Available online at www.sciencedirect.com
SCIENCE DIRECT®
Vision Research 46 (2006) 1230–1241
Vision Research
www.elsevier.com/locate/visres

A reversed structure-from-motion effect for simultaneously viewed stereo-surfaces

Julian Martin Fernandez *, Bart Farell

Institute for Sensory Research, Syracuse University, Syracuse, NY 13244, USA

Received 22 July 2004; received in revised form 24 October 2005

Abstract

A spatially flat stimulus is perceived as varying in depth if its velocity structure is consistent with that of a three-dimensional (3D) object. This is structure from motion (SFM). We asked if the converse effect also exists. A motion-from-structure effect would skew an object's perceived velocity structure to make it more consistent with the 3D structure provided by its depth cues. This proposed phenomenon should be opposite in sign from velocity constancy and could potentially interfere with it. Previous tests of velocity constancy compared stimuli presented at different times, not simultaneously. This explains why a reversal of SFM has not been previously reported, as it is expected to appear only for simultaneous presentations. We tested this prediction using random-dot stereograms to define two adjacent moving surfaces separated in stereoscopic depth. We found that subjects did not perceive velocity constancy with either simultaneous or sequential stimulus presentations. For sequential presentations, subjects matched retinal speeds, in agreement with previous work. However, for simultaneous presentations, the nearer surface was seen as moving faster when both surfaces were moving with the same retinal speed, an effect opposite in polarity from velocity constancy and a signature of the motion-from-structure phenomenon.

Keywords: Velocity constancy; Structure from motion; Binocular disparity; Binocular vision; Speed discrimination; Motion from structure

1. Introduction

Structure from motion (SFM) is a non-veridical percept in which the retinal motion field of a 3D object is perceived as having a depth structure closely related to that of the object, despite its being stereoscopically flat. Thus, we can think of SFM as a process in which one visual dimension (velocity) modifies the perception of another dimension (depth), making it more consistent with the object properties implied by the first, inducing, dimension.

We ask here whether the SFM phenomenon can be reversed, that is, if the depth structure of a stimulus can affect its apparent speed, skewing the perceived motion to be more consistent with the 3D structure provided by the depth cues. We will call this hypothetical phenomenon "motion from structure" (MFS), with the understanding that it refers to a perceived modification of existing stimulus motion, not an illusory induction of motion into a stationary stimulus.

Fig. 1A illustrates SFM. Two fields are plotted here, one a field of depth values (given by disparity, shading, or texture gradient, for example) and the other a field of velocity values. The depth field is uniform, consistent with a flat, frontoparallel surface. The velocity field is non-uniform; speeds peak at the center of the display and fall progressively along the flanks. Non-rigid perceptual interpretations of the surface spatial structure are possible, but human observers generally prefer the SFM interpretation of a rigid rotating three-dimensional cylinder. In Fig. 1B, the curvatures of the depth and velocity fields have been switched; the velocity field is uniform, while the depth field is consistent with a cylindrical surface. The MFS interpretation is exemplified by perceiving the *velocity* field as peaking in the center. Of course, SFM does not imply that the perceived spatial structure is consistent with the physical

---

[*] Corresponding author. Tel.: +1 3154439714.
E-mail address: julian_fernandez@isr.syr.edu (J.M. Fernandez).

