Vision Res. Vol. 32, No. 3, pp. 587–590, 1992
Printed in Great Britain.

0042-6989/92 \$5.00 + 0.00

Research Note

Retinal Speed Gradients and the Perception of Surface Slant

MIKE HARRIS,*† TOM FREEMAN,* JIM HUGHES*

Received 17 April 1991

Previous work has demonstrated a difference in human sensitivity to compressive and shearing speed gradients. This raises the possibility that the ability to estimate the slant of a surface may vary with its direction of tilt. No such variance was found here, which may indicate that slant estimation depends upon deformation rather than upon compression or shear.

INTRODUCTION

Movements of the observer produce smooth transformations of the retinal image which contain useful information about the three-dimensional layout of the world (e.g. Helmholtz, 1925; Gibson, 1950). In principle, spatial gradients in the speed of retinal motion encode the direction of tilt and the degree of slant of object surfaces (e.g. Koenderink, 1986; Koenderink & van Doorn, 1976; Longuet-Higgins & Prazdny, 1980). In practice, a number of studies have shown that human observers can make use of these speed gradients in judging the shape of three dimensional objects (e.g. Braunstein & Andersen, 1984; Hildreth, Grzywacz, Adeslon & Inada, 1990; Rogers & Graham, 1979; Treue, Husain & Andersen, 1991), distinguishing between convex and concave edges (e.g. Braunstein & Andersen, 1981; Braunstein & Tittle, 1988; Farber & McConkie, 1979), and in estimating surface slant (e.g. Braunstein, 1968).

Studies of perceived surface slant have generally combined horizontal translations of the observer with a vertical direction of surface tilt because, under these circumstances, it is easy to manipulate speed and texture gradients independently (e.g. Braunstein, 1968). However, for such stimuli, speed varies only at right angles to the direction of motion (producing a shearing gradient) whereas, in natural retinal images, speed usually varies both along and at right angles to the direction of motion (producing compressive and shearing gradients, respectively). This may be important because there is some evidence of a difference in the perception of these different types of gradient. Using a motion analogue of the Craik–O'Brien–Cornsweet stimulus, Rogers and Graham (1983) found that compressive gradients produced a greater illusion of surfaces at different depths than did shearing gradients, suggesting that sensitivity to shallow gradients is greater for shears than for compressions. Indeed, Rogers and Graham point out that sensitivity does fall off less rapidly at low spatial frequencies for shears than for compressions.

Since information about surface slant is usually carried by rather shallow speed gradients, we might also expect the ability to estimate surface slant to vary with the direction of tilt. In fact, the existence of a difference in sensitivity suggests a simple way to address the more general issue of whether and how the human visual system makes use of the differential invariants *div*, *curl* and *def* in encoding retinal flow (Koenderink & van Doorn, 1976). *Div* measures local change in area and is sensitive only to speed gradients along the direction of motion (i.e. compressions). Conversely, *curl* measures local rotation and is sensitive only to speed gradients at right angles to the direction of motion (i.e. shears). *Def* measures local change in shape without change in area and is sensitive to both types of gradient.

For the simple cases used here, these differential invariants can be adequately understood in terms of a vector, **F**, describing the direction of tilt and amplitude of slant of the surface relative to the line of sight, and a vector, $\mathbf{A}_T$, describing the direction and speed of motion at right angles to the line of sight [see Koenderink (1986) for an elegant demonstration and Koenderink (1985) for a complete mathematical account]:

$$Div = -\mathbf{F}\cdot\mathbf{A}_T$$

$$Curl = -\mathbf{F}\times\mathbf{A}_T$$

$$Def = FA_T.$$

The axis of compression of the *def* component bisects **F** and $\mathbf{A}_T$. For horizontal movement of the observer and the small surfaces used here, $\mathbf{A}_T$ is everywhere approximately horizontal. When the surface is tilted vertically,

*Cognitive Science Research Centre, School of Psychology, The University of Birmingham, Birmingham B15 2TT, England.
†To whom all correspondence should be addressed.

