The visual perception of surface orientation from optical motion

JAMES T. TODD
Ohio State University, Columbus, Ohio

and

VICTOR J. PEROTTI
Rochester Institute of Technology, Rochester, New York

Observers viewed monocular animations of rotating dihedral angles and were required to indicate their perceived structures by adjusting the magnitude and orientation of a stereoscopic dihedral angle. The motion displays were created by directly manipulating various aspects of the image velocity field, including the mean translation, the horizontal and vertical velocity gradients, and the manner in which these gradients changed over time. The adjusted orientation of each planar facet was decomposed into components of slant and tilt. Although the tilt component was estimated with a high degree of accuracy, the judgments of slant exhibited large systematic errors. The magnitude of perceived slant was determined primarily by the magnitude of the velocity gradient scaled by its direction. The results also indicate that higher order temporal derivatives of the moving elements had little effect on observers' judgments.

A fundamental issue in the theoretical analysis of three-dimensional (3-D) structure from motion concerns the number of distinct views that are required for different types of perceptual judgments. Whereas the first-order relations between pairs of views provide sufficient information to distinguish rigid motion from nonrigid motion, they are inherently ambiguous with respect to an object's 3-D structure. Ullman (1977, 1979, 1983) proved that an arbitrary two-frame motion sequence under parallel projection has an infinite one-parameter family of possible 3-D interpretations (see also Bennett, Hoffman, Nicola, & Prakash, 1989; Koenderink & van Doorn, 1991). In order to obtain a unique computation of euclidean metric structure, the motion sequence must contain a minimum of three distinct views of at least four points. These theoretical limits define an absolute upper bound on what can be computed from pure motion information—even for an ideal observer who can measure the projected position of each point and perform necessary mathematical operations with perfect accuracy.

There is a growing body of evidence to indicate, however, that observer sensitivity to higher order relations among three or more views is extremely imprecise (e.g.,

Snowden & Braddick, 1991; Todd, 1981; Werkhoven, Snippe, & Toet, 1992) and that the visual perception of 3-D structure from motion is based primarily on first-order temporal derivatives of moving images (e.g., Liter, Braunstein, & Hoffman, 1993; Perotti, Todd, Lappin, & Phillips, 1998; Perotti, Todd, & Norman, 1996; Todd & Bressan, 1990; Todd & Norman, 1991). Because of the inherent theoretical limitations of this information, observers typically exhibit large errors in judgments of euclidean metric structure from motion (e.g., Braunstein, Liter, & Tittle, 1993; Liter & Braunstein, 1998; Liter et al., 1993; Perotti et al., 1998; Todd & Norman, 1991), and they have difficulty discriminating different structures within the one-parameter family even when a motion sequence contains more than two distinct frames (e.g., Eagle & Blake, 1995; Todd & Bressan, 1990; Todd & Norman, 1991). There is, on the other hand, another important aspect of the available data that is not easily explained by the mathematical ambiguities of pure velocity information: When observers are asked to estimate 3-D structures of moving objects, their judgments are often highly reliable even though they may exhibit systematic biases (Braunstein et al., 1993; Liter & Braunstein, 1998; Liter et al., 1993; Todd & Norman, 1991). If the available information is infinitely ambiguous, then why should an object appear to have any specific structure at all? To the extent that it does, there would have to be some other constraint or heuristic at work to restrict the set of possible perceptual interpretations.

The research described in the present article was designed to investigate the specific aspects of image motion that determine the perceived orientation of a planar surface patch rotating in depth under orthographic projection.
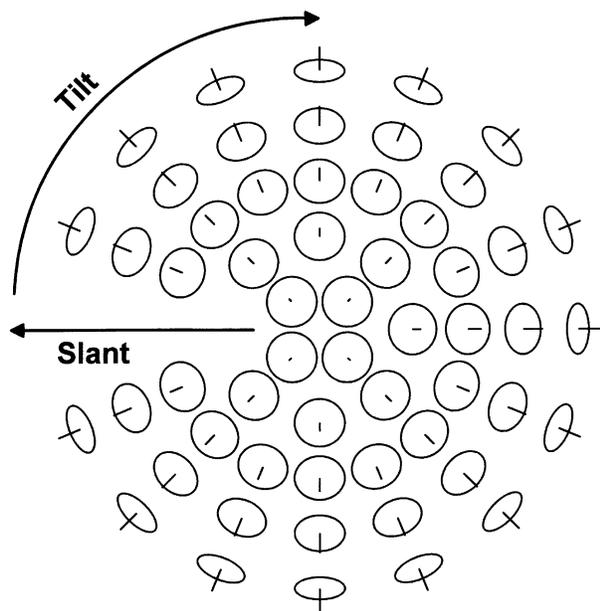

Figure 1. A set of planar patches arranged on a sphere to illustrate the slant and tilt components of surface orientation.

A planar surface in 3-D space can be mathematically specified by the following equation:

$$Z(x,y) = Z_0 + x \sin \tau \tan \sigma + y \cos \tau \tan \sigma, \quad (1)$$

where σ (slant) is the angle between the surface normal and the line of sight, τ (tilt) is the projected orientation of the surface normal in the image plane relative to the vertical, and Z_0 is the distance from the surface to the image plane along the line of sight (see Figure 1). Equation 1 defines the depth (Z) relative to the image plane of any surface point as a simple function of its horizontal position (X) and its vertical position (Y).

Ullman (1977) has shown that all rigid rotations of an object under orthographic projection can be transformed mathematically to the special case of rotation about a vertical axis in the image plane (see also Todd & Bressan, 1990). If a planar surface is rotated in this manner, the image velocity (V) of each point is given by the following equation:

$$V(x,y) = \omega (Z_0 + x \sin \tau \tan \sigma + y \cos \tau \tan \sigma), \quad (2)$$

where ω is its angular velocity in 3-D space. Gradients of the image velocity field are obtained from its partial spatial derivatives in the horizontal and vertical directions. These can be expressed as follows, where subscripts indicate the direction of differentiation:

$$V_X = \omega \sin \tau \tan \sigma, \quad (3)$$

$$V_Y = \omega \cos \tau \tan \sigma. \quad (4)$$

With appropriate rearrangements of these equations,

$$\tan \tau = \frac{V_X}{V_Y}, \quad (5)$$

$$\tan \sigma = \frac{\sqrt{V_x^2 + V_y^2}}{\omega}, \quad (6)$$

it is possible to show that the tilt component of orientation is uniquely specified by the ratio of two velocity gradients, but that the slant component is underdetermined. In order to obtain an accurate estimate of slant from pure velocity information, it would be necessary to know the angular velocity ω.

One physiologically plausible way of measuring image velocity gradients has been suggested by Koenderink and van Doorn (1977). Their idea is to approximate the local differential structure of a surface by taking several closely spaced velocity measurements, which are then summed with a predefined weighting function. The power of this kind of operator is that the individual velocities can be combined in a variety of ways to assess different aspects of image motion in a local region. For example, Figure 2 shows how the mean translation (V_0) and the horizontal and vertical velocity gradients (V_X and V_Y) can be measured from the outputs of four velocity detectors (V_1 through V_4) that are spatially separated by D_X and D_Y, respectively, in the horizontal and vertical directions.

How might these measures of local image velocity be used to produce *reliable* estimates of surface orientation? Domini, Caudek, and Gerbino (1995) and Domini and Caudek (1999) have recently proposed that judgments of slant are based primarily on the local pattern of deformation (def), where

$$\mathrm{def} = \sqrt{V_x^2 + V_y^2}. \quad (7)$$

It is important to note that def also appears in the numerator of Equation 6. If observers adopted some default value of ω in order to estimate local orientation, then judgments of slant should increase monotonically with def. Although this strategy could produce judgments that are reliable, they would not necessarily be accurate, depending on how much the true angular velocity deviates from its assumed value.