0042-6989/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2005.10.026

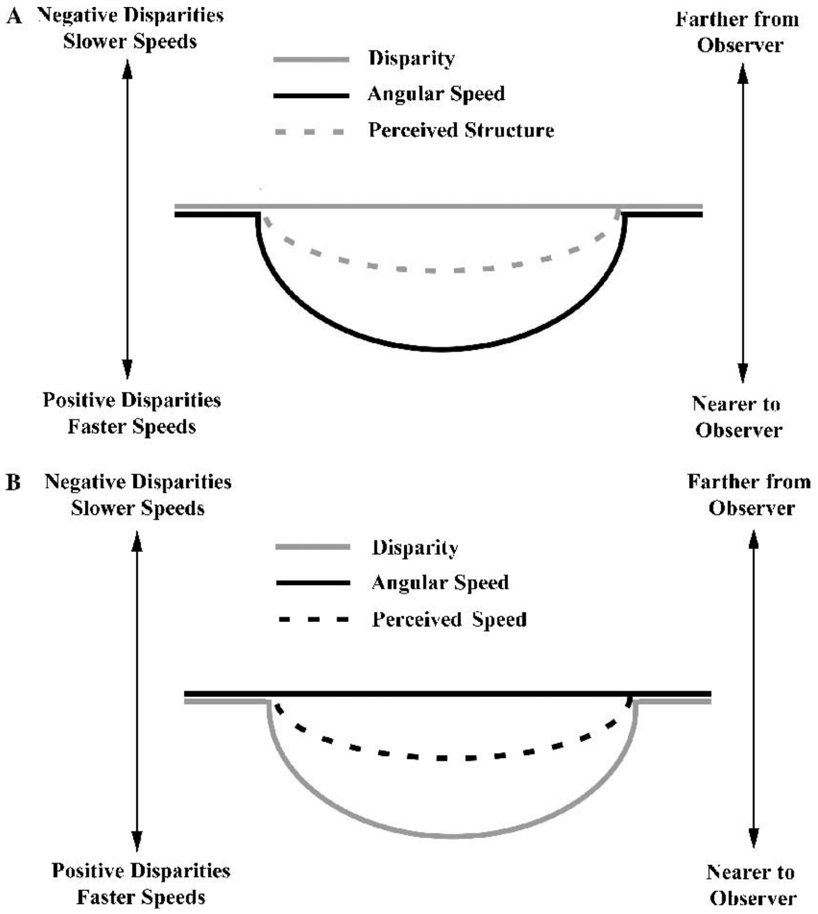


Fig. 1. (A) Structure from motion (SFM): the depth field (gray line) is consistent with a flat, frontoparallel surface. The velocity field (black line) is non-uniform; speeds peak at the center of the display and fall progressively along the flanks. Typically, the 3D SFM interpretation (dotted gray line) is that of a rigid rotating cylinder. (B) Motion from structure (MFS): the depth field (gray line) is consistent with a cylindrical surface. The velocity field (black line) is uniform. The MFS phenomenon is predicted to produce the perceived velocity field (dotted black line) as peaking at the center. Notice the symmetry, from (A) and (B) between SFM and MFS.

velocity field; nor does MFS imply that the perceived velocity structure is consistent with the physical disparity field. What is implied instead is a perceptual shift in the direction of consistency. Thus, the perceived surfaces, as drawn in Fig. 1, are compromises between the physical velocities and disparities in the display.

Both SFM and MFS need to be distinguished from velocity constancy (VC). The retinal velocity of a moving object varies with the frontoparallel component of the object's physical velocity and with the distance of the object from the observer. Yet two objects traveling at the same linear speed in frontoparallel planes at different distances from the observer appear to have the same speed, despite the difference in their retinal velocities (for a review, see Howard & Rogers, 1995). Velocity constancy holds in the presence of adequate cues to depth, such as changes in size, density or texture (Rock, Hill, & Fineman, 1968; Zohary & Sittig, 1993) or the presence of background reference frames (Epstein, 1978). However, velocity constancy is not always observed. McKee and Welch (1989) used binocular disparity as a cue to distance and found no evidence that it supports velocity constancy. In addition, Zohary and Sittig (1993) found that neither convergence nor accommodation supports velocity constancy when random-dot kinematograms were used as stimuli.

The finding that velocity constancy is not always observed can be turned to experimental advantage. As will be demonstrated later, MFS is expected to have the opposite sign as VC. Consequently, the two could potentially interfere with each other. By choosing disparity as the depth cue to probe for the existence of MFS, we are drawing on the previous evidence that disparity alone does not provide an adequate depth cue for VC. Thus, disparity will not generate a VC effect that could interfere with MFS.

In previous tests of velocity constancy, the stimuli to be compared were presented at different times, not simultaneously. Comparisons were either made across temporal intervals within a trial or between stimuli situated at different sides of the head, so that only one stimuli was seen at a given time. We are not aware of any study of velocity constancy involving simultaneous comparisons. This fact could explain why no MFS effect has been reported previously, even in the absence of VC. Structure from motion seems to require simultaneity to work, something that can be seen in the Ramachandran illusions with rotating coaxial cylinders (Ramachandran, Cobb, & Rogers-Ramachandran, 1988). Because of the analogies between SFM and MFS, we also expect that MFS might appear only in simultaneous stimulus presentations. This suggests (Fernandez, Watson, & Qian, 2002) but does not necessarily imply that the simultaneously presented stimuli being judged are processed as a single segmented unit, a proto-object in the sense defined by Pylyshyn (2003). Perception of a single rigid object might be necessary for linking the depth and velocity structures of a stimulus to each other, so that knowledge of one influences the observer's estimate of the other. Without single proto-object segmentation, depth and velocity structures would not be necessarily related, and knowledge of one would not be predictive of the other. This hypothesis, although not tested here, is consistent with psychophysical and neuropsychological evidence of object-based information access (for a review, see Pylyshyn, 2003, chap. 4).