F is everywhere approximately vertical and the resultant shearing flow pattern contains substantial *curl*, no *div*, and *def* with the axis of compression oriented diagonally. When the surface is tilted horizontally, **F** is everywhere approximately horizontal and the resultant compressive flow pattern contains substantial *div*, no *curl*, and the same amplitude of *def* with the axis of compression oriented horizontally. Thus, a simple comparison of horizontal and vertical tilt may reveal which, if any, of these differential invariants is important to the task of estimating surface slant.

METHOD

Stimuli

The stimulus consisted of a random pattern of 256 bright dots generated by DEC LSI 11/23 computer and displayed via a CED 501 laboratory interface upon the dark screen of an oscilloscope (Hewlett–Packard 2161, P31 phosphor) at a frame rate of 60 Hz. In order to control for the effects of shape and texture gradient, two main conditions were used. In the first, "full cues", condition, the dots were evenly distributed within a circle upon a virtual surface at right angles to the line of sight. The virtual surface was then rotated as required and the dots were then projected into the image so that the resulting stimulus preserved both the texture gradients and the distortion of shape produced by the rotation. In the second, "reduced cues" condition, the dots were evenly disributed within a circle on the screen and then projected back onto a virtual surface which had already been rotated to the required degree. This condition thus contained neither static texture nor static shape cues although, in the "gradient" condition (see