In an effort to test this hypothesis in the present experiments, we examined the effects of several possible measures of image motion, including def, to evaluate their relative importance for the perception of surface orientation. In Experiment 1, we employed displays called *constant flow fields* to eliminate higher order information from temporal variations of V_X and V_Y (see Perotti et al., 1998; Perotti et al., 1996). The instantaneous velocities of these displays were mathematically indistinguishable from real object motion, but their higher order temporal derivatives did not provide a unique rigid interpretation. To confirm the generality of this approach, we also performed a second experiment using computer simulations
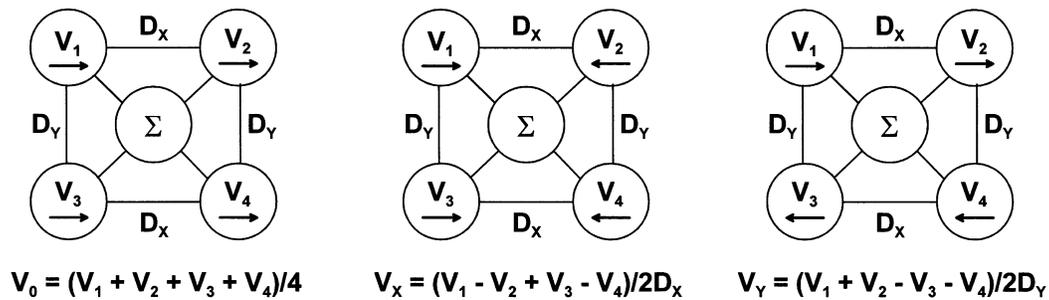
$$V_0 = (V_1 + V_2 + V_3 + V_4)/4 \qquad V_X = (V_1 - V_2 + V_3 - V_4)/2D_X \qquad V_Y = (V_1 + V_2 - V_3 - V_4)/2D_Y$$

Figure 2. A schematic diagram of three local differential operators for measuring different aspects of local image motion.

of real rotating objects whose velocity fields changed over time.

EXPERIMENT 1

Method

Apparatus. The optical patterns were created and displayed on a Silicon Graphics Crimson VGXT workstation with hardware texture-mapping capabilities. Stereoscopic viewing hardware was also used. The stereoscopic half-images were presented using LCD (liquid crystal) shuttered glasses that were synchronized with the monitor's refresh rate. The left and right views of a stereo pair were displayed at the same position on the monitor screen, but they were temporally offset. The left and right lenses of the LCD glasses shuttered synchronously with the display so that each view of the stereo pair was seen only by the appropriate eye. The CRT was refreshed at 120 Hz. Thus, each view of a stereoscopic half-image was updated at half that, or 60 Hz. The viewing distance was 57 cm, such that the 1,280 pixel wide $\times$ 1,024 pixel high display screen subtended 35.2º $\times$ 28.2º of visual angle.

Stimuli. All stimuli appeared as dihedral angles, here defined as two planes meeting at a horizontal line across the center of the display screen. Each plane was covered with a rocky texture pattern, using a process that is defined below. To prevent the use of a bounding contour as a possible source of information, the edges of all stimuli were occluded by a rectangular aperture whose horizontal and vertical dimensions were 800 and 640 pixels, respectively (i.e., 22.0º $\times$ 17.6º). Figure 3 shows a pair of images from a typical apparent motion sequence, which can be viewed stereoscopically to reveal the appearance of a dihedral angle.

To create a continuously moving texture pattern, the image velocities at the corners of the rectangular aperture (V_1 and V_2) and at the midpoints of its left and right boundaries (V_3 and V_4) were used
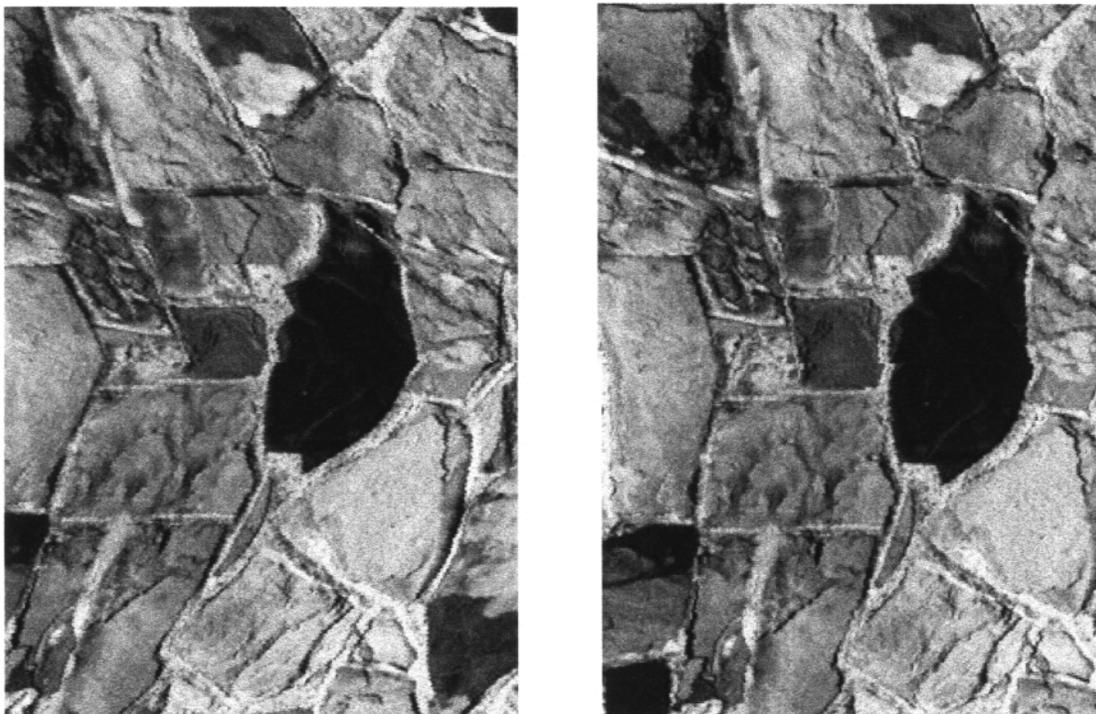

Figure 3. A stereogram of a textured dihedral angle similar to those used in the present experiments.

