

The perception of 3-D structure from contradictory optical patterns

J. FARLEY NORMAN and JAMES T. TODD
Ohio State University, Columbus, Ohio

Two experiments investigated observers' perceptions of 3-D structure when optical sources of information were contradictory. When motion and stereoscopic disparities specified different surfaces, the perceptual outcome depended strongly on the direction of curvature present within each modality. Previous research has shown that the perception of surface slant and curvature is anisotropic for both motion and stereo and that it depends on the direction in which it takes place. In the present experiments, the modality with the "effective" direction of curvature tended to dominate or suppress the perception of surfaces in the other modality with less effective curvatures. The results have implications for models which attempt to combine 3-D data from different optical sources.

Human observers have the remarkable ability to perceive environmental 3-D structure from many different sources of optical information, including binocular disparities of identifiable features (stereopsis), motion parallax, differential projected image motions due to the rotation of a solid object (the kinetic depth effect), and patterns of shading and highlights across the projected area of surfaces. Research has shown that any one of these optical sources is sufficient to produce compelling impressions of 3-dimensional shape. When they are available simultaneously in a natural environment, the human visual system somehow integrates this information into a coherent, unified percept, although the exact process by which this takes place remains largely a matter of speculation (see, however, Landy, Maloney, Johnston, & Young, 1995; Tittle & Braunstein, 1993).

Most of the time, these different sources of optical information provide similar input to higher visual centers. For example, the simultaneous differences between the two eyes' views are replicated over time in a single eye's view as a 3-D object rotates by an amount equivalent to the convergence angle. However, this similarity does not always exist, especially in the modern world. Contradictory visual situations often arise, although we are not usually consciously aware that this is so. The viewing of paintings is one of the most obvious examples. While you are admiring a Renaissance painting, the pattern of binocular disparities specifies that you are looking at a flat 2-dimensional surface. The shading, highlights, texture gradients, and perspective variations are more parsimoniously consistent with a 3-D interpretation (see

Nakayama & Shimojo, 1992, for a discussion of accidental vs. nonaccidental properties). Looking at pictures is inherently a contradictory visual situation, as was noted by Wheatstone (1838, p. 372). A similar example is the act of watching television. The deformations and motions within the changing image indicate that you are viewing 3-D objects and surfaces, but again the stereoscopic disparities provide you with information that the TV screen is slightly curved, or perhaps flat.

In such contradictory situations in which different inputs are indicating different things to the visual system, that system has several alternatives. It can combine the differing inputs in some way, for example, by an average. This presupposes that the two inputs have a common currency or format, or can be converted into the same format; otherwise there is no way to even define the concept of an average. For 3-D vision, such a common high-level currency might be depths, slants, or curvatures. Alternatively, one of the discrepant inputs could be suppressed while the remaining information is used for further analysis. It is a well-known phenomenon that the human visual system often uses suppression to eliminate conflicts. In particular, people with strabismus or anisometropia have conflicting visual inputs to the two eyes. In strabismus, one eye's view is quite different, because while the normal eye is fixating a desired target, the deviating eye is looking somewhere else. This means that different objects appear at the fovea and other parts of the visual field. Under these conditions, people may see double images, or diplopia, at a given location in the visual field. However, the most common solution imposed by higher cortical visual centers is to suppress the perception of one of the conflicting targets (see Schor, 1977; Sireteanu, 1982, for a thorough discussion of strabismic suppression; and Sireteanu & Fronius, 1981). Similar suppressions also occur during binocular rivalry for non-strabismic observers (Blake, 1989; Fox & Check, 1972; Fukuda & Blake, 1992).

This research was supported by Air Force Office of Scientific Research Grant F49620-93-1-0116 to J. T. We would like to thank Jim Tittle for helpful discussions relating to these issues. Correspondence should be addressed to J. F. Norman, Department of Psychology, Ohio State University, 142 Townshend Hall, Columbus, OH 43210-1222 (e-mail : fnorman@magnus.acs.ohio-state.edu).

The main purpose of the present experiments was to evaluate observers' perceptions of 3-D structure in situations in which the stereoscopic disparities and motions (kinetic depth effect) within a combined optical pattern specified completely different surfaces, with curvatures in orthogonal directions. For example, if one modality indicated to an observer that he was viewing a sinusoidal surface with horizontal corrugations, the other modality simultaneously specified a vertically corrugated sinusoidal surface. In these circumstances, there are only three potential outcomes at any given time: (1) the stereo surface or (2) the motion-defined surface may be visible, suppressing the perception of the other modality's surface, or (3) the depths, orientations, or curvatures of the two individual surfaces are combined in some fashion (such as an average) to produce the perception of an egg-crate-like surface with curvatures in both directions that is not defined in either modality alone.

One factor that may potentially play a major role in determining what is seen in contradictory circumstances is the direction of curvature. Rogers and Graham (1983) have shown that the stereoscopic perception of curved surfaces is anisotropic. Surfaces that are curved in a vertical direction (such as a horizontal cylinder) are more salient than those with curvatures in the horizontal direction (vertical cylinder). Norman and Lappin (1992) and Cornilleau-Pérès and Droulez (1989) showed that a similar anisotropy existed for curved surfaces defined by the kinetic depth effect. However, for motion there is one important difference. The two stereoscopic views of an observer can always be related by a rotation about a vertical axis, since the eyes are separated horizontally on the head. The motions produced by the kinetic depth effect do not have this limitation—rotations can take place about any axis. Therefore, it is the direction of curvature relative to the axis of rotation that is the important determining factor for perception, rather than absolute orientation in space per se (Droulez & Cornilleau-Pérès, 1990; Norman & Lappin, 1992). For example, a vertical cylinder may be more salient perceptually, rather than less, if it is rotated about a horizontal axis instead of a vertical one.

