL rendering routines were used to generate the surface slanted in depth by defining a texture MIP MAP with the cloud texture image and applying 3-D transformations to the surface to which the texture was applied. Where the texture slant was x° and the motion slant y°, the surface was oriented to a slant angle of $(x-y)^\circ$. This image was translated by a motion offset appropriate for a given frame of the animation. The transformed image was then oriented to a slant angle of y°. Because the texture was mapped after the transformations took place, there was no loss of image quality. (Effectively, the transformations merely specified the way in which the texture was distorted.)

5.1.4 *Procedure.* The procedure for experiment 3 was similar to that for experiment 2. The standard stimulus was a plaid grating, for which texture slant and motion slant were both set at 45°. Test stimuli were presented eight times at each of nine levels (from 25° to 65° in steps of 5°). The slant discrimination bias and slant discrimination threshold for each stimulus condition were estimated from a psychometric function based on 72 measures. An additional 20% of trials were feedback trials. Stimulus duration was 1.0 s, with observer response times of up to 2.0 s. The presentation block was repeated twice to give three runs from which to take mean biases and thresholds.

5.2 *Results*

Figure 5 shows the slant discrimination bias data. The data for plaid textures (indicated with open triangles) are taken from experiment 2 and are the means of the results for observers JO and CB. The data for cloud textures (indicated with open circles) are the means of the results from experiment 3 for the two observers, JO and JS. Data for the texture-varying condition are plotted with unbroken lines, and for the motion-varying condition with dashed lines. Biases for the cloud textures are in all cases less than for the equivalent plaid condition, indicating that surfaces with the cloud texture appeared less slanted than the plaid standard.

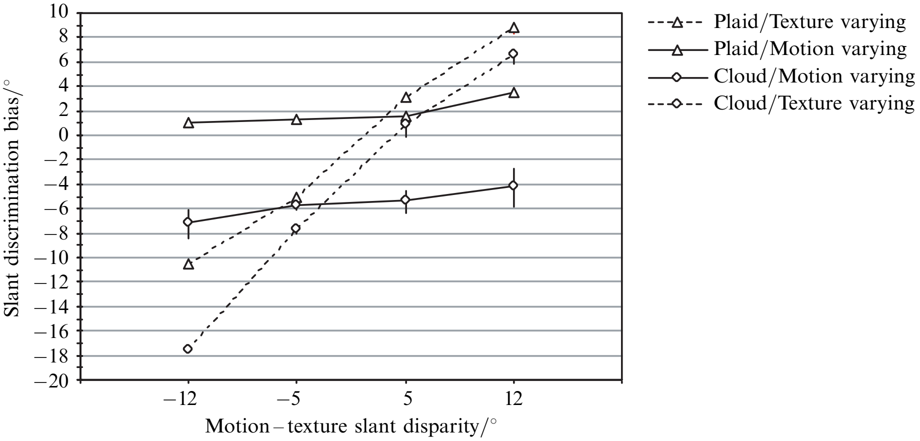


Figure 5. The data for plaid textures (indicated with open triangles) are taken from experiment 2 and are the means of the results for observers JO and CB. The data for cloud textures (indicated with open circles) are the means of the results from experiment 3 for the two observers, JO and JS. Data for the texture-varying condition are plotted with dashed lines, and for the motion-varying condition with unbroken lines. Biases for the cloud textures are in all cases less than for the equivalent plaid condition, indicating that surfaces with the cloud texture appeared less slanted than the plaid standard. When texture slant is held constant, and motion slant varies, the slope of the bias function for the cloud texture is almost identical to that for the plaid texture. When motion slant is held constant and texture slant varies the slopes of the bias functions for cloud and plaid textures are again almost identical.