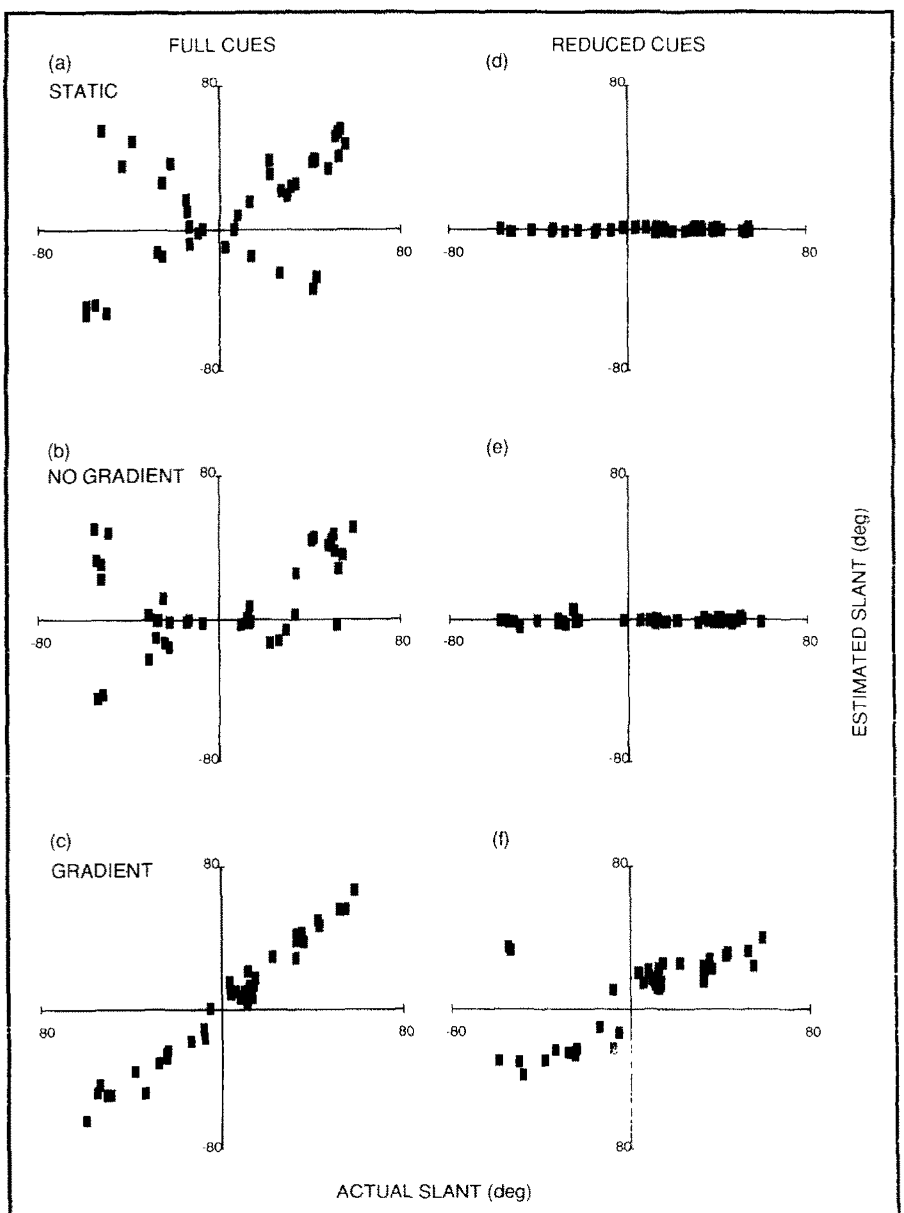

FIGURE 1. Slant estimates for observer MGH for horizontal movement and a vertical direction of tilt. Full texture and shape cues: (a) static, (b) no gradient and (c) gradient. Reduced texture and shape cues: (d) static, (e) no gradient and (f) gradient.

below), both texture and shape varied during stimulus motion.

Three motion conditions were investigated. In the first "gradient", condition, the position of each dot in virtual space was shifted horizontal between frames and then projected back into the image. This condition thus preserved the speed gradients which would normally be produced by horizontal movements of a slanting surface or, equivalently, of the observer's viewpoint. In the second, "no gradient" condition, each dot was shifted horizontally by the same amount between frames, irrespective of its position in virtual space. This condition contained the same average amount of motion but contained no speed gradients. In the third, "static" condition, a single frame was randomly chosen from the "gradient" condition and repeatedly displayed.

All stimuli were viewed monocularly in a darkened room from a distance of 57 cm. The observer's head was supported by a chin rest. The centre of the virtual surface was located 285 cm from the viewpoint and, when at right angles to the line of sight, subtended 5 deg. In the moving conditions, the virtual surface moved sinusoidally at 1 c/sec through a total horizontal distance of 28.5 cm. A stationary dot in the centre of the display was provided as a fixation point, but observers were not instructed to adopt any particular fixation strategy.

Procedure

The three authors acted as observers. Each undertook a total of 24 sessions consisting of two replications of the 6 conditions (2 types of cues by 3 types of motion) for vertical and horizontal directions of surface tilt. Each session investigated a single condition and consisted of 5 practice and 20 experimental trials. The observers undertook the sessions in different random orders. On each trial the virtual surface was rotated through a random amount up to 60 deg from the image plane. The observer's task was to estimate this slant by adjusting the angle of a straight line subtending 3 deg located 7.5 deg below the centre of the display. The stimulus was presented continuously until the observer had made an estimate and informed the computer by pressing a button.

RESULTS

All of the observers reported that conditions involving speed gradients provided an immediate and compelling impression of three-dimensionality, akin to that produced by a random dot stereogram (see also Rogers & Graham, 1979). For conditions which did not include speed gradients but which preserved texture gradients and shape cues, it was possible to infer three-dimensionality, but only indirectly. These impressions were born out by the data. Figure 1 shows typical settings made by one of the observers for the 6 conditions with a vertical direction of tilt.

As expected, with reduced cues there was no impression of surface slant either in the static [Fig. 1(d)] or the no gradient [Fig. 1(e)] condition, since these stimuli contain no cues to surface layout. For the gradient condition in the absence of static texture and shape cues, surface slant could be consistently judged but tended to be underestimated [Fig. 1(f)]. With full cues and static presentation [Fig. 1(a)] surface slant could be inferred fairly accurately but observers sometimes mistook the polarity of the slant, suggesting that they were making more use of shape than texture cues. Performance was not improved simply by the addition of retinal motion, as demonstrated by the no gradient condition [Fig. 1(b)], although it remains possible that any improvement due to motion is cancelled by the accompanying increased cues to surface flatness. However, when speed gradients were added, slope was estimated very accurately and without ambiguity [Fig. 1(c)].

The full data are summarized in Fig. 2 which shows, separately for each condition, the mean slopes of the three observers derived from linear regression. The regressions were performed both with the raw data (Fig. 2, solid symbols) and with their absolute values (open symbols). This manipulation had no great effect upon the conditions including speed gradients but the absolute values yielded substantially better slant estimates in the static and no gradient conditions with full texture and shape cues. This improvement reflects the perceived ambiguity of slant polarity and demonstrates that, even though observers report that the degree of slant can only be indirectly inferred under these conditions, the inference is fairly accurate.