Uomori and Nishida (1994) placed stereopsis and motion into conflict by using cylinders oriented in different directions and examined how the resulting perceived shape changed over time. They found that the motion-defined cylinder was perceived initially, but as time continued, the stereo surface gradually became dominant. This is an interesting result, but it may have little generality since the design of their experiment did not include the direction of curvature as an important factor. In their displays, the surface defined by the kinetic depth effect contained the relatively weak and ineffective direction of curvature (a vertical cylinder rotating about a vertical axis). This may explain why the stereoscopic horizontal cylinder with the effective direction of curvature dominated at longer time intervals. In one condition, they did use horizontally oriented cylinders for the motion stim-

uli. However, in this case as well, the motion-defined surface had the ineffective direction of curvature, since they rotated the horizontal cylinders about a horizontal axis. In their experiments they did not use motion-defined cylinders that had the more "effective" curvatures in a direction parallel to the axis of rotation.

If human observers perceived depths or surface orientations equally well in all conditions, there would be no anisotropy regarding the perception of curved surfaces defined by motion and/or stereopsis. In this case, it would make no difference which direction of curvature was present for motion and stereo. But, curvatures in some directions are more "effective" and can be detected and discriminated better than identical curvatures in other (degenerate) orientations. Hence, in a contradictory situation, where one is attempting to evaluate how motion and stereo combine together, one should include direction of curvature as a fundamental and important factor. Otherwise, one may significantly bias the results. Either stereo or motion may become dominant only because one modality has the effective orientation, while the other does not. Accordingly, the present set of experiments were designed in an effort to resolve the ambiguities surrounding this issue. Determining how the higher levels of the visual system behave given contradictory 3-D inputs should allow us to differentiate between alternative explanations for how the visual system integrates 3-D structural information from different optical sources.

EXPERIMENT 1

Method

Apparatus. The optical patterns were created and displayed on a Silicon Graphics Crimson VGXT workstation. Stereoscopic viewing hardware was also used. The stereoscopic half-images were presented with 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, so that each view of a stereoscopic half-image was updated at half that, or 60 Hz. The viewing distance was 114.6 cm, so that the 1,280-pixel-wide \times 1,024-pixel-high display screen subtended $16.9^\circ \times 13.5^\circ$ of visual angle.

Stimulus displays. The contradictory surfaces depicted in this experiment were defined by the binocular disparities and motions of 1,000 hardware antialiased luminous points. The effective resolution of the display was at least 0.1 pixel. The motion-defined surface was a horizontally or vertically oriented sinusoidal surface oscillating around a vertical axis, while the stereoscopically defined surface was sinusoidally corrugated in an orthogonal direction. The spatial frequency of the corrugations was constant, at 0.48 c/deg of visual angle, which is near the peak of the stereoscopic and motion modulation transfer functions (Norman, Lappin, & Zucker, 1991; Rogers & Graham, 1982). The apparent motion sequences for the motion-defined surfaces consisted of 15 individual views—a "frontal" view and 7 views on either side from frontal. The motion surface rotated 2.5° between every frame transition. Therefore, the surface oscillated between $\pm 17.5^\circ$. The

frames were presented at a temporal rate of 30 Hz. The patterns were viewed through a circular aperture (5.5-cm radius, or 2.75° of visual angle) in a square-shaped occluder which hid the deforming boundaries of the moving pattern. The occluder had a large crossed disparity, so that it appeared to be hovering in depth in front of the stereoscopic/motion defined surfaces. The stereoscopic patterns were consistent with an interpupillary distance of 6.3 cm. An example of a stereoscopic stimulus pattern is shown in Figure 1.

It is not possible in the real world to find a 3-D object that has curvatures in different directions for motion and stereo. It is important, therefore, to describe the particular method with which we created these optical patterns. The simplest way to produce such a contradictory pattern would be to make an apparent motion sequence for one eye (for example, the left), where the positions of the moving points in each frame are exactly appropriate for the 3-D shape defined by motion, while at the same time copying those positions in each frame to the other (right) eye's view, adding the disparities for each point necessary to specify the contradictory stereoscopic surface. If one used this technique, one eye would have motions that were the correct projections of points on an object's surface, while the motions in the other eye would not correspond to the projections of any rigidly moving 3-D structure. To prevent such qualitative differences between the two eye's views, we divided the stereoscopic disparities so that the positions of the points in each eye were displaced by half of the disparity needed for them to specify the desired stereoscopic surface. Despite this image distortion, all of the observers reported that the apparent motion sequences for the left and right eyes, when viewed in isolation, appeared as rigidly rotating 3-D surfaces. The stereoscopic surface defined by the disparities of the points was itself stationary, although the positions of the points moved over time. In this respect, the stereoscopic displays were similar to dynamic random-dot stereograms (see, e.g., Fox & Patterson, 1981, and Lehmkuhle & Fox, 1980), in which the stereoscopic form itself appears stable and compelling despite the appearance and disappearance of individual points over time.

Procedure. On any given trial, observers initially viewed a standard surface defined by one modality at a given amplitude in depth—the amplitude of the other modality's surface (the test surface) was initially set to zero (a flat frontoparallel plane). The standard surface was presented at one of four amplitudes (0.25, 0.575, 0.9, and 1.225 cm, zero-peak), and the observers' task was to adjust the amplitude of the test surface until its depth variations (orthogonal to the corrugations of the standard surface) were just visible—an ascending method of limits. The observers made their adjustments in amplitude of the test surface with the workstation's mouse. The ob-

servers had an unlimited amount of time in which to make their adjustments. The largest amplitude standard surface was 1.225 cm, zero-peak. If an observer, while performing the adjustment, could not perceive the test surface at an amplitude of 2.0 cm, zero-peak, then the response for that trial was clipped to 2.0 cm, and the next trial was then initiated. The four standard amplitudes (zero-peak) of 0.25, 0.575, 0.9, and 1.225 cm corresponded to stereoscopic disparities of $0.8'$, $1.9'$, $3.0'$, and $4.0'$ of arc at the 114.6-cm viewing distance. These values were chosen following extensive pilot observation. One can see that the absolute sensitivity of our observers was good, because the lowest amplitude standard surfaces were clearly visible at a maximum disparity of $0.8'$ or $48''$ of arc.