When texture slant is held constant, and motion slant varies (the two data series plotted with solid lines), the slope of the bias function for the cloud texture is almost identical to that for the plaid texture. The slope for the plaid texture is 0.08 (see previous experiment) and for the cloud texture is 0.11 (linear regression: slant discrimination $\text{bias} = 0.11 \times \text{disparity} - 5.58$; both t-values ns; $r^2_{\text{adj}} = 0.20$). When motion slant is held constant and texture slant varies (the dashed lines) the slopes of the bias functions for cloud and plaid textures are again almost identical. The slope for the plaid textures (from the previous experiment) is 0.85, and for the cloud texture is 0.98 (linear regression: slant discrimination $\text{bias} = 0.98 \times \text{disparity} - 4.42$; for disparity, $t = 15.88$, $p < 0.0001$; and for the intercept, $t = -7.76$, $p = 0.0002$; $r^2_{\text{adj}} = 0.97$).

The texture and disparity types have no significant effect on slant discrimination thresholds, indicating that the slant of the irregularly textured surfaces was no more difficult to perceive than that of the grating-textured surfaces, and that a disparity between motion and texture slants did not make it more difficult to make a judgment of a single coherent slant angle.

5.3 *Discussion*

Despite early indications from pilot trials, there is no evidence here that the weighting attributed to the texture cue is reduced when the texture is irregular, even though the cloud-textured surfaces appeared less slanted than the plaids. It might be the case that an even more irregular texture would lend itself to lower weights on the texture cue, but, as discussed previously, there is no obvious metric of texture irregularity appropriate to slant-from-texture estimation. Furthermore, is should be noted that the cloud texture used in the experiment was selected as the least regular of a series of 30 natural texture images (Adobe Systems Inc. 1994). However textures are measured, it would seem likely that few are more irregular than the cloud texture used in this experiment. Indeed, images supplied by BAe Military Flight Simulators Ltd, which were retouched photographs of landscapes (forest textures, sand textures, mountain terrain textures, etc) were rejected on the grounds of these textures being far too regular.

It can therefore be assumed that the difference in results between experiment 2 and those of Young et al is due to the difference in the nature of the motion cues, in this case motion parallax of a planar surface (with velocities in a range lower than used in previous experiments) and in the case of Young et al (1993) a rotation about a horizontal axis of a hemicylinder.

6 Experiment 4: Gradient discrimination thresholds

The previous experiments have demonstrated that texture is a stronger cue than motion to the slant of planar surfaces within a given set of stimulus parameters, and that the regularity of the texture cue has no effect on this weighting. It is therefore necessary to determine the extent to which the visual system is more sensitive to changes in spatial-frequency gradients than it is to changes in velocity gradients. This serves a number of purposes. First, it will be possible to determine if the weights assigned to texture and motion cues, as established in experiment 2, do equalise if they are rescaled in terms of perceptual units. Second, this will serve as validation for the stimuli used, especially the motion cue, which all observers reported gave a strong impression of depth, yet the results of none indicate that this cue was particularly influential in their slant discrimination judgments. If thresholds for detecting changes in velocity gradients are very much higher than those for detecting changes in spatial-frequency gradients, it might suggest that the motion stimulus adopted is in some way suboptimal in terms of mean velocities used.

6.1 *Method*

6.1.1 *Design.* The experiment comprised two parts, spatial-frequency-gradient discrimination and velocity-gradient discrimination. Each consisted of three stimulus types. For the texture, the 1-D horizontal, 1-D vertical, and 2-D plaid textures were used. The motion stimuli consisted of compressive, shearing, and diagonal shearing/compressive motions of the plaid textures. Static textures were generated at a mean slant of 0°. Motion stimuli were generated in which texture slant was set at 0°. Levels were established during pilot trials. Nine levels at 10° intervals were used.

The static texture stimuli were generated as in experiment 1. The motion stimuli were generated as in experiment 2. A motion stimulus was effectively a surface of texture slant 0° and motion slant $n°$, where $-40 < n < 40$; that is, a grating to which an animated shear was applied. Slant levels were determined in pilot trials.