DISCUSSION

Speed gradients in the absence of static texture and shape cues yield a consistent and compelling impression of surface slant. As has been suggested by other authors (e.g. Braunstein, 1968), the fact that slant tends to be systematically underestimated under these conditions may be due to the confounding cues to flatness provided

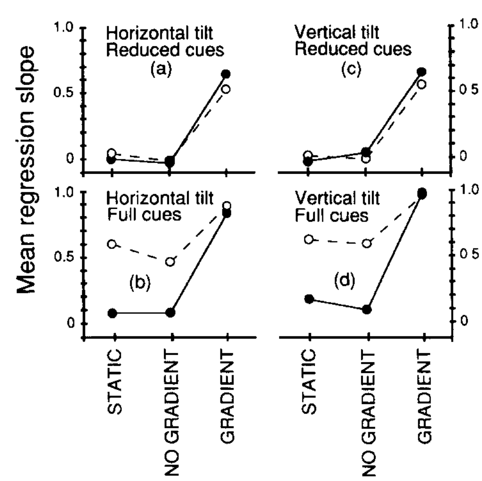

FIGURE 2. Mean slopes for 3 observers of slant estimates, derived from linear regression on the type of data depicted in Fig. 1. Vertical tilt: (a) reduced texture and shape cues and (b) full texture and shape cues. Horizontal tilt: (c) reduced cues and (d) full cues. Solid symbols: raw data. Open symbols: absolute values.

by the uniform texture and undistorted shape of the stimulus. Static texture gradients and shape distortions allow slant to be inferred fairly accurately but observers frequently mistake the polarity of the slant. Motion of the stimulus does not, by itself, improve performance but, when texture and shape are combined with speed gradients, the impression of surface slant is compelling, unambiguous, and very accurate.

Despite the findings of Rogers and Graham (1983), compression and shear are equally effective cues to surface slant. In view of the greater sensitivity to shallow shearing gradients than to compressive gradients, we might have expected estimates of surface slant to be better for vertical than for horizontal tilt. However, even for very small slants, where the sensitivity difference should be greatest and sensitivity should limit performance, the graphs of the raw data (e.g. Fig. 1) do not suggest any systematic difference in either the slopes or the variability of the estimates. It thus seems that speed gradients both along and at right angles to the direction of motion are equally effective in the judgement of surface slant, which in turn suggests that *def* is more likely to be used in this task than either *div* or *curl*. The findings of Rogers and Graham (1983) suggest that these latter two differential invariants may be more important in segmenting complex flow fields on the basis of discontinuities in the speed of motion.

REFERENCES

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 and Psychophysics*, *29*, 145–155.

Braunstein, M. L. & Andersen, G. J. (1984). Shape and depth perception for parallel projections of three-dimensional motion. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 749–760.

Braunstein, M. L. & Tittle, J. S. (1988). The observer-relative velocity field as the basis for effective motion parallax. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 582–590.

Farber, J. M. & McConkie, A. B. (1979). Optical motions as information for unsigned depth. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 494–500.

Gibson, J. J. (1950). *The perception of the visual world.* Boston: Houghton Mifflin.

Helmhotz, H. (1925). *Physiological optics* (Vol. 3). Rochester, N.Y.: Optical Society of America.

Hildreth, E. C., Grzywacz, N. M., Adeslon, E. H. & Inada, V. K. (1990). The perceptual buildup of three-dimensional structure from motion. *Perception and Psychophysics*, *48*, 19–36.

Koenderink, J. J. (1985). Space, form and optical deformations. In Ingle, D. T., Jeanerod, M. & Lee, D. M. (Eds), *Brain mechanisms and spatial vision.* Dordrecht: Martinus Nijhoff.

Koenderink, J. J. (1986). Optic flow. *Vision Research*, *26*, 161–180.

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

Longuet-Higgins, H. C. & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London, Series B*, *208*, 385–397.

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

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

Treue, S., Husain, M. & Andersen, R. A. (1991). Human perception of structure from motion. *Vision Research*, *31*, 59–75.

Acknowledgement—This work was partly supported by a SERC Research Studentship to Jim Hughes.