There were 16 experimental conditions formed by the orthogonal combination of two types of standard surfaces (stereo and motion), two orientations of the standard surface (horizontal and vertical corrugations; if the standard surface was horizontal, then the test surface was vertical, and vice-versa), and four levels of standard surface amplitude. The order of the 16 conditions was randomly varied within a block of trials. Five adjustments were collected for each of the 16 experimental conditions.

Observers. The displays were presented to 4 observers, two of whom were the authors (J.F.N and J.T.T.). Observer J.S.T. was familiar with the types of issues being investigated. However, V.J.P., while having previous experience as a psychophysical observer, was naive with regard to the particular purposes of this experiment. All observers had normal or corrected-to-normal vision.

Results and Discussion

The results for the motion standard surfaces are shown in Figure 2, for each observer separately (observers adjusted the stereo surface). The open symbols indicate the condition in which the motion standard surface was vertically oriented and the stereo test surface was horizontally oriented. In this case, the motion standard surface had the less potent curvature, because of the anisotropy of orientation. It was unable to effectively suppress the perception of the stereo test surface. The corrugations of the sinusoidal test surface in this condition were perceived at relatively low amplitudes. In contrast, consider the condition in which the motion standard surface had the effective horizontal corrugations (filled circles). In this condition, the standard surfaces dominated, and the perception of the stereo test surface was inhibited over a

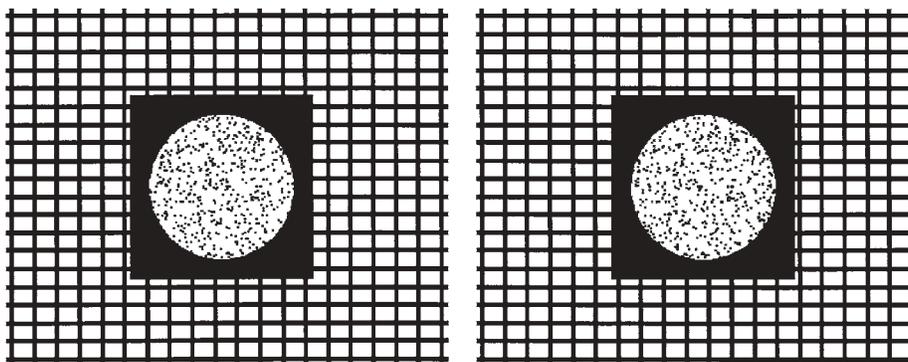


Figure 1. An example of a stereoscopic stimulus used in Experiment 1. This stereogram is designed for crossed free-fusion. It is not possible to illustrate the contradictory nature of the combined stereo/motion optical patterns that were actually displayed. However, in this example, the stereoscopic surface has horizontal peaks and troughs—the contradictory motion-defined surface would also be sinusoidally corrugated, but would have vertical peaks and troughs.

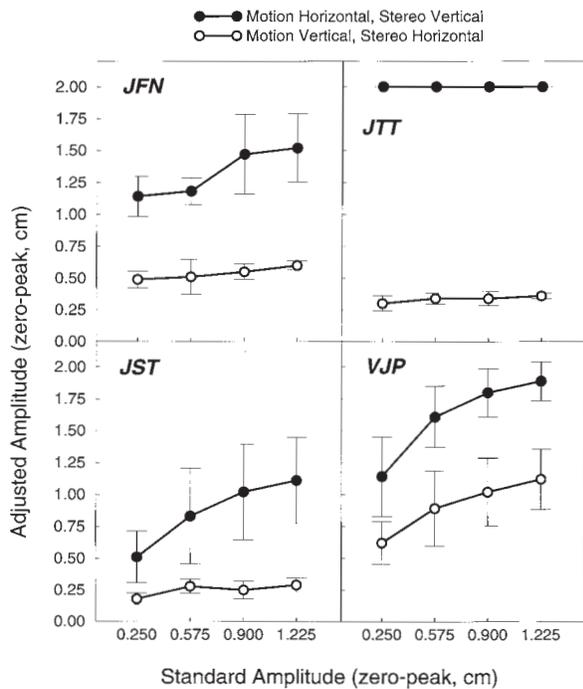


Figure 2. Results for the motion standard surfaces (observers adjusted stereoscopic test surfaces), presented separately for each observer. The adjusted amplitude necessary for the test surfaces' corrugations in depth to be perceived is plotted as a function of standard surface amplitude. The error bars represent ± 1 SD. The filled circles indicate that the motion standard surfaces were horizontally corrugated. The opaque symbols represent vertically corrugated standard surfaces.

much larger range of test amplitudes. Indeed, this was a very striking effect—in both conditions, the amplitude of the motion standard surface was the same (0.25, 0.575, 0.9, or 1.225 cm, zero-peak).

The only difference between the two cases was the orientation of the standard and test surfaces. The large difference in height between the two curves shows that the changes in the orientation of the motion standard surfaces had large effects on the visibility of the stereoscopic test surfaces. The strength of the suppression when the standard surface had the effective orientation was proportional to the amplitude of the standard—larger standard amplitudes led to greater suppression.