In this study, we measured the perceived relative speed of stimuli whose depth structure was defined by binocular disparity. We used both simultaneous and successive stimulus presentations to test the predictions of the reversed SFM effect. We found that, for sequential presentations, subjects matched retinal speeds, in agreement with previous work. But for simultaneous presentations, the nearer of the two surfaces was seen as moving faster when both surfaces were moving with the same retinal speed. This effect, which demonstrates the existence of MFS, is opposite in polarity from velocity constancy.

2. Quantitative predictions

The terminology used in the rest of the article includes "retinal speed" and "angular speed", which will be used interchangeably. "Linear 3D speed" refers to the physical speed of a given object in 3D space. "Relative angular speed" is defined as the difference between two angular speeds normalized by one of those speeds and expressed as $\Delta v/v$.

In our experiments, data take the form of points of perceived speed equality as a function of the depth separation between two surfaces. These data are expressed in terms of the surfaces' relative angular speeds. The shape of this curve reveals whether judgments were based on angular speeds or linear 3D speeds (i.e., velocity constancy), or whether they deviate from both angular and linear 3D speeds. Thus, we can generically distinguish three possible outcomes for our experiment:

1. *Equal angular speeds.* Perceived relative speed agrees with the relative retinal speeds.
2. *Velocity constancy.* Perceived relative speed agrees with the relative linear 3D speeds. In this case, the closer surface will be seen to move slower than the farther surface when the relative angular speed is zero.
3. *Motion from structure.* Perceived relative speed agrees with neither retinal nor 3D linear speeds. The closer surface will be seen to move faster than the farther surface when the relative angular velocity is zero.

These alternatives are illustrated in Fig. 2. Represented here is the top view of two surfaces moving in frontoparallel planes at different distances from the observer, where Distance 1 < Distance 2. The observer is located at the vertex in each panel. The surface motions are drawn to show the physical speeds that yield equal perceived speeds for the two surfaces under each of the three hypotheses considered here: equal angular speeds, velocity constancy, and motion from structure. For equal angular speed perception (left panel), the linear 3D speed of the farther surface is larger than that of the nearer surface at the point of equal perceived speed, but the angular speeds (represented by the angles of the apertures) are the same. For velocity constancy (center panel), the linear 3D speeds (represented by the arrows' length) are the same at the point of equal perceived speed, but the angular speed is larger for the nearer surface. For motion from structure (right panel), at the point of equal perceived speed both the linear 3D speed and the angular speed of the farther surface are larger than those of the nearer surface.

The left and center panels are intuitively understandable, for they follow directly from the definitions of velocity constancy and equal angular speeds. The MFS prediction shown in the right panel is discussed in the next section. Let us first put the predictions for equal angular speeds and velocity constancy in quantitative form. We want to compute the actual relative angular speed of the stimuli when subjects perceive no difference in relative

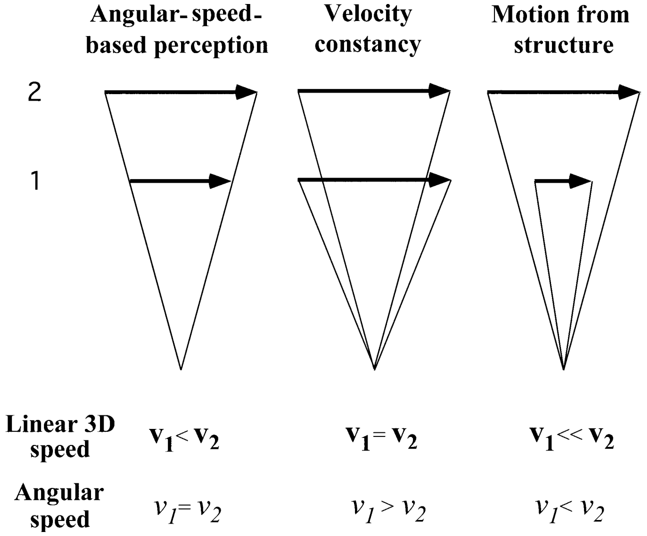


Fig. 2. Predictions of three models relating linear 3D speed ($\mathbf{v}$) and retinal (i.e., angular) speed (v). In this diagram it is assumed that the speeds of stimuli 1 and 2 are perceived to be equal. The length of each arrow gives its linear 3D speed, and the angular aperture gives its retinal speed. Stimuli 1 and 2 are 'near' and 'far' surfaces, respectively.

speed. Without loss of generality we will use only translations in our description of motion, because translations and rotations are mathematically equivalent for SFM calculations (Fernandez et al., 2002). Let us first develop some of the relationships between relative depth and the relative speed of translations.