6.1.2 *Subjects.* The experiment was completed by the same two observers who completed experiment 3: one of the authors (JO) and a volunteer (JS). Both had corrected vision.

6.1.3 *Procedure.* The general procedure was as described for previous experiments, except that slant discrimination bias was not measured. It was only the ability to discriminate between stimuli that was of concern. Accordingly, observers were instructed to use whatever cues were available to make their discriminations, and it did not matter if they were unsure of the absolute slant of a surface, since their discrimination concerned only the polarity of this slant.

6.2 *Results*

Slant discrimination thresholds are displayed in figure 6. Thresholds for the motion stimuli are shown in black, and for the texture stimuli in white. The three groups on the x-axis are horizontal, vertical, and plaid texture/motion. The effect here across observers and across texture/motion types is evident: the thresholds for motion are higher than the thresholds for textures. The mean difference between texture and motion thresholds across stimulus type and observers is 6.44°. Motion thresholds are 31.3% higher than texture thresholds in the equivalent condition.

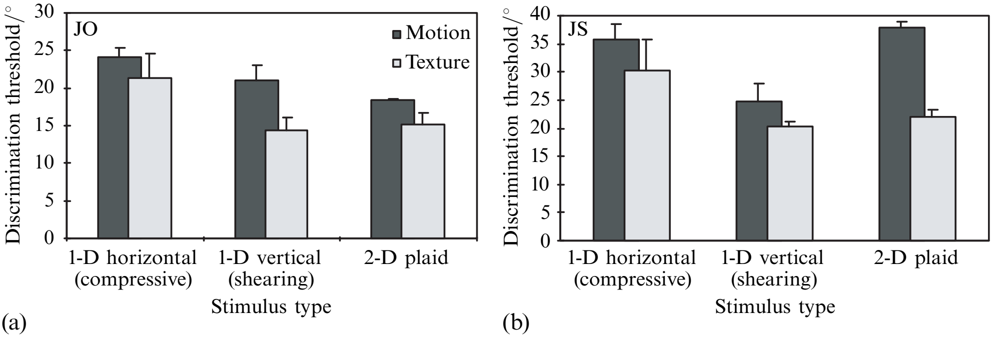


Figure 6. Thresholds for discriminating changes in spatial frequency and in velocity gradients for observers JO (a) and JS (b). Spatial-frequency thresholds are shown in white and velocity in black. For all three conditions, thresholds are higher for velocity gradients than for spatial-frequency gradients.

6.3 *Discussion*

Changes in velocity gradients are more difficult to perceive than changes in spatial-frequency gradients by a factor of 30%, but the difference is not as large as the weightings found in experiment 2 suggest, were a combinative mechanism of cue combination in operation. Motion thresholds are only 30% higher than texture thresholds, yet, in combination, motion gets less than 10% of the weight.

It is not the case, therefore, that, if the results from experiment 2 are rescaled in terms of perceptual units instead of degrees, weights equalise.

Furthermore, the 6.44° mean difference between spatial-frequency and velocity gradient thresholds vindicates the observers' impressions that the motion cue used was a strong cue to depth. This is despite that fact that the range of velocities present in the motion stimuli was generally lower than in other studies (eg Braunstein 1968) and that the stimulus duration was also shorter than employed elsewhere.

This result, coupled with that of experiment 2 (motion assigned a $< 10\%$ weight), suggests that the most appropriate model for cue combination in planar slant perception with motion and texture cues is one of veto (Clark and Yuille 1990). Subjects tend to use the most discriminable cue for surface slant estimation.

7 General discussion

The purpose of this paper is to illustrate an example of when texture is a stronger cue than motion in depth perception. Previous studies, such as those of Braunstein (1968) and Young et al (1993), described above, have demonstrated cases in which motion information is stronger than texture information. It has been shown here that when texture and motion cues contain equivalent information about surface slant, almost no weight is attributed to the motion cue by the visual system. This study has been motivated by computational studies of how slant can be calculated from both texture (Malik and Rosenholtz 1994) and motion information (Freeman et al 1996), using similar procedures.