time 1 I━━━━━━━━━━━━━━▶ time 2
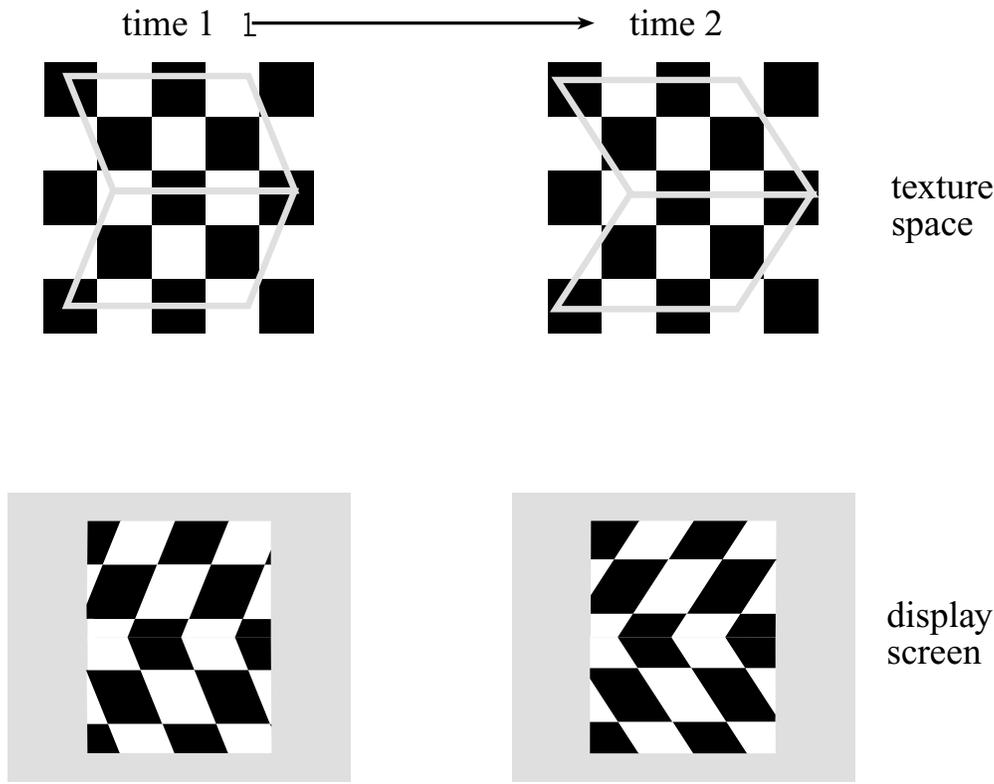

texture space

display screen

Figure 4. A schematic diagram of the texture-mapping process used to create the patterns of motion in Experiment 1.

to define the deformation of a chevron in texture space. Image velocities V_1 and V_2 controlled the motion of the top left, top right, bottom left, and bottom right corners of the chevron, and V_3 and V_4 controlled the motion at the two middle corners. Figure 4 gives a schematic illustration of the texturing and display process. In any given frame of a movement sequence, the texture within the chevron was mapped into the viewing window of the display screen through linear interpolation. In the next subsequent frame, the corners of the chevron in texture space were displaced horizontally by the values assigned for V_1 through V_4, which transformed the pattern in the viewing window.

Each animation sequence was composed of 24 distinct frames that oscillated back and forth in continuous alternation at a rate of 20 frames per second. The displays included 26 different conditions with varying combinations of V_0, V_X, and V_Y, which are described in Table 1.

The image velocities V_1 through V_4 for each display were computed from these parameters using the following equations:

$$V_1 = V_0 + \frac{D_X V_X}{2} + \frac{D_Y V_Y}{2}, \tag{8}$$

$$V_2 = V_0 - \frac{D_X V_X}{2} + \frac{D_Y V_Y}{2}, \tag{9}$$

$$V_3 = V_0 + \frac{D_X V_X}{2} - \frac{D_Y V_Y}{2}, \tag{10}$$

$$V_4 = V_0 - \frac{D_X V_X}{2} - \frac{D_Y V_Y}{2}, \tag{11}$$

where D_X and D_Y are the horizontal and vertical distances between the corners of the chevron (i.e., $D_X = 22°$, and $D_Y = 8.8°$).

Note that Table 1 lists a simulated tilt for each condition but no simulated slant. Because the displacements in texture space were identical at each frame transition, the resulting patterns of motion in image space defined a constant flow field (see Perotti et al., 1998; Perotti et al., 1996). Although the instantaneous pattern of velocities was mathematically indistinguishable from a rotating dihedral angle, the higher order temporal derivatives had no possible rigid interpretation as an object rotating in depth. It is interesting to point out in this regard that displays with nonzero values of V_0 could technically be interpreted as dihedral angles undergoing perspective translation (see Liter & Braunstein, 1998). Assuming that the moving surfaces were at the same distance (Z_0) as the plane of the display screen, then a simulated slant could be computed from the following equation:

$$\tan \sigma = \frac{Z_0 \text{def}}{V_0}. \tag{12}$$

It is also important to recognize, however, that the values of V_0 we employed were much too small relative to the values of V_X and V_Y for the displays to be plausibly interpreted as translating surfaces. When slant is computed from Equation 12 using the parameters in Table 1, almost all of the conditions had simulated slants in excess of 80°. Although this interpretation was mathematically possible, all of the observers reported that the depicted surfaces appeared perceptually to be rotating in depth.

Procedure. At the beginning of each trial, the observers were presented with a motion display and were asked to estimate the

Table 1
A Summary of Display Parameters for
the 26 Conditions of Experiment 1

Condition	V_0 (deg/sec)	V_X (1/sec)	V_Y (1/sec)	Tilt (deg)
1	0.505	0.000	0.000	
2	0.500	0.021	0.000	90.00
3	0.484	0.029	0.000	90.00
4	0.452	0.036	0.000	90.00
5	0.391	0.041	0.000	90.00
6	0.000	0.046	0.000	90.00
7	0.500	0.000	0.051	0.00
8	0.495	0.021	0.051	21.80
9	0.478	0.029	0.051	29.50
10	0.445	0.036	0.051	34.72
11	0.380	0.041	0.051	38.66
12	0.484	0.000	0.073	0.00
13	0.478	0.021	0.073	15.79
14	0.459	0.029	0.073	21.80
15	0.421	0.036	0.073	26.10
16	0.338	0.041	0.073	29.50
17	0.452	0.000	0.089	0.00
18	0.445	0.021	0.089	13.00
19	0.421	0.029	0.089	18.09
20	0.368	0.036	0.089	21.80
21	0.000	0.041	0.089	24.79
22	0.391	0.000	0.103	0.00
23	0.380	0.021	0.103	11.31
24	0.338	0.029	0.103	15.79
25	0.000	0.036	0.103	19.11
26	0.000	0.000	0.115	0.00

magnitude and orientation of the depicted dihedral angle at the center of the rotation sequence. Once they had a clear sense of the perceived structure, they were instructed to press a mouse button that replaced the moving display with a stereogram, which they were required to adjust so that it matched the appearance of the moving dihedral angle. Vertical movements of the mouse controlled the angle between the two planes; horizontal movements controlled the degree of rotation about a vertical axis. Once an appropriate setting was obtained, the display could be toggled back into a motion pattern to reexamine the original stimulus. The observers were allowed to toggle back and forth as many times as was necessary until they were satisfied with their judgments, which they indicated by pressing a different mouse button. The observers were also instructed to close one eye while examining the motion sequences in order to enhance the 3-D appearance. The 26 display conditions were presented five times each in a random sequence over a period of two experimental sessions.

Observers. Judgments were obtained for 4 different observers, including the 2 authors and 2 other naive observers who were unfamiliar with the purpose of the experiment or how the displays had been generated. Each observer had corrected-to-normal vision.

Results

From the observers' judgments on each trial, we computed the adjusted slant and tilt of each planar facet to estimate the perceived orientations of the moving surfaces. In order to assess the reliability of these judgments, we computed the standard deviation of the repeated trials in each condition. The results of this analysis are presented in Table 2, which shows the average deviation of judged slant and tilt collapsed over conditions for each observer. In general, the average spread of the observers' judgments

in both slant and tilt was approximately 4°. We also evaluated the variations between observers by correlating their responses across the different conditions. The results of this analysis are shown in Table 3. Note in the table that the correlation coefficients for each possible pair of observers were all above .93, thus indicating that there was a high degree of consistency in their judgments.

Additional correlations were performed to evaluate how the observers' perceptions were scaled by the different parameters of image motion. Let us first consider their judgments of the tilt component of surface orientation, which is mathematically specified in the first-order pattern of image velocities (see Equation 5). Domini et al. (1995) and Domini and Caudek (1999) have reported that observers' tilt judgments for rotating planar surfaces are almost perfectly accurate, and that result is confirmed by the present experiment. Figure 5 shows the average adjusted tilt plotted as a function of the simulated tilt in all of the conditions collapsed over observers. A regression analysis of these data revealed that the adjusted and simulated tilts were almost perfectly correlated ($r = .99$).