The angular speed of a given point on a translating rigid object moving in the frontoparallel direction can be approximated by

$$v \approx \frac{T}{z}, \tag{1}$$

where v is angular speed, T is linear 3D speed, and z is the distance along the line of sight between the observer and the observed point on the surface (see Fig. 3). This approximation is valid for object sizes below ~10° (i.e., angles for which the orthographic projection approximation is valid). We want to calculate the relative angular speed between two points on the object at different depths, z and z_c. Developing Eq. (1) in Taylor series and keeping only the first order term on $\Delta z = z - z_c$, we obtain

$$\Delta v \approx -T\frac{\Delta z}{z_c^2}, \tag{2}$$

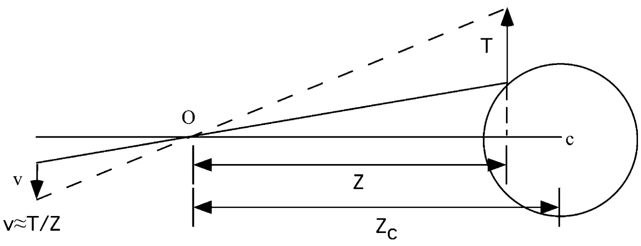
which gives the difference in angular speed between the two surfaces for $\Delta z \ll z_c$. From Eqs. (1) and (2) we obtain the relative angular speed between any two points on a rigid translating object (assuming $z \approx z_c$, which is valid in our approximation)

$$\frac{\Delta v}{v} \approx -\frac{\Delta z}{z}. \tag{3}$$

2.1. Angular-speed-based perception

This is the simplest case. If subjects respond in terms of angular speeds (AS), then, when they perceive zero relative speed between the two points (the *null* setting), the actual relative angular speed between those two points will also be zero. We represent this as

$$\left(\frac{\Delta v}{v}\right)_{\mathrm{AS}}^{\mathrm{null}} = 0. \tag{4}$$



v: retinal speed
T: component of 3d-speed on the frontoparallel plane
O: nodal point of eye

Fig. 3. Geometry of an object centered at c translating relative to the observer (top view). Left edge represents the retinal surface.

2.2. Velocity constancy

For velocity constancy, we want to find the relative angular speeds when subjects perceive a match between the linear 3D speeds. Humans observing velocity constancy perceive linear 3D speeds instead of retinal speeds, so they would report zero relative speed (the *null* setting) when the linear 3D speeds are the same. This happens to be the case in Eq. (3), which gives the relative angular speed when the 3D speeds are the same (i.e., a rigid translation). Thus, we can use Eq. (3) to get, for velocity constancy, the relative angular speeds when subjects perceive null relative speed

$$\left(\frac{\Delta v}{v}\right)_{\mathrm{VC}}^{\mathrm{null}} \approx -\frac{\Delta z}{z}. \tag{5}$$

2.3. Motion from structure

Motion from structure brings the perceived angular speeds closer to the speeds consistent with the surface structure specified by the depth cues. MFS is based on the assumption that the perceived angular speeds belonging to an object should deviate from veridical to be in closer agreement with Eq. (3). When the different velocities are perceived as belonging to different, independently moving objects, the departure from veridically perceived angular speed should not appear.

More concretely, for a translating object, the perceived relative angular speeds will be a weighted average of the angular speeds and the "depth consistent" speeds of the stimuli. The "depth consistent" speeds are the theoretical angular speeds consistent with a rigidly translating object whose shape is that specified by the depth cues (Eq. (3)). These speeds will be labeled as *depth*. The speeds labeled as *stimulus* correspond to the stimulus angular speeds. Thus, we have

$$\left(\frac{\Delta v}{v}\right)_{\mathrm{MFS}}^{\mathrm{perceived}} = \omega\left(\frac{\Delta v}{v}\right)_{\mathrm{MFS}}^{\mathrm{depth}} + (1-\omega)\left(\frac{\Delta v}{v}\right)_{\mathrm{MFS}}^{\mathrm{stimulus}}, \tag{6}$$

where $0 < \omega < 1$ is a weighting factor. Notice that Eq. (6) is a combination of terms that do (depth consistent speeds term) and do not (stimulus speed term) depend on the rigidity assumption. Thus, in general, the result will not be consistent with a rigid object. The rigidity assumption introduces a bias in the perceived speeds, but does not make them consistent with those of a rigid object. According to Eq. (6), perceived speeds will be consistent with those of a rigid object only if the stimulus is a rigid object, but in such a case near and far speeds will never match.