Phenomenologically, in the region below the curves, where the amplitude of the test surface was insufficient to break the suppression, only the depth variations of the standard surface were seen. The perception of the test surface was completely inhibited. The observers did not typically see an egg-crate surface, which might have been expected if the two different surfaces' depths, orientations, or curvatures had been averaged together. This suppression of one modality by another was often very strong; for example, for three out of the four observers, a horizontally oriented motion standard surface of 0.25-cm amplitude was able to suppress the perception of a 1.0-cm amplitude (four times larger) stereo surface with

vertical peaks and troughs. Sometimes, just before the threshold amplitude was crossed from invisibility to visibility, the observers could detect the presence of the test surfaces' disparities or motions, because of such factors as perceived rigidity, and so forth. But if this occurred, the observers were instructed to keep increasing the amplitude of the test surface until actual variations in *depth* in the test direction became visible.

Figure 3 shows the analogous results for the stereoscopic standard surfaces. One can see that the anisotropy of orientation is evident in these conditions as well, but is weaker. The stereo standard surfaces with the horizontal corrugations did suppress the test surface more effectively than the standard surfaces with vertical corrugations. The results for J.F.N., J.T.T., and J.S.T. were essentially identical, whereas the anisotropy seemed especially weak for Observer V.J.P. In informal further testing with V.J.P, we used higher stereo standard amplitudes (1.8 and 2.45 cm, zero-peak). At those higher standard amplitudes, V.J.P. showed significant anisotropy effects, similar to those for the other observers.

It is interesting that in our conditions, in which stereo and motion specified curvatures in orthogonal directions, the combination rule seemed to be primarily a veto, rather than some other combination rule. For large regions under the curves, the test surfaces had amplitudes that would normally be easily visible, if they were

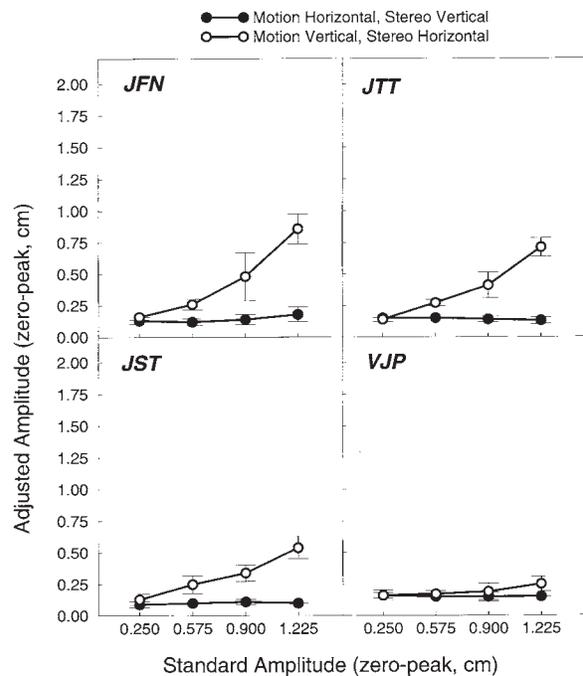


Figure 3. Results for the stereoscopic standard surfaces (observers adjusted motion surfaces), presented separately for each observer. The adjusted amplitude necessary for the test surfaces' corrugations in depth to be perceived is plotted as a function of standard surface amplitude. The error bars represent ± 1 SD. The filled circles indicate that the stereo standard surfaces were vertically corrugated. The opaque symbols represent horizontally corrugated standard surfaces.

not presented in contradictory situations. When the standard and test surfaces had similar amplitudes, one surface tended to dominate the other, depending primarily on the relative orientations of each. However, as the test amplitude was increased, at some point the observers crossed a "threshold" and the curvatures of the test surface became visible. This does not necessarily mean that the two suprathreshold curvatures were always integrated into a coherent percept of a rigidly moving egg-crate surface. In fact, this rarely occurred. To illustrate, consider an example in which the standard stereoscopic surface has vertical corrugations and one is adjusting (from zero) the amplitude of a motion surface with horizontal peaks and troughs. One initially sees the stereo vertical surface only. At some point in the adjustment, the horizontal corrugations of the motion surface become visible. The perception of the previously visible stereo surface will then frequently be suppressed. As the amplitude of the test surface increases from zero, the perceived surface changes from stereo vertical only to motion horizontal only. One may or may not see an "intermediate" egg crate; if one does perceive this intermediate surface with curvatures in both directions, it will only occur for a narrow range of standard and test surface amplitudes.

Informal testing with these contradictory patterns also revealed that when both standard and test surfaces have suprathreshold amplitudes, rivalry may be seen. If one views a display with an easily visible motion standard surface with horizontal corrugations and adjusts the amplitude of a stereo test vertical surface until it breaks the suppression, one will not usually perceive a coherently rotating egg crate. Rather, the two surfaces often engage in rivalry, and one will alternately perceive one surface followed by the other, in a repeating cycle.

The ability of the motion standard surfaces to suppress stereoscopic test surfaces was especially strong for Observer J.T.T. Notice in Figure 2 that the motion surface with the effective horizontal corrugations completely suppressed the perception of the stereoscopic vertical corrugations. In this condition, J.T.T. never perceived the test surface, no matter what the amplitude, and his test amplitude adjustments were therefore clipped at 2.0 cm, zero-peak. However, in informal testing with noncontradictory patterns, J.T.T. was able to perceive the corrugations of the stereoscopic vertical surfaces as easily as the other observers.