Because the higher order temporal derivatives of our constant flow fields had no possible rigid interpretation, there is no "correct" value of simulated slant with which we can compare the observers' judgments. We can, however, examine their performance with respect to the hypothesis of Domini and Caudek (1999) that the magnitude of perceived slant should vary proportionally with def. Figure 6 shows the average adjusted slants, collapsed over the 4 observers, plotted as a function of def for all 26 conditions. Although the correlation of judged slant with def is relatively high ($r = .92$), it is clear from the graph that the residuals of this analysis are not randomly distributed. To better understand the regularities of these data, it is useful to perform a more detailed examination of how various components of the velocity field influenced the observers' judgments. Each symbol shape in Figure 6 represents a different possible value of the vertical velocity gradient (V_Y), and the shading of these symbols represents the horizontal velocity gradient (V_X)—that is, the darkness of the symbols increases with the magnitude of V_X. In conditions where $V_Y = 0$ (represented by circles), the magnitude of perceived slant varied positively with V_X. However, in all of the other conditions with nonzero values of V_Y, increasing the magnitude of V_X had a negative effect on perceived slant.

From an ecological point of view, this result is quite surprising. Other things being equal, increasing the slant of a rotating surface will produce a corresponding increase

Table 2
The Average Deviations Between Repeated Trials of
the Same Condition for the 4 Observers of Experiment 1

	J.T.	V.P.	J.N.	J.S.
Tilt	2.95	3.30	4.45	5.79
Slant	3.49	4.32	4.10	5.46

Table 3
The Correlations Between Observers in
Experiment 1 for Adjusted Slant and Tilt

	J.T.–V.P.	J.T.–J.N.	J.T.–J.S.	V.P.–J.N.	V.P.–J.S.	J.N.–J.S.
Tilt	.99	.97	.98	.95	.99	.95
Slant	.96	.96	.93	.97	.93	.97

in the magnitudes of both V_X and V_Y (see Equations 3 and 4), and this is also true for translating surfaces under perspective projection (see Equation 12). Why then should an increase in V_X lead to a reduction in perceived slant? Although we were initially skeptical of this finding, our confidence was bolstered when we recognized that the same result has been obtained in several previous studies—albeit in a somewhat different guise. In one such study by Braunstein et al. (1993), moving dihedral angles were presented with varying values of the horizontal gradient (they used the term *compression*) combined with fixed values of the vertical gradient. As the horizontal gradient was increased, there were systematic reductions in the perceived relative orientation between the two planar facets of the depicted dihedral angles. That is, there was a reduction in perceived slant. A similar result has more recently been reported by Domini and Caudek (1999). They measured the perceived orientations of rotating planar surfaces with a fixed magnitude of def and varying directions of tilt. As the direction of the velocity gradient was varied from vertical to horizontal so that the magnitude of V_X was increased relative to V_Y, the perceived slants of the surfaces were systematically lowered (see also Liter & Braunstein, 1998; Perotti, Todd, Tittle, & Norman, 1994; Tittle, Todd, Perotti, & Norman, 1995).

One possible explanation of these findings is that the relative magnitudes of perceived velocity gradients are anisotropic, such that gradients parallel to the direction of average motion are perceived to be smaller than equivalent gradients in a perpendicular direction (e.g., Cornilleau-Pérès & Droulez, 1989; Norman & Lappin, 1992; Norman & Todd, 1995). If def is the primary determinant of perceived slant, but it is computed from directionally biased gradient measures, then observers' slant judgments would be significantly influenced by surface tilt. Although this could explain the findings of Domini and Caudek (1999) for varying tilts with fixed values of def, it cannot account for the results of the present experiment or those reported by Braunstein et al. (1993). In the present experiment and in Braunstein et al.'s study, increasing the magnitude of V_X for fixed values of V_Y produced significant reductions in perceived slant, thus indicating that V_X can be weighted negatively. Moreover, the results of the present experiment also indicate that the scaling of V_X cannot be a fixed parameter in the perceptual analysis of surface orientation, since its effect can be either positive or negative depending on the value of V_Y.

If the observers' responses were not based exclusively on def, then what alternative source of information might they have used to reliably estimate the slants of the depicted surfaces? In order to fit the data from this experiment, any potential scaling function to be considered must satisfy three criteria: (1) The function must evaluate to zero when V_X and V_Y are both zero; (2) V_X must produce a positive contribution to perceived slant when V_Y is zero (i.e., when $\tau = 90°$); and (3) V_X must produce a negative contribution to perceived slant when V_Y is sufficiently greater than zero (i.e., when $\tau < 90°$). One pos-
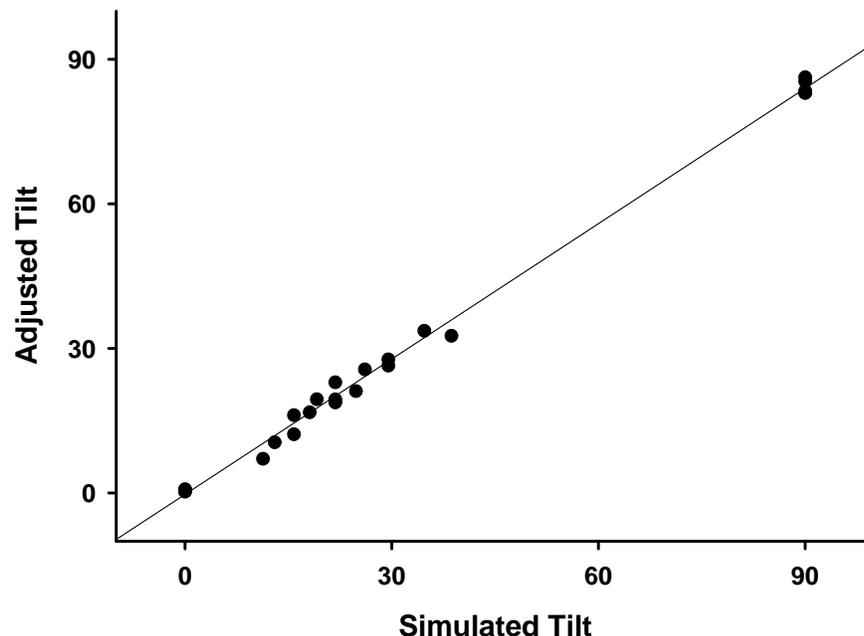

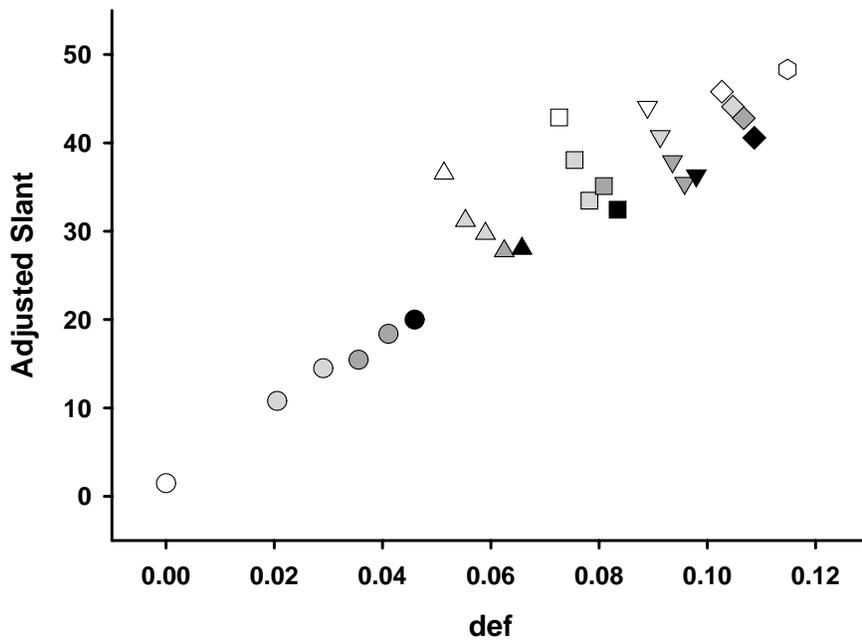Figure 5. The mean adjusted tilt in Experiment 1 plotted as a function of simulated tilt.