By definition, the *stimulus* speed is at the *null* setting when the perceived relative speed is zero. Thus, for the *null* setting, Eq. (6) becomes

$$\left(\frac{\Delta v}{v}\right)_{\mathrm{MFS}}^{\mathrm{null}} = -\frac{\omega}{1-\omega}\left(\frac{\Delta v}{v}\right)_{\mathrm{MFS}}^{\mathrm{depth}}. \tag{7}$$

But $\left(\frac{\Delta v}{v}\right)_{\mathrm{MFS}}^{\mathrm{depth}}$, the "depth consistent" speeds specified by the depth cues, can be obtained from Eq. (3), which gives the relative angular speed for two points on an object undergoing a rigid translation. Thus, using Eq. (3), Eq. (7) becomes

$$\left(\frac{\Delta v}{v}\right)_{\mathrm{MFS}}^{\mathrm{null}} \approx \frac{\omega}{(1-\omega)}\frac{\Delta z}{z}. \tag{8}$$

In Eq. (8), the predicted relative angular speed of the stimulus, for null perceived relative speed, has the opposite sign from that predicted by velocity constancy (Eq. (5)). Eq. (8) predicts that a nearer stimulus with a slower angular velocity matches a farther stimulus with a faster angular velocity. To make this result more intuitive let us look again at Fig. 1B. The black dotted line shows perceived speeds. Thus, in this example, MFS makes perceived speeds (black straight line) faster at points around the center of the cylinder (near points) than at points around the edges (far points). To make two given points, one near and one far, appear to be moving with the same speed, we need to diminish the angular speed of the near point relative to that of a far point; that is, we need to compensate in the opposite direction the effect of MFS. This is what is expressed by Eq. (8).

3. Methods

Using a chin rest to stabilize head position, observers viewed random-dot stereo-kinematograms (RDSK) through a mirror stereoscope. Stimuli in all experiments consisted of random configurations of square dark dots ($29\ \mathrm{cd\ m^{-2}}$, ~50% Weber contrast) moving inside a rectangular area and 'wrapping-around' to the opposite side after moving beyond the borders of the rectangle. Background luminance was $56\ \mathrm{cd\ m^{-2}}$. Because of an attenuator between the computer and the monitor it controlled, only the monitor's green gun was used; thus the color of both the dots and the background was green. The two surface regions to be compared had the same angular size, dot angular size and density, and differed only in disparity and speed.

Each RDSK comprised 20 frames presented at half the 75 Hz refresh rate of the monitor. This rate yielded smooth apparent motion. The stimulus duration was 533 ms. Number of dots per surface was 100 and dot size was 7.5′ on a side. The optical path was 95 cm. Subjects were instructed to maintain fixation at a central 15′ square black dot. The speed of the far surface was 10°/s and stimulus duration was 533 ms; speed discrimination has been shown to be near-optimal under these conditions (De Bruyn & Orban, 1988). Because they impair the quality of the speed signal, short-lived dots were not used. Instead, the individual dots survived the entire stimulus duration. However, our results show that subjects did not compare speeds by tracking dots. A dot-tracking strategy would produce responses based on retinal speeds; that is, no dependence of relative angular speeds on relative depth would be obtained.

3.1. Simultaneous presentations

The stimulus consisted of two contiguous rectangular areas of 4° (horizontal) × 2° (vertical) size, one above and the other below the fixation point (Fig. 4). All of the dots on the upper surface moved at the same constant speed; likewise, all the dots on the lower surface moved at the same constant speed. Dot speeds varied between the upper and lower surfaces and between trials, as explained below. Each trial consisted of a single presentation of the stimulus and observers were instructed to indicate in which surface, upper or lower, the dots moved faster. We determined the relative angular speeds that result in equal perceived speeds as a function of the relative depth (provided by binocular disparity) between the two surfaces. Relative angular speed is defined as

$$\frac{\Delta v}{v} = \frac{v_{\mathrm{far}} - v_{\mathrm{near}}}{v_{\mathrm{far}}}, \tag{9}$$

where v is angular speed.