EXPERIMENT 2

In Experiment 1, the anisotropy obtained while using motion standard surfaces was found to be larger than the analogous anisotropy for stereoscopic standard surfaces. One possible explanation of this difference might be that the stereoscopic disparity field was static, whereas the velocity field was changing over time. Almost all 3-D objects generate varying velocity fields in an optical projection when they undergo rotation in depth. Consider,

for example, the velocity measured at a particular place in the visual field. It changes over time as new parts of a 3-D object successively occupy and leave that part of the visual field. The only surfaces that do not generate such time-varying optic flow fields are surfaces of revolution such as cylinders, spheres, and hyperboloids that rotate about their axes of symmetry. For these surfaces undergoing these particular motions, every location in the optical projection will always indicate a constant velocity. For examples of such surfaces and motions, see Braustein, Hoffman, and Saidpour (1989).

Since constant flow fields do not change over time, they are qualitatively similar to instantaneous binocular disparity fields. We were curious to know whether the difference between the strength of suppressions for the motion and stereoscopic standard surfaces observed in Experiment 1 had occurred because one modality (stereo) had a static field, whereas the other (motion) had a dynamically changing flow field. This qualitative difference was therefore removed in Experiment 2 by using constant flow fields to define the motion. An advantage of this technique is that the same equations can be used to generate both the stereoscopic disparity and the motion fields. The resulting motion and stereoscopic patterns were much more similar to each other than those used in Experiment 1.

Method

Apparatus. The apparatus was identical to that used in Experiment 1.

Stimulus displays. The contradictory stimuli were similar in many respects to those of Experiment 1, with two significant differences. First, the curved surfaces in Experiment 2 were elliptical cylinders rather than sinusoidally corrugated surfaces. And second, the motion surface was defined by a constant flow field.

The constant vector fields for both motion and stereo were defined in the following fashion. Todd and Norman (1991) showed that for small displacements of a surface rotating about an axis in the image plane under parallel projection, the optical motion of a given point i is equal to the product of its position in depth (Z_i) relative to the rotation axis and the angular velocity ε :

$$D_i = \varepsilon Z_i. \quad (1)$$

In order to simulate the instantaneous velocity field (or disparity field) of a rotating 3-D surface, we can substitute into Equation 1 any surface of the form $Z = f(x, y)$, a Monge surface. The surfaces used in this experiment were vertically or horizontally oriented elliptical cylinders. The cylinders were presented either in a frontoparallel orientation or at a 30° slant. In particular, for a vertically oriented cylinder slanted in depth by an angle ϕ about the horizontal axis,

$$Z = \left(D_0 + b - b\sqrt{1 - \frac{X^2}{a^2}} \right) \cos \phi + \left[Y + \left(D_0 + b - b\sqrt{1 - \frac{X^2}{a^2}} \right) \sin \phi \right] \tan \phi. \quad (2)$$

Similarly, for a horizontal cylinder slanted about the vertical axis,

$$Z = \left(D_0 + b - b \sqrt{1 - \frac{Y^2}{a^2}} \right) \cos \phi + \left[X + \left(D_0 + b - b \sqrt{1 - \frac{Y^2}{a^2}} \right) \sin \phi \right] \tan \phi, \quad (3)$$

where D_0 , a , b , and ϕ are constants, such that, for a given trial, velocities and disparities are a constant function of x and y in the image. The simulated rotations were about a vertical axis in the image plane. D_0 refers to the simulated distance from the rotation axis to the cylinder surface (see Figure 4) and was set to 4.0 cm. a and b refer to the length of the semi-axes of the elliptical cylinder, and ϕ refers to the simulated slant angle. a was held constant at 8.0 cm, while b was varied to manipulate the eccentricities of the elliptical cylinders used as standard and test surfaces.

By substituting (2) and (3) into (1) we get expressions for velocity and disparity. The ε value used to generate the disparity field was the actual convergence angle of 4.75° for viewing the displays at a distance of 76.0 cm. The ε value to generate the motion field was 1.0° , simulating the instantaneous rotation of $30^\circ/\text{sec}$. The surfaces appeared to move in an oscillatory manner, so that the motions resembled those used in Experiment 1. This was accomplished by switching the direction of the points' motions from right to left and back again every 15 frames, although the magnitudes of the velocities were always constant over the entire motion sequence. The displays were viewed through a square aperture 14 cm on each side which occluded the outermost portions of the simulated cylinders. All other stimulus parameters and methods were identical to those used in Experiment 1.

There were 48 experimental conditions formed by the orthogonal combination of two modalities of standard surfaces (stereo and motion) \times two orientations of the standard surface (horizontal or vertical elliptical cylinder) \times three levels of standard surface amplitude (low, medium, and high amplitudes, corresponding to b values of 4.0, 8.0, and 12.0 cm, respectively) \times two levels of standard surface slant (0° and 30°) \times two levels of test surface slant (0° and 30°).

Procedure. The psychophysical procedure was identical to that of Experiment 1. On any given trial, the observers saw a standard surface defined by one modality at a given amplitude and were instructed to adjust the amplitude, from zero, of the test surface oriented in an orthogonal direction until the depth variations of the test surface were just visible. The observers made their test amplitude adjustments using the workstation's mouse. Each observer made five adjustments for each of the 48 experimental conditions. The observers had an unlimited amount of time to make their adjustments.

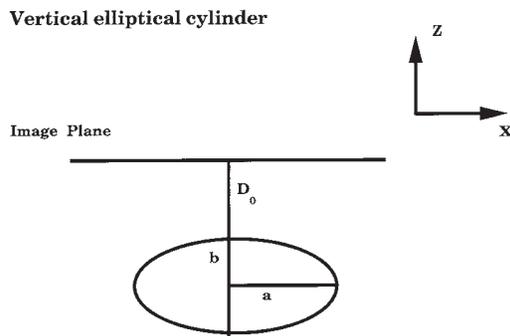


Figure 4. A schematic illustration of the elliptical cylinders used as stimuli in Experiment 2.