Figure 6. The mean adjusted slant in Experiment 1 plotted against the magnitude of def. Each symbol represents a different possible value of the vertical velocity gradient (V_Y) and the shading of those symbols represents the horizontal velocity gradient (V_X). The darkness of the symbols increases with the magnitude of V_X.
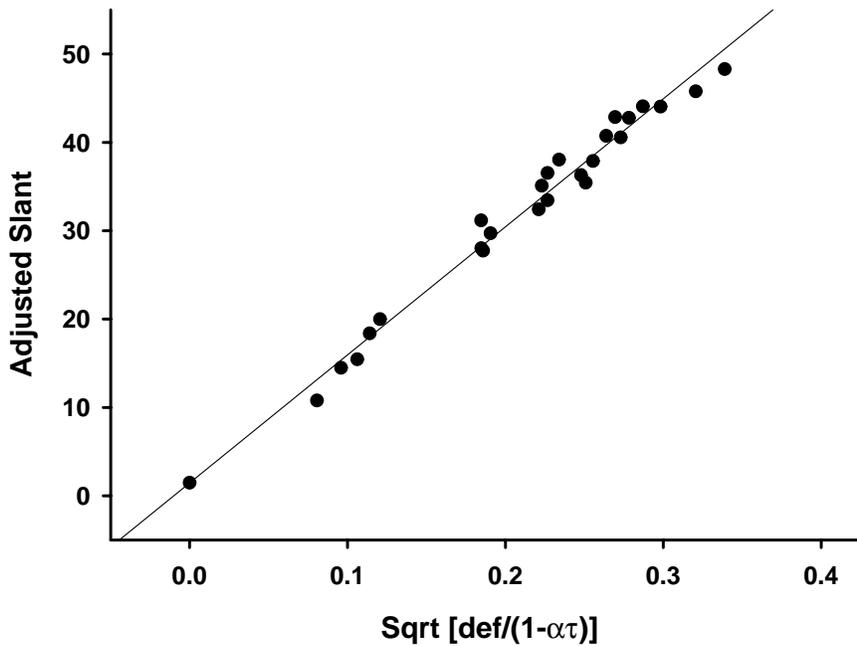
Figure 7. The mean adjusted slant in Experiment 1 plotted against the scaling function described by Equation 13.

Table 4
A Summary of Display Parameters for
the 24 Conditions of Experiment 2

Condition	V_0 (deg/sec)	V_X (1/sec)	V_Y (1/sec)	ω deg/sec	Slant (deg)	Tilt (deg)
1	0.000	0.040	0.057	3.75	46.86	34.72
2	0.000	0.040	0.057	7.50	28.08	34.72
3	0.000	0.056	0.081	3.75	56.47	34.72
4	0.000	0.056	0.081	7.50	37.04	34.72
5	0.438	0.000	0.057	3.75	41.26	0.00
6	0.438	0.000	0.057	7.50	23.68	0.00
7	0.619	0.000	0.081	3.75	51.13	0.00
8	0.619	0.000	0.081	7.50	31.81	0.00
9	0.000	0.032	0.081	3.75	53.19	21.80
10	0.000	0.032	0.081	7.50	33.74	21.80
11	0.000	0.046	0.115	3.75	62.11	21.80
12	0.000	0.046	0.115	7.50	43.37	21.80
13	0.357	0.000	0.081	3.75	51.13	0.00
14	0.357	0.000	0.081	7.50	31.81	0.00
15	0.505	0.000	0.115	3.75	60.32	0.00
16	0.505	0.000	0.115	7.50	41.26	0.00
17	0.000	0.023	0.099	3.75	57.33	13.00
18	0.000	0.023	0.099	7.50	37.94	13.00
19	0.000	0.032	0.141	3.75	65.61	13.00
20	0.000	0.032	0.141	7.50	47.79	13.00
21	0.253	0.000	0.099	3.75	56.65	0.00
22	0.253	0.000	0.099	7.50	37.22	0.00
23	0.357	0.000	0.141	3.75	65.04	0.00
24	0.357	0.000	0.141	7.50	47.05	0.00

sible scaling function that satisfies all three of these constraints is described by the following equation:

$$\sqrt{\frac{\text{def}}{1+\alpha\tau}}, \qquad (13)$$

in which def is scaled by tilt (τ) and a free parameter (α). Figure 7 shows the observers' slant judgments plotted against this measure, using a value of .024 for the free parameter α. A regression analysis of these data revealed that the correlation is almost prefect ($r = .99$).

It is important to keep in mind that the stimuli used in Experiment 1 were constant flow fields, whose higher order temporal derivatives were not mathematically interpretable as rigid rotations, and it is possible therefore that the results may not generalize to other more ecologically valid patterns of motion. In an effort to address this issue, Experiment 2 was designed to investigate the perception of surface orientation from the orthographic projections of rotating dihedral angles whose optical flow fields deformed over time.

EXPERIMENT 2

Method

The apparatus and procedure were identical to those described for Experiment 1. Each display was specified by four parameters, V_0, V_X, V_Y, and ω, which were used to compute the instantaneous planar facets of a rotating dihedral angle from Equations 2, 5, and 6. The resulting dihedral angles were then textured with the same rocky pattern shown in Figure 3 and rotated back and forth in depth over a 24-frame sequence with a frame-to-frame angular displacement as specified by the value of ω. The surface orientation defined by the initial values of V_0, V_X, and V_Y always occurred in the middle frame of the apparent motion sequence; though unlike Experiment 1, these different components of the velocity field were changed over time as appropriate for the simulated 3-D rotation (see Liter & Braunstein, 1998). The moving dihedral angles were displayed under orthographic projection within a rectangular aperture whose horizontal and vertical dimensions were 800 and 640 pixels, respectively (i.e., 22.0º × 17.6º).

The displays included 24 different conditions with varying combinations of V_0, V_X, V_Y, and ω, which are described in Table 4. There were several important factors in the selection of these parameters that deserve to be highlighted. First, we included two different levels of total motion energy, in which all of the motion parameters were varied in a fixed proportion. This allowed us to examine whether perceived orientation is dependent on the absolute value of various motion gradients or on their relative values as a proportion of the total amount of motion in any given display. The vertical velocity gradient was varied across conditions over a range of values from .057 to .141. On half of the displays, all of the remaining motion energy was provided by the horizontal velocity gradient (i.e., $V_0 = 0$); on the other half of the displays, the remaining motion energy was entirely due to translation (i.e., $V_X = 0$). Finally, we also employed two different values of angular velocity, ω. Note that this had no effect whatsoever on the instantaneous pattern of image velocities in the middle of each apparent motion sequence. It only affected how those patterns changed over time. Thus, if perceived orientation were to vary with the value of ω, it could be interpreted as evidence that these judgments are influenced by higher order temporal derivatives in the depicted motion.