To find perceptually equal stimulus speeds we used a double staircase procedure. One staircase started with retinal speeds for the near and far surfaces in a ratio of 1.35 (near surface faster) and the other staircase started with the retinal speeds in a ratio of 0.75 (near surface slower). Each staircase was controlled in "one-up, one-down" fashion: each time the subject responded "faster" to the near surface, the speed ratio was diminished by a factor of 0.02 and, conversely, when the subject responded "faster" to the far surface, the speed ratio was increased by a factor of 0.02. In both cases the speed of the dots on the far surface was kept constant and only the speed of the dots on the near surface was changed (the opposite case was also investigated and no difference in results was found). The speed of the far surface never changed across trials nor with changes in disparity, and was set to a value of 10°/s. After several trials the initial large difference between speeds diminished and eventually fluctuated

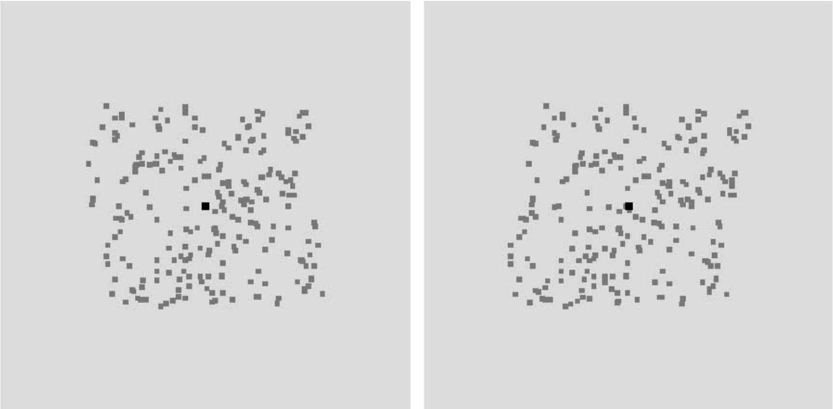


Fig. 4. Sample images of the stimulus used. When viewed stereoscopically, two planar rectangular surfaces situated at different depths can be perceived. The dots in both surfaces moved at constant speed on each surface in the horizontal direction. Speed, in general, was different in both surfaces. Both upper and lower surfaces had a size of 4° × 2°.

around a plateau of no perceived difference. Here the proportion of "faster" responses was evenly distributed between the two surfaces. The staircase ended after 80 trials or after 15 reverses, whichever took longer (although it never took more than 80 trials). The average speed of the near surface in the last 20 trials was then compared to the speed of the far surface to compute $\Delta v/v$ as defined in Eq. (9). Each plotted point corresponds to at least four runs, two for each initial-speed-difference ratio. The two staircases produced statistically equivalent results as assessed by a two tailed t test ($t = -0.41$, $p = 0.6821$, $\mathrm{df} = 310$).

During a given run of trials the disparity of the surfaces was held constant, but the closer surface could appear on top or bottom at random across trials. The surfaces were positioned in depth so as to straddle the fixation point, which was always midway between them. Motion direction was also randomly varied between leftward and rightward from trial to trial (but the same direction in both surfaces) to discourage anticipatory tracking. Subjects indicated in which surface, top or bottom, the dots were moving faster by moving a mouse to click labeled on-screen boxes that appeared after stimulus offset.

In addition, we carried out two parametric studies to check whether our results were specific to particular stimulus values. The first study tested multiple speed values for the far surface (2°/s, 5°/s, 10°/s, and 20°/s for S1; and 2°/s, 5°/s, and 10°/s for S6). The second study introduced a vertical gap separating the two surfaces (0°, 1°, 2.5°, and 4°, for both S1 and S6), with the speed of the far surface fixed at 10°/s. In both of these studies, the disparity difference between surfaces was set to 38′ and the stimulus duration was shortened to 160 ms. In other respects, the stimuli and procedures were identical to those described above.

3.2. Sequential presentations

Rather than display two abutting surfaces, the sequential presentation condition attempted to display the surfaces of two distinct objects. These surfaces were shown in sequence, each surface appearing in one of two intervals. In other respects the surfaces were like those of the simultaneous presentation condition. Stimulus duration was 533 ms for each interval, and the intervals were separated by 400 ms. The surface in the first interval was assigned at random one of the four possible combinations of near vs. far and top vs. bottom, and the surface in the second interval was assigned the complementary combination (e.g., if the surface in the first interval was assigned top and near, then the surface in the second interval was assigned bottom and far). The surfaces were positioned in depth such that the fixation point was always midway in depth between them. Motion direction was randomly varied between leftward and rightward from trial to trial (but the same direction was assigned to dots in both surfaces in a given trial). As a control, we repeated this experiment with the area of each surface doubled so the two stimuli occupied the same 4° × 4° retinal region across the two intervals. In all other respects we followed the same procedure as in the simultaneous presentations condition.