Observers. The 4 observers were the same as those who participated in Experiment 1.

Results

The results can be seen in Figures 5 and 6 for motion and stereoscopic standard surfaces, respectively. Despite the usage of very different optical patterns, the pattern of results is essentially identical to the results of Experiment 1, in which corrugated sinusoidal surfaces and projected 3-D rotation were used. We have collapsed over the four levels of the standard and test slant manipulations, since there was no systematic effect of these variables. For 3 observers (J.T.T., J.S.T., and V.J.P.), the anisotropy observed when motion standard surfaces were used was again larger than the anisotropy obtained with stereoscopic standard surfaces. It would therefore appear that although the orientation of surfaces (anisotropy) is a primary determining factor, there is an asymmetry such that motion-defined surfaces more readily suppress stereoscopic surfaces than vice versa.

GENERAL DISCUSSION

We have seen that the human visual system in its analysis of 3-D shape tends to use suppression when faced with conflicting inputs, at least under conditions similar to those used in our experiments. We found that the major factor determining which modality was visible or suppressed was the orientation of each modality's constituent curvatures. Whichever modality (motion or stereo) had the "effective" direction of curvatures was perceived while the other modality's surface was suppressed. Thus, it would appear that the anisotropy of orientation reported by Rogers and Graham (1983), Cornilleau-Pérès and Droulez (1989), and Norman and Lappin (1992) in research involving both motion and stereo has an influence on how multiple sources of information are combined together.

In general, we did not find that the perceived surface was the result of a simple linear combination of the depths of the stereo and motion-defined surfaces. This combination rule has been referred to as *weak fusion* (Clark & Yuille, 1990; Landy et al., 1995; Young, Landy, & Maloney, 1993). The specific predictions of the weak fusion model have not been supported psychophysically, and in a series of papers, Landy and his colleagues (Landy et al., 1995; Landy, Maloney, & Young, 1991; Maloney & Landy, 1989; Young et al., 1993) have proposed a model of combination described as "modified weak fusion." In this model, each modality forms its own low-level representation, and early interactions between modalities allow each modality to specify missing parameters to others so that all of the low-level representations can be converted to a common format. Following this conversion, which they refer to as *promotion*, the final single representation of depth is a linear combination of the individual depth maps. The weights for each modality's contribution vary over both space and time depending on the reliability or robustness of each modality's information

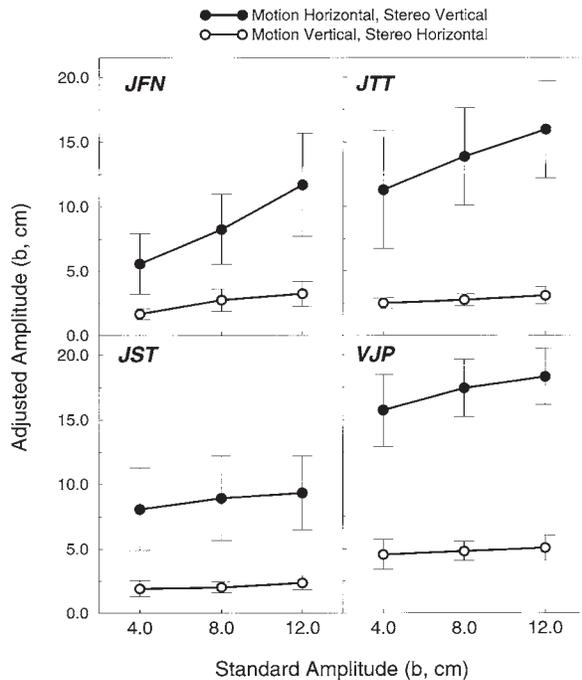


Figure 5. Results for the motion standard surfaces, presented separately for each observer. The adjusted amplitude necessary for the test surfaces' corrugations in depth to be perceived is plotted as a function of standard surface amplitude. The error bars represent ± 1 SD. The filled circles indicate that the motion standard surfaces were horizontally corrugated. The opaque symbols represent vertically corrugated standard surfaces.

and on how discordant any given modality is compared with the others. The modified weak fusion model allows for any modality's weight to become zero (a veto situation) if it is supplying information highly discrepant from the others. In this respect, our results fit within this general scheme, because the weaker modality having the less effective direction of curvatures was suppressed over a wide range of amplitudes by the modality presented with a 3-D surface with the more effective curvatures.

The research of others shows that different combination rules are used by the visual system when attempting to combine or integrate less contradictory inputs. For example, Rogers and Collett (1989) showed that if motion parallax specified sinusoidally corrugated surfaces while the stereoscopic disparities were consistent with a flat plane, the perceived amplitude of the display was about half of that observed with motion parallax alone. The perceived surface was approximately an average of the two contradictory inputs. However, when the stereoscopic pattern was also sinusoidally corrugated (curvatures in the same direction; the motion- and stereo-defined sine waves were either in phase or 180° out of phase with different amplitudes), Rogers and Collett found that the stereoscopic disparities were the primary determinant of the resulting perceived amplitude, whereas the motion parallax in the display was interpreted by the observers as rotation in depth. Tittle and Braunstein

(1991) also found that perceived amplitudes were primarily determined by stereoscopic disparities, rather than by the motions in a contradictory optical pattern. The motions and disparities in their conditions simulated ellipsoids with differing extents in depth—the curvatures in the motion and stereo fields were different in magnitude, not in direction.