The superficial conflict between our own results and those of Braunstein (1968), can be explained in terms of the difference between the texture cues used by us and those used by Braunstein: the texture used here was a ray-traced surface, in which there was no conflict between different texture characteristics, such as size or density. The difference between the results reported here and those of Young et al (1993) can be attributed to the type of motion used, namely motion parallax produced with planar surfaces, as opposed to a rotation in depth produced with curved surfaces (O'Brien 1997). It is not the case that the texture was found to have a higher weight because the motion cue used in the experiments reported above was weak, since the results of experiment 4 indicate that slant discrimination thresholds for motion slant were only 35% higher than for texture slant, even though the velocities employed were quite low, and the stimulus duration short.

We did not find a substantive difference in slant perception for the two orthogonal 1-D texture cues, whereas most studies report that the perspective cue (vertical grating) is stronger than the compression cue (horizontal grating). The findings are consistent, however, with computational models of slant from texture (Witkin 1981; Stone 1993; O'Brien 1997), in which both types of 1-D cue present a single further unknown (eg distance to the centre of the plane), which if resolved can make the slant of a 1-D texture surface as easy to calculate as that of a 2-D textured surface.

The experiments reported above were not intended to test a particular model of depth cue combination, but, in terms of Clark and Yuille (1990), our results suggest a case of cue veto, while in terms of the weighted linear combination of the modified weak fusion model (Maloney and Landy 1989), the weight of the motion cue is much lower than would be expected if cues are dynamically re-weighted in terms of the reliability of the information available from each cue.

Measuring the perception of planar surface slant as a function of the disparity between motion and texture cues has proved effective in determining reliable weights for the two cues for regular and irregular textures. Since the effect of cue weighting was so pronounced, the high weighting of the texture cue was evident with only a small disparity between the slant portrayed by texture cues and that portrayed by motion cues. The limitation of perturbation analysis is that cues should be only slightly disparate, otherwise the fact that the cues indicate different slants becomes salient. This limitation did not hinder the experiments reported above.

It has been established that when texture and motion cues provide equivalent information about surface slant, the addition of the motion component has no effect on slant discrimination thresholds, but a 2-D motion component added to a 2-D texture can increase the perceived slant angle by approximately 1°. When there is a disparity between the slant depicted by texture and by motion cues, perceived slant is almost identical to the texture slant. This is true even for a highly irregular texture pattern. This result cannot, however, be explained by the difference in discrimination thresholds for texture and speed gradients.

Acknowledgements. The first author was a recipient of a grant from the Biotechnology and Biological Sciences Research Council (UK) and British Aerospace plc.

References

Adobe Systems Inc., 1994 "Photoshop 3 Deluxe CD-ROM", Part No. 0199 8440 894 (Mountain View, CA: Adobe Systems)

Attneave F, Olson R K, 1966 "Inferences about visual mechanisms from monocular depth effects" *Psychonomic Science* **4** 133 – 134

Blostein D, Ahuja N, 1989 "Shape from texture: Integrating texture-element extraction and surface estimation" *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11** 1233 – 1253

Braunstein M L, 1968 "Motion and texture as sources of slant information" *Journal of Experimental Psychology* **78** 247 – 253

Braunstein M L, Andersen G J, 1981 "Velocity gradients and relative depth perception" *Perception & Psychophysics* **29** 145 – 155

Braunstein M L, Liter J C, Tittle J S, 1993 "Recovering three-dimensional shape from perspective translations and orthographic rotations" *Journal of Experimental Psychology: Human Perception and Performance* **19** 598 – 614

Clark J J, Yuille A L, 1990 *Data Fusion for Sensory Information Processing Systems* (Boston, MA: Kluwer)

Cutting J E, Millard R T, 1984 "Three gradients and the perception of flat and curved surfaces" *Journal of Experimental Psychology: General* **113** 198 – 216