3.3. Controls

We also carried out several other control conditions, all with simultaneous stimulus presentations. To control for eye movements during stimulus presentation, we repeated the experiments, for two subjects, using 160 ms (6 frame) movies. And as a control for possible effects of subjects' attending to a particular depth or making anticipatory vergence eye movements, a random pedestal disparity was added to both surfaces. Thus, both surfaces could have positive disparity and appear behind the fixation point, negative disparity and appear in front of the fixation point, or one could be positive and the other negative, straddling fixation. The results turned out to be the same as when the pedestal disparity was zero.

3.4. Subjects

All subjects had normal or corrected-to-normal vision. Subjects included both authors, who were experienced psychophysical observers, and four other subjects who were naive as to the purpose of the experiments and had no previous experience in psychophysical tests. No feedback about the correctness of responses was provided to subjects in any of the experiments.

3.5. Perceived-depth calibration

In a separate series of measurements we assessed the magnitude of the perceived relative depth generated by the disparity of the stimulus. We used these perceived depth measures to calibrate the predictions of MFS and of velocity constancy.

Depth magnitude estimates were obtained by the method of adjustment. Subjects were presented with two planes separated in depth. The planes were identical to those used in the main experiment. After each presentation, subjects adjusted, with the click of a mouse, the separation in depth between the two planes, each plane being represented as a line (bird's-eye view), which had the same width (4°) as the planes previously seen.

At least 10 depth estimates were obtained for each depth separation. In separate conditions, stimulus duration were 533 and 160 ms. These durations correspond to those used in the main experiment. The simulated depth separations between the planes were the same as in the main experiment. The optical path was 95 cm. Subjects were instructed to maintain fixation at a central 15′ square black dot. Two subjects, S1 and S6, participated in the experiments.

Perceived relative depth ($\Delta z/z$), can be obtained from the adjusted separations in depth in the following way. We obtain from the adjustments the perceived depth-to-width ratio (μ) defined:

$$\mu = \frac{\Delta z}{\Delta X}, \tag{10}$$

where Δz is the perceived elongation in depth, and ΔX is the perceived width. Also, note that

$$\Delta X = z\Delta x, \tag{11}$$

where z is the perceived distance, and Δx the angular width. Thus, from Eqs. (10) and (11) we get the perceived relative depth as

$$\frac{\Delta z}{z} = \mu\Delta x. \tag{12}$$

In addition, we repeated the calibration using cylinders instead of planes. Subjects were presented with a series of vertically oriented elliptical cylinders which varied in their elongation in depth from trial to trial. The procedure was the same as that explained for planes. We used cylinders as a control because we knew from a previous study that subjects could perform the task with these stimuli. In a study of structure from motion, we presented subjects with the same task, but this time the cylinders' shape was defined only by motion parallax. Subjects matched shapes veridically in this task along the whole range of relative depths.

A two-way repeated measures ANOVA found no significant effect of stimulus type and no significant interaction between stimulus type and disparity. The data shown in Section 4 are those obtained using planes.

4. Results

4.1. Simultaneous versus sequential presentations

4.1.1. Simultaneous presentations

In the simultaneous-presentation condition, subjects were shown two contiguous surfaces defined by moving dots, one above the other. The relative angular speed (Eq. (9)) between these surfaces was measured at the point of subjective speed equality. Results for six subjects, distributed across two plots for clarity, appear in Fig. 5. Here, relative angular speed is displayed as a function of the rel-