In contrast to the findings in the two previously cited studies, Johnston, Cumming, and Landy (1994) found motion and stereo (curvatures were different magnitudes, but in the same directions) to be weighted more equally. In their experiments, motion significantly affected the perceived amplitudes of contradictory patterns. In all three of these earlier sets of studies, the researchers found combination rules different from that found in the present experiments, although Bülthoff and Mallot (1988) have also observed veto situations. It would seem that different interactions can occur between motion and stereopsis, depending on the severity of the conflict. We used curvatures in orthogonal directions, whereas Rogers and Collett (1989), Tittle and Braunstein (1991), and Johnston et al. (1994) used similar curvature directions, but discrepant magnitudes.

While the results of the specific experiments reported here are consistent in general with modified weak fusion, other research exists which casts some doubt on the ability of one modality to “promote” or specify missing

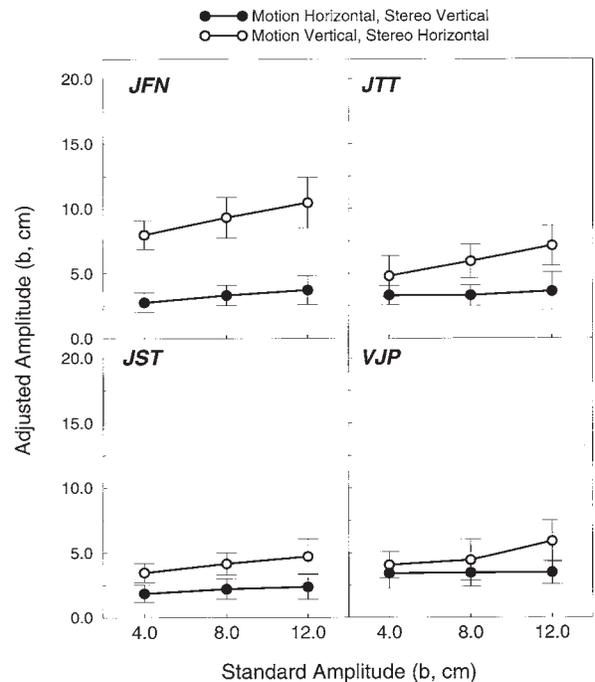


Figure 6. Results for the stereoscopic standard surfaces, presented separately for each observer. The adjusted amplitude necessary for the test surfaces' corrugations in depth to be perceived is plotted as a function of standard surface amplitude. The error bars represent ± 1 SD. The filled circles indicate that the stereo standard surfaces were vertically corrugated. The opaque symbols represent horizontally corrugated standard surfaces.

parameters that another modality needs to convert to a common format such as a depth map. This process would allow one to obtain a "more accurate overall estimate" by combining separate estimates (Landy et al., 1995). In a recent set of experiments by Norman, Todd, Perotti, and Tittle (in press), the perception of 3-D length was no better for optical patterns in which motion and stereo were combined than when they were presented separately. Rather, each observer's performance for the combined conditions was equivalent to the best individual modality. In addition, significant distortions of visual space were found even in full-cue, real-world conditions. Human observers may not be able to form an accurate depth map even when many different optical sources of information are available.

Finally, it may not be possible for the human visual system to promote representations derived from image shading to a format commensurate with other low-level representations. If human observers perceived the orientation of surface normals in a reasonably accurate manner, it would be possible (as Landy et al., 1995, have suggested) to integrate the surface orientations to recover a depth map that could then be used in a weighted average along with the other modalities' depth maps. Norman, Todd, and Phillips (1995) have shown for surfaces defined by shading that while the tilt component of surface orientation is perceived accurately, the slant component is not. If one integrated the observers' perceived surface orientations, a depth map would be obtained that would conflict with the depths obtained from other modalities. Koenderink, van Doorn, and Kappers (1992) have also observed this dichotomy in the human recovery of surface orientation between slant and tilt.

The present results are consistent with the more general psychophysical literature relating to conflict situations. Ever since Charles Wheatstone's (1838) discovery that binocularly disparate views lead to the perception of depth and solidity, there has been a general consensus that stereopsis ("solid" vision) is the primary determining factor in perceiving the 3-dimensional structure of the environment, and that other optical sources of information are secondary. Indeed, as late as the 1960s and 1970s, it was still a subject of great debate whether motion parallax, for example, was actually used by human observers to perceive depth and the shape of solid objects. In this context, it is important to note that researchers such as Gogel (1954), Gibson (as cited by Gregory, 1966, pp. 182-183), Braunstein, Andersen, Rouse, and Tittle (1986), Stevens and Brookes (1988), and Buckley and Frisby (1993) placed stereoscopic disparity in conflict with other sources of information about 3-D structure, such as occlusion (Braunstein et al. 1986; Gibson, as cited in Gregory, 1986) and texture gradients (Buckley & Frisby, 1993; Gogel, 1954; Stevens & Brookes, 1988). In all cases, it was found that the monocular sources could override stereopsis when they were in conflict. Our results also show that stereoscopic information about shape can be overcome by motion information in certain circumstances. It would appear, therefore, that while

stereopsis is an important determinant of perceived shape, it is only one of a multitude of other informative optical variables. Learning exactly how all of these multiple sources are integrated together into a coherent perception of our visual world will continue to challenge researchers for some time to come.