Judgments were obtained for the same 4 observers who participated in Experiment 1, including the 2 authors and 2 naive observers who were unfamiliar with the purpose of the experiment or how the displays had been generated. Each observer had corrected-to-normal vision.

Results

From the observers' judgments on each trial, we computed the adjusted slant and tilt of each planar facet to estimate the perceived orientations of the moving surfaces. The standard deviations of these judgments over repeated trials of the same condition are presented in Table 5 for each observer, and the correlations between observers are presented in Table 6. The average spread of the observers' judgments within a given condition was approximately 3º, which was slightly lower than in Experiment 1, though the variability between observers was slightly higher.

Previous investigations have provided evidence that perceived structure from motion is based primarily on the instantaneous field of image velocities and that higher order temporal derivatives of moving elements are of negligible importance (e.g., Liter et al., 1993; Perotti et al., 1998; Todd & Bressan, 1990; Todd & Norman, 1991). It follows from this hypothesis that observers' judgments of surface tilt should be highly accurate, since the tilt component of surface orientation is uniquely specified in the

Table 5
The Average Deviations Between Repeated Trials of
the Same Condition for the 4 Observers of Experiment 2

	J.T.	V.P.	J.N.	J.S.
Tilt	1.90	2.30	1.61	3.82
Slant	2.74	3.41	2.31	3.89

Table 6
The Correlations Between Observers in
Experiment 2 for Adjusted Slant and Tilt

	J.T.–V.P.	J.T.–J.N.	J.T.–J.S.	V.P.–J.N.	V.P.–J.S.	J.N.–J.S.
Tilt	.97	.98	.99	.98	.95	.97
Slant	.83	.81	.76	.90	.74	.88

instantaneous velocity field (see Equation 5). The results of the present experiment confirm this prediction. Figure 8 shows the average adjusted tilts, collapsed over observers, plotted as a function of simulated tilt for all 24 conditions. A regression analysis of these data revealed that the observers' judgments of tilt were almost perfectly accurate ($r = .99$), which is consistent with the results of Experiment 1 and the earlier findings of Domini et al. (1995) and Domini and Caudek (1999).

The slant component of surface orientation, on the other hand, can only be uniquely determined on the basis of an analysis of higher order temporal derivatives. If observers are unable to perform such an analysis, as suggested by previous research, then the accuracy of their slant judgments should be significantly impaired. This prediction is also confirmed by the results of the present experiment. In contrast to the high level of accuracy in the tilt component of perceived orientation, the observers' judgments of slant exhibited large systematic errors. Figure 9 shows the average adjusted slant plotted against the simulated slant for all of the different conditions. Note in this case that the two variables are only weakly correlated

such that the simulated slant accounts for less than 10% of the variance in the observers' judgments.

In the design of this experiment, we were particularly interested in the effects of varying the rate of angular velocity ω. Because this parameter was manipulated independently of V_0, V_X, and V_Y, its only effect on the pattern of depicted motion was in how the velocity field changed over time. Note in Figure 9 that the two different values of angular velocity (represented by open and filled symbols) produced large differences in the magnitude of simulated slant (see Equation 6), but these variations had no significant effect on the observers' judgments. On the basis of this finding, it seems reasonable to conclude that the higher order temporal derivatives among three or more views of the apparent motion sequences had little or no effect on their perceived 3-D structures.

Although the observers may not have been accurate in their judgments of slant, the results in Table 5 show clearly that they could perform these judgments with a high degree of reliability. What source of information could be responsible for defining a specific perceived orientation in each condition? In an effort to address this issue, Domini et al. (1995) and Domini and Caudek (1999) proposed that perceived slant is based primarily on the magnitude of def. We tested this hypothesis in the present experiment using an analysis of linear regression, and the results revealed that def accounted for only 20% of the variance in the observers' judgments (see Figure 10). There are two distinct factors evident in Figure 10 that are primarily responsible for this low correlation. For a given magnitude
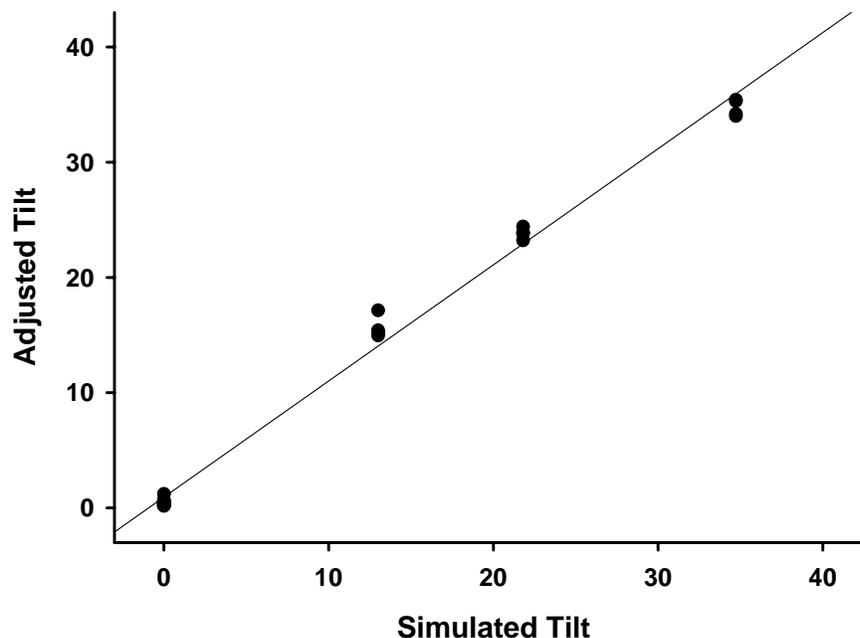

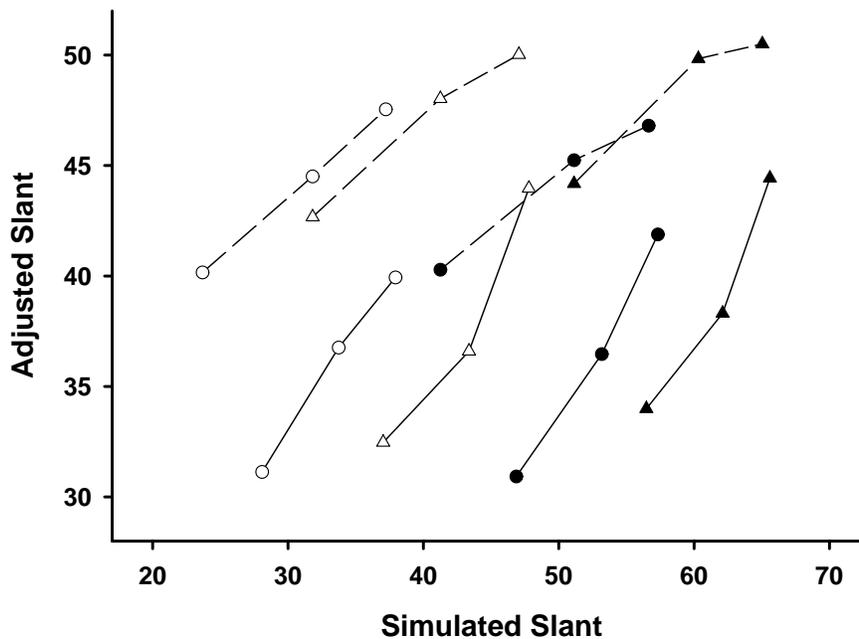Figure 8. The mean adjusted tilt in Experiment 2 plotted as a function of simulated tilt.