Epstein W, Bontrager H, Park J, 1962 "The induction of nonveridical slant and the perception of shape" *Journal of Experimental Psychology* **63** 472 – 479

Farber J M, McConkie A B, 1979 "Optical motion as information for unsigned depth" *Journal of Experimental Psychology: Human Perception and Performance* **5** 494 – 50

Flach J M, Hagen B A, Larish J F, 1992 "Active regulation of altitude as a function of optical texture" *Perception & Psychophysics* **51** 557 – 568

Freeman T C A, Harris M G, Meese T S, 1996 "On the relationship between deformation and perceived surface slant" *Vision Research* **36** 312 – 322

Gibson J J, 1950 "The perception of visual surfaces" *American Journal of Psychology* **63** 367 – 384

Gillam B J, 1968 "Perception of slant when perspective and stereopsis conflict: experiments with aniseikonic lenses" *Journal of Experimental Psychology* **78** 299 – 305

Gillam B J, 1970 "Judgements of slant on the basis of foreshortening" *Scandinavian Journal of Psychology* **11** 31 – 34

Glassner A S, 1989 *An Introduction to Ray Tracing* (San Diego, CA: Academic Press)

Gruber H E, Clark W C, 1956 "Perception of slanted surfaces" *Perceptual and Motor Skills* **6** 97 – 106

Koenderink J J, Doorn A J van, 1976 "Local structure of movement parallax of the plane" *Journal of the Optical Society of America* **66** 717 – 723

Landy M S, Maloney L T, Johnston E B, Young M J, 1991 "In defense of weak fusion: Measurement and modelling of depth cue combination", in *Mathematical Studies in Perception and Cognition* (New York: New York University Press) pp 91 – 93

Landy M S, Maloney L T, Johnston E B, Young M, 1995 "Measurement and modeling of depth cue combination: In defense of weak fusion" *Vision Research* **35** 389 – 412

Malik J, Rosenholtz R, 1994 "Computing local surface orientation and shape from texture for curved surfaces", report UCB/CSD 93/775, Computer Science Division, University of California, Berkeley, CA

Maloney L T, Landy M S, 1989 "Psychophysical estimation of the human depth combination rule", in *Visual Communication and Image Processing IV* Ed. W A Pearlman *Proceedings of the SPIE* **1199** 1154 – 1163

O'Brien J M D, 1997 *Texture and Motion Cues in Slant Perception* PhD thesis, University College London, London

Rogers B, Graham M, 1979 "Motion parallax as an independent cue for depth perception" *Perception* **8** 125 – 134

Rosinski R R, 1974 "On the ambiguity of visual stimulation: A reply to Eriksson" *Perception & Psychophysics* **16** 259 – 263

Saidpour A, Andersen G J, Braunstein M L, 1997 "Effect of texture and viewing condition on viewer-centred and object-centred judgments of scene depth" *Investigative Ophthalmology & Visual Science* **38**(4) S1010

Smith A H, 1956 "Gradients of outline convergence and distortion as stimuli for slant" *Canadian Journal of Psychology* **10** 211 – 218

Stone J V, 1993 "Shape from local and global analysis of texture" *Philosophical Transactions of the Royal Society of London, Series B* **339** 53 – 65

Vickers D, 1971 "Perceptual economy and the impression of visual depth" *Perception & Psychophysics* **10** 23 – 27

Watt A, 1991 *Understanding Vision* (London: Academic Press)

Wilcox R R, 1995 "Anova—a paradigm for low-power and misleading measures of effect size" *Review of Educational Research* **65** 55 – 57

Witkin A P, 1981 "Recovering surface shape and orientation from texture" *Artificial Intelligence* **17** 17 – 45

Young M J, Landy M S, Maloney L T, 1993 "A perturbation analysis of depth perception from combinations of texture and motion cues" *Vision Research* **33** 2685 – 2696