REFERENCES

- BLAKE, R. (1989). A neural theory of binocular rivalry. *Psychological Review*, **96**, 145-167.
- BRAUNSTEIN, M. L., ANDERSEN, G. J., ROUSE, M. W., & TITTLE, J. S. (1986). Recovering viewer-centered depth from disparity, occlusion, and velocity gradients. *Perception & Psychophysics*, **40**, 216-224.
- BRAUNSTEIN, M. L., HOFFMAN, D. D., & SAIDPOUR, A. (1989). Parts of visual objects: An experimental test of the minima rule. *Perception*, **18**, 817-826.
- BUCKLEY, D., & FRISBY, J. P. (1993). Interaction of stereo, texture, and outline cues in the shape perception of three-dimensional ridges. *Vision Research*, **33**, 919-933.
- BÜLTHOFF, H. H., & MALLOT, H. A. (1988). Integration of depth modules: Stereo and shading. *Journal of the Optical Society of America A*, **5**, 1749-1758.
- CLARK, J. J., & YUILLE, A. L. (1990). *Data fusion for sensory information processing systems*. Boston: Kluwer.
- CORNILLEAU-PÉRÈS, V., & DROULEZ, J. (1989). Visual perception of surface curvature: Psychophysics of curvature detection induced by motion parallax. *Perception & Psychophysics*, **46**, 351-364.
- DROULEZ, J., & CORNILLEAU-PÉRÈS, V. (1990). Visual perception of surface curvature: The spin variation and its physiological implications. *Biological Cybernetics*, **62**, 211-224.
- FOX, R., & CHECK, R. (1972). Independence between binocular rivalry suppression duration and magnitude of suppression. *Journal of Experimental Psychology*, **93**, 283-289.
- FOX, R., & PATTERSON, R. (1981). Depth separation and lateral interference. *Perception & Psychophysics*, **30**, 513-520.
- FUKUDA, H., & BLAKE, R. (1992). Spatial interactions in binocular rivalry. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 362-370.
- GOGEL, W. C. (1954). Perception of the relative distance position of objects as a function of other objects in the field. *Journal of Experimental Psychology*, **47**, 335-342.
- GREGORY, R. L. (1966). *Eye and brain: The psychology of seeing*. New York: McGraw-Hill.
- JOHNSTON, E. B., CUMMING, B. G., & LANDY, M. S. (1994). Integration of stereopsis and motion shape cues. *Vision Research*, **34**, 2259-2275.
- KOENDERINK, J. J., VAN DOORN, A. J., & KAPPERS, A. M. L. (1992). Surface perception in pictures. *Perception & Psychophysics*, **52**, 487-496.
- 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.
- LANDY, M. S., MALONEY, L. T., & YOUNG, M. J. (1991). Psychophysical estimation of the human depth combination rule. In P. S. Schenker (Ed.), *Sensor Fusion III: 3-D Perception and Recognition. Proceedings of the SPIE*, **1383**, 247-254.
- LEHMKUHLE, S., & FOX, R. (1980). Effect of depth separation on meta-contrast masking. *Journal of Experimental Psychology: Human Perception & Performance*, **6**, 605-621.
- MALONEY, L. T., & LANDY, M. S. (1989). A statistical framework for robust fusion of depth information. In W. A. Pearlman (Ed.), *Visual Communications and Image Processing IV: Proceedings of the SPIE*, **1199**, 1154-1163.
- NAKAYAMA, K., & SHIMOJO, S. (1992, September 4). Experiencing and perceiving visual surfaces. *Science*, **257**, 1357-1363.
- NORMAN, J. F., & LAPPIN, J. S. (1992). The detection of surface curvatures defined by optical motion. *Perception & Psychophysics*, **51**, 386-396.
- NORMAN, J. F., LAPPIN, J. S., & ZUCKER, S. (1991). The discriminability of smooth stereoscopic surfaces. *Perception*, **20**, 789-807.

- NORMAN, J. F., TODD, J. T., PEROTTI, V. J., & TITTLE, J. S. (in press). The visual perception of 3-D length. *Journal of Experimental Psychology: Human Perception & Performance*.
- NORMAN, J. F., TODD, J. T., & PHILLIPS, F. (1995). The perception of surface orientation from multiple sources of optical information. *Perception & Psychophysics*, **57**, 629-636.
- ROGERS, B. J., & COLLETT, T. S. (1989). The appearance of surfaces specified by motion parallax and binocular disparity. *Quarterly Journal of Experimental Psychology*, **41A**, 697-717.
- ROGERS, B. [J.], & GRAHAM, M. [E.] (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision Research*, **22**, 261-270.
- ROGERS, B. J., & GRAHAM, M. E. (1983). Anisotropies in the perception of three-dimensional surfaces. *Science*, **221**, 1409-1411.
- SCHOR, C. M. (1977). Visual stimuli for strabismic suppression. *Perception*, **6**, 583-593.
- SIRETEANU, R. (1982). Binocular vision in strabismic humans with alternating fixation. *Vision Research*, **22**, 889-896.
- SIRETEANU, R., & FRONIUS, M. (1981). Naso-temporal asymmetries in human amblyopia: Consequence of long-term interocular suppression. *Vision Research*, **21**, 1055-1063.
- STEVENS, K. A., & BROOKES, A. (1988). Integrating stereopsis with monocular interpretations of planar surfaces. *Vision Research*, **28**, 371-386.
- TITTLE, J. S., & BRAUNSTEIN, M. L. (1991). Shape perception from binocular disparity and structure-from-motion. In P. S. Schenker (Ed.), *Sensor Fusion III: 3-D Perception and Recognition. Proceedings of the SPIE*, **1383**, 225-234.
- TITTLE, J. S., & BRAUNSTEIN, M. L. (1993). Recovery of 3-D shape from binocular disparity and structure from motion. *Perception & Psychophysics*, **54**, 157-169.
- TODD, J. T., & NORMAN, J. F. (1991). The visual perception of smoothly curved surfaces from minimal apparent motion sequences. *Perception & Psychophysics*, **50**, 509-523.
- UOMORI, K., & NISHIDA, S. (1994). The dynamics of the visual system in combining conflicting KDE and binocular stereopsis cues. *Perception & Psychophysics*, **55**, 526-536.
- WHEATSTONE, C. (1838). Contributions to the physiology of vision. Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London*, **128**, 371-394.
- 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.

(Manuscript received August 22, 1994;
revision accepted for publication January 30, 1995.)