Figure 9. The mean adjusted slant in Experiment 2 plotted as a function of simulated slant. Conditions with zero values of V_X are represented by dashed lines, while those with zero values of V_0 are represented by solid lines. The triangles and circles represent different magnitudes of total motion energy, and the shading of these symbols represents different values of angular velocity.

of the vertical velocity gradient V_Y, nonzero values of the horizontal gradient V_X (represented by solid lines) produced less perceived slant than did those displays for which $V_X = 0$ (represented by dashed lines). Note that this effect is in the opposite direction of what would be expected if perceived slant were based solely on def. That is to say, for a given value of V_Y, the magnitude of def increases with V_X. Another relevant factor to consider in this regard is the manipulation of total motion energy. Given the finding of Domini and Caudek (1999) that the effect of def varies with tilt, it is useful to consider the subset of displays in which def was varied while tilt remained constant. The relevant comparisons are identified in Figure 10 by connected symbols. Those connected by dashed lines represent conditions with identical values of tilt, and those connected with solid lines represent conditions with identical values of both tilt and V_0 (see Table 4). The average difference in def between these matched conditions was approximately 30%, but the difference in their judged slants was only 6.8%. It would appear from these comparisons that the effects of def in this experiment were relatively small and that most of the variance in the observers' slant judgments was due to other factors.

In light of these observations, we were curious whether the scaling function employed in Experiment 1 would be similarly successful in the present study. Figure 11 shows the mean adjusted slant in each condition plotted against the measure described in Equation 13, using a value of

.066 for the free parameter α. As in Experiment 1, this scaling function is almost perfectly correlated with the observers' judgments ($r = .98$) and can account for over 96% of the variance among the different conditions.

DISCUSSION

During the past decade, there has been a growing body of evidence that the perception of 3-D structure from motion is based primarily on first-order temporal derivatives of moving elements within a visual image. One line of research that has supported this conclusion involves a comparison of tasks that are or are not theoretically possible on the basis of pure velocity information (see Todd & Bressan, 1990; Todd & Norman, 1991). For example, it can be shown mathematically that the tilt component of surface orientation is uniquely specified by a ratio of velocity gradients in orthogonal directions (see Equation 5) but that the slant component is inherently ambiguous without additional information about the rate of rotation ω. The results obtained in the present experiments and those obtained by Domini and Caudek (1999) reveal a similar distinction between slant and tilt in observers' perceptions of 3-D structure. Whereas the tilt component of judged orientation can by highly accurate, the slant component typically exhibits large systematic errors. A closely related finding has also been reported for the perception of surface curvature from motion (Dijkstra, Snoeren, & Gielen,
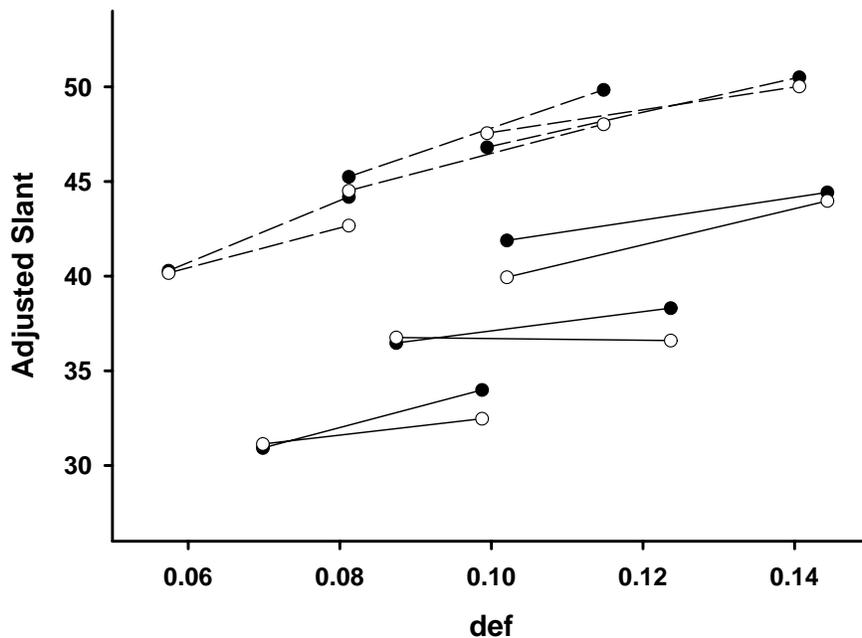
Figure 10. The mean adjusted slant in Experiment 2 plotted against the magnitude of def. The shading of the symbols represents different values of angular velocity. Those connected by dashed lines represent conditions with identical values of tilt, and those connected with solid lines represent conditions with identical values of both tilt and V_0.

1994; Perotti et al., 1998). Observers are quite accurate at judgments of shape index, which is mathematically specified in the instantaneous velocity field, but they exhibit large errors in judgments of curvedness, which is not.

Another source of evidence that perceived structure from motion is based primarily on velocity information comes from varying the higher order components of motion independently of the pattern of image velocities. This approach was used in the present research by comparing simulated rotations with constant flow fields and by comparing the same velocity patterns with varying values of ω to alter the manner in which they change over time. Although these manipulations had large effects on the simulated 3-D structures, they had little or no influence on perceived slant, thus indicating that the pattern of image velocities provided the primary source of information for observers' judgments. In a related experiment involving rotating dihedral angles, Eagle and Blake (1995) have shown that that observers can detect differences in higher order relations among three or more frames if they are sufficiently large, but the Weber fractions obtained in that study were over 100%. It is perhaps not surprising therefore that observers might rely on other sources of information that can be measured more precisely.

Whatever information observers use to estimate the 3-D structures of moving objects, it appears to exhibit some form of automatic gain control. That is, if the velocities of all moving elements in a display are increased or decreased by a uniform proportion, it has a relatively small

effect on observers' perceptions of depth or orientation (see Braunstein et al., 1993; Loomis & Eby, 1988, 1989; Todd & Norman, 1991). It is important to recognize, however, that this gain control is not perfect. For example, in one study by Todd and Norman (1991), a uniform doubling of the image velocities produced a 23% gain in the perceived amplitudes of sinusoidally corrugated surfaces, whereas a doubling of the simulated amplitudes produced an 87% gain. Similar effects from proportional increases of image velocity have also been reported by Liter et al. (1993) on the perceived relative depths of rotating random dots and by Domini and Caudek (1999) on the perceived slants of rotating planar patches. Although it was overshadowed by other factors in the present investigation, this same effect can be observed in Figure 9 as a result of our motion energy manipulation. For the conditions represented by triangles in this figure, the image velocities were on average 30% larger than in the corresponding conditions represented by circles. This produced a 6.8% increase in the magnitude of adjusted slant, which is proportionally comparable to the gains reported in previous investigations. The functional significance of this automatic gain control is to minimize the effects of angular velocity on the perception of 3-D structure from motion. In general, changes in 3-D structure will not produce a proportional scaling of the image velocity field, except in the degenerate case in which an object is stretched along the line of sight (see Norman & Todd, 1993; Todd & Norman, 1991).
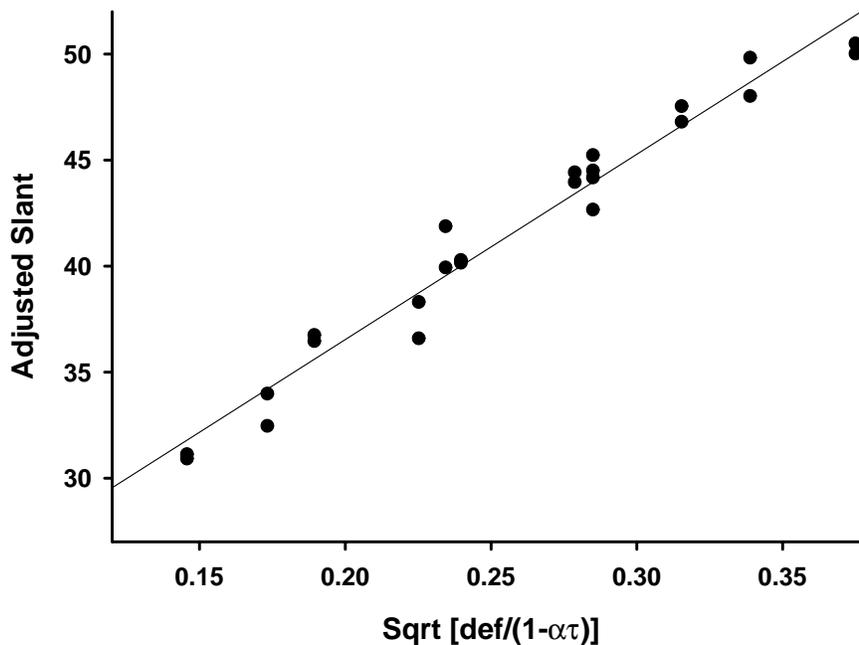
Figure 11—The mean adjusted slant in Experiment 2 plotted against the scaling function described by Equation 13.