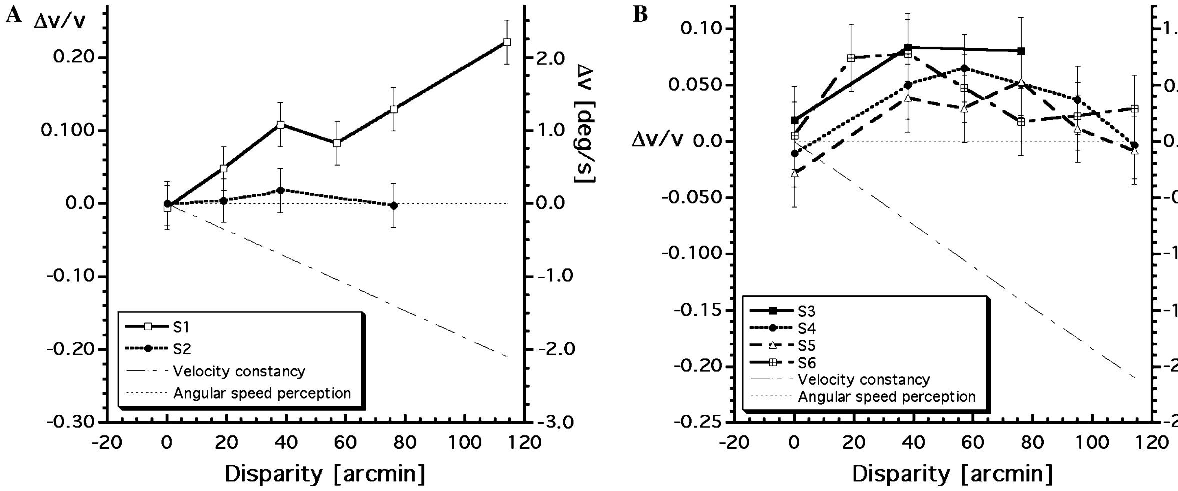


Fig. 5. Variation of relative angular speed at equally perceived speeds, as a function of the relative disparity between surfaces for the simultaneous-presentation condition. Dotted lines: predicted veridical relative angular speed perception. Dash-dotted lines: predicted velocity constancy under the assumption of veridically perceived depth. (A) Results for subjects S1 and S2. Observer S1 (squares) shows a strong deviation from veridical relative angular speed perception with the opposite polarity to that of velocity constancy. Observer S2 (circles) shows veridical relative angular speed perception. (B) Results for subjects S3, S4, S5, and S6. All subjects show a deviation from veridical relative angular speed perception opposite in polarity to that of velocity constancy and in the direction expected from MFS.

ative disparity between the surfaces. Also shown in Fig. 5 are the predictions for veridical angular-speed-based perception and for velocity constancy (this last is based on the assumption that there is a linear relationship between disparity and perceived depth; see, e.g., Fernandez et al., 2002, Eq. (6)).

Despite considerable variation across subjects, the matching angular speeds exceed the velocity constancy prediction for all subjects, and for all subjects but one they exceed the veridical angular speed perception prediction. The data of one subject (S2, Fig. 5A) are consistent with veridical angular speed perception. The data of another subject (S1, Fig. 5A) are consistent with MFS with a value of the slope close but with opposite sign to that of velocity constancy; the relative angular speed required for a perceived velocity match increases in direct proportion to the disparity between the surfaces. Data for the four other subjects (Fig. 5B) fall between those of S1 and S2 and show a common pattern: matching speeds increase at small disparities and then revert to veridical relative angular speed at large disparities. All the data points for intermediate relative depths are above the line showing the prediction for veridical relative angular speed and hence they are opposite in direction from velocity constancy. A repeated measures ANOVA confirm a statistically significant effect of binocular disparity for all subjects except S2, and S5 for whom significance is marginal (S1: $F(5,23) = 24.22$, $p < 0.001$; S2: $F(3,15) = 0.48$, $p = 0.7$; S3: $F(2,11) = 9.71$, $p = 0.013$; S4: $F(5,23) = 5.58$, $p = 0.004$; S5: $F(5,23) = 2.53$, $p = 0.074$; S6: $F(6,27) = 3.56$, $p = 0.016$).

The prediction for the perception of veridical angular speed is given by the horizontal line passing through the origin in Fig. 5; relative angular speed at the point of subjective speed equality is predicted to be constant across all disparity values. The prediction for velocity constancy shown in Fig. 5 assumes the standard linear relationship between relative disparity and relative depth. This gives a line, computed from Eq. (5), with negative slope anchored at $\Delta v/v = 0$ for a disparity of zero. The motion-from-structure prediction is a line with positive slope (Eq. (8)). However, the values of the slope for both velocity constancy and MFS depend on the perceived relative depth of the stimuli as a function of disparity. Non-veridical depth from disparity can arise from incorrect scaling for distance (Foley, Ribeiro-Filho, & Da Silva, 2004; Johnston, 1991) and from the lowered gain of stereoscopic depth at short stimulus presentations (Patterson, Moe, & Hewitt, 1992) and at disparities near and beyond the diplopia threshold (Ogle, 1952). These considerations suggest that perceived depth would deviate most from veridical depth at the upper range of disparities in our experiment. For velocity constancy and MFS, then, a quantitative comparison of data with predictions must await the perceived depth measures presented below.

4.2. Sequential presentations

In the sequential-presentation condition, the surfaces were presented in series either at different retinal locations—one above the other—or at the same retinal location with the stimulus size doubled. Results for these two sequential conditions are shown in Fig. 6 for the four subjects on whom data were collected. They all showed the same pattern: relative angular speed perception was veridical; speed matches were independent of disparity. A repeated measures ANOVA revealed no statistically significant effect of binocular disparity on matched relative angular speeds for any subject or condition. Moreover,