A particularly perplexing aspect of the present results is the reduction of perceived slant that occurred with increasing magnitudes of the horizontal velocity gradient V_X (i.e., the gradient perpendicular to the axis of rotation). Other things being equal, increasing the slant of a rotating surface will produce a corresponding increase in the magnitude of V_X (see Equation 3), yet increases in V_X can produce reductions of perceived slant. Although this might seem to make little sense from an ecological perspective, this effect has been confirmed in several different experiments. For example, in one such study by Braunstein et al. (1993), the introduction of a horizontal velocity gradient with fixed values of V_Y significantly lowered the perceived relative orientation between the two planar facets of a rotating dihedral angle. Similar manipulations have been performed indirectly by varying the 3-D orientations of dihedral angles undergoing perspective rotation (Tittle et al., 1995) or translation (Liter & Braunstein, 1998). In both cases, there is a reduction in perceived relative orientation between the two planar facets as the horizontal component of the velocity gradient takes on a greater and greater proportion of the total motion energy (see also Domini & Caudek, 1999).

It is important to keep in mind when evaluating this issue that all of the studies described above have employed the same basic stimulus configuration: a single planar patch or a dihedral angle whose edge is parallel to the direction of image motion. Thus, given that the apparent scaling mechanism used by observers (see Figures 7 and 11) has no obvious theoretical justification, it is best to be cautious before drawing too strong a conclusion about the generality of these findings. An interesting problem for future research will be to investigate the effects of velocity gradients in different directions for other types of surfaces, such as quadrics, and for random configurations of connected line segments.

REFERENCES

BENNETT, B., HOFFMAN, D., NICOLA, J., & PRAKASH, C. (1989). Structure from two orthographic views of rigid motion. *Journal of the Optical Society of America*, **6**, 1052-1069.

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 & Performance*, **19**, 598-614.

CORNILLEAU-PÉRÈS, V., & DROULEZ, J. (1989). Visual perception of curvature: Psychophysics of curvature detection induced by motion parallax. *Perception & Psychophysics*, **46**, 351-364.

DIJKSTRA, T. M. H., SNOEREN, P. R., & GIELEN, C. C. A. M. (1994). Extraction of three-dimensional shape from optic flow: A geometric approach. *Journal of the Optical Society of America A*, **11**, 2184-2196.

DOMINI, F., & CAUDEK, C. (1999). Perceiving surface slant from deformation of optic flow. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 426-444.

DOMINI, F., CAUDEK, C., & GERBINO, W. (1995). Perception of surface attitude in SFM displays. *Investigative Ophthalmology & Visual Science*, **36**, s360.

EAGLE, R. A., & BLAKE, A. (1995). Two-dimensional constraints on three-dimensional structure from motion tasks. *Vision Research*, **35**, 2927-2941.

KOENDERINK, J. J., & VAN DOORN, A. J. (1977). How an ambulant observer can construct a model of the environment from the geometrical structure of the visual inflow. In G. Hauske & F. Butenandt (Eds.), *Kybernetik* (pp. 224-247). Munich: Oldenburg.

KOENDERINK, J. J., & VAN DOORN, A. J. (1991). Affine structure from motion. *Journal of the Optical Society of America A*, **8**, 377-385.

LITER, J. C., & BRAUNSTEIN, M. L. (1998). The relationship of vertical and horizontal velocity gradients in the perception of shape, rotation and rigidity. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1257-1272.

LITER, J. C., BRAUNSTEIN, M. L., & HOFFMAN, D. D. (1993). Inferring structure from motion in two-view and multi-view displays. *Perception*, **22**, 1441-1465.

LOOMIS, J. M., & EBY, D. W. (1988). Perceiving structure from motion: Failure of shape constancy. In *Proceedings from the Second International Conference on Computer Vision* (pp. 383-391). Washington, DC: IEEE.

LOOMIS, J. M., & EBY, D. W. (1989). Relative motion parallax and the perception of structure from motion. In *Proceedings from the Workshop on Visual Motion* (pp. 204-211). Washington, DC: IEEE.

NORMAN, J. F., & LAPPIN, J. S. (1992). The detection of surface curvatures defined by optical motion. *Perception & Psychophysics*, **51**, 386-396.

NORMAN, J. F., & TODD, J. T. (1993). The perceptual analysis of structure from motion for rotating objects undergoing affine stretching transformations. *Perception & Psychophysics*, **53**, 279-291.

NORMAN, J. F., & TODD, J. T. (1995). Perception of 3-D structure from contradictory optical patterns. *Perception & Psychophysics*, **57**, 826-834.

PEROTTI, V. J., TODD, J. T., LAPPIN, J. S., & PHILLIPS, F. (1998). The perception of surface curvature from optical motion. *Perception & Psychophysics*, **60**, 377-388.

PEROTTI, V. J., TODD, J. T., & NORMAN, J. F. (1996). The visual perception of rigid motion from constant flow fields. *Perception & Psychophysics*, **58**, 666-679.

PEROTTI, V. J., TODD, J. T., TITTLE, J. S., & NORMAN, J. F. (1994). Perception of 3D form from instantaneous flow components. *Investigative Ophthalmology & Visual Science*, **35**, 1317.

SNOWDEN, R. J., & BRADDICK, O. J. (1991). The temporal integration and resolution of velocity signals. *Vision Research*, **31**, 907-914.

TITTLE, J. S., TODD, J. T., PEROTTI, V. J., & NORMAN, J. F. (1995). Systematic distortion of perceived three-dimensional structure from motion and binocular stereopsis. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 663-678.

TODD, J. T. (1981). Visual information about moving objects. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 795-810.

TODD, J. T., & BRESSAN, P. (1990). The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Perception & Psychophysics*, **48**, 419-430.

TODD, J. T., & NORMAN, J. F. (1991). The visual perception of smoothly curved surfaces from minimal apparent motion sequences. *Perception & Psychophysics*, **50**, 509-523.

ULLMAN, S. (1977). *The interpretation of visual motion*. Unpublished doctoral dissertation, Massachusetts Institute of Technology.

ULLMAN, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.

ULLMAN, S. (1983). Recent computational studies in the interpretation of structure from motion. In J. Beck & A. Rosenfeld (Eds.), *Human and machine vision* (pp. 459-480). New York: Academic Press.

WERKHOVEN, P., SNIPPE, H., & TOET, A. (1992). Visual processing of optical acceleration. *Vision Research*, **32**, 2313-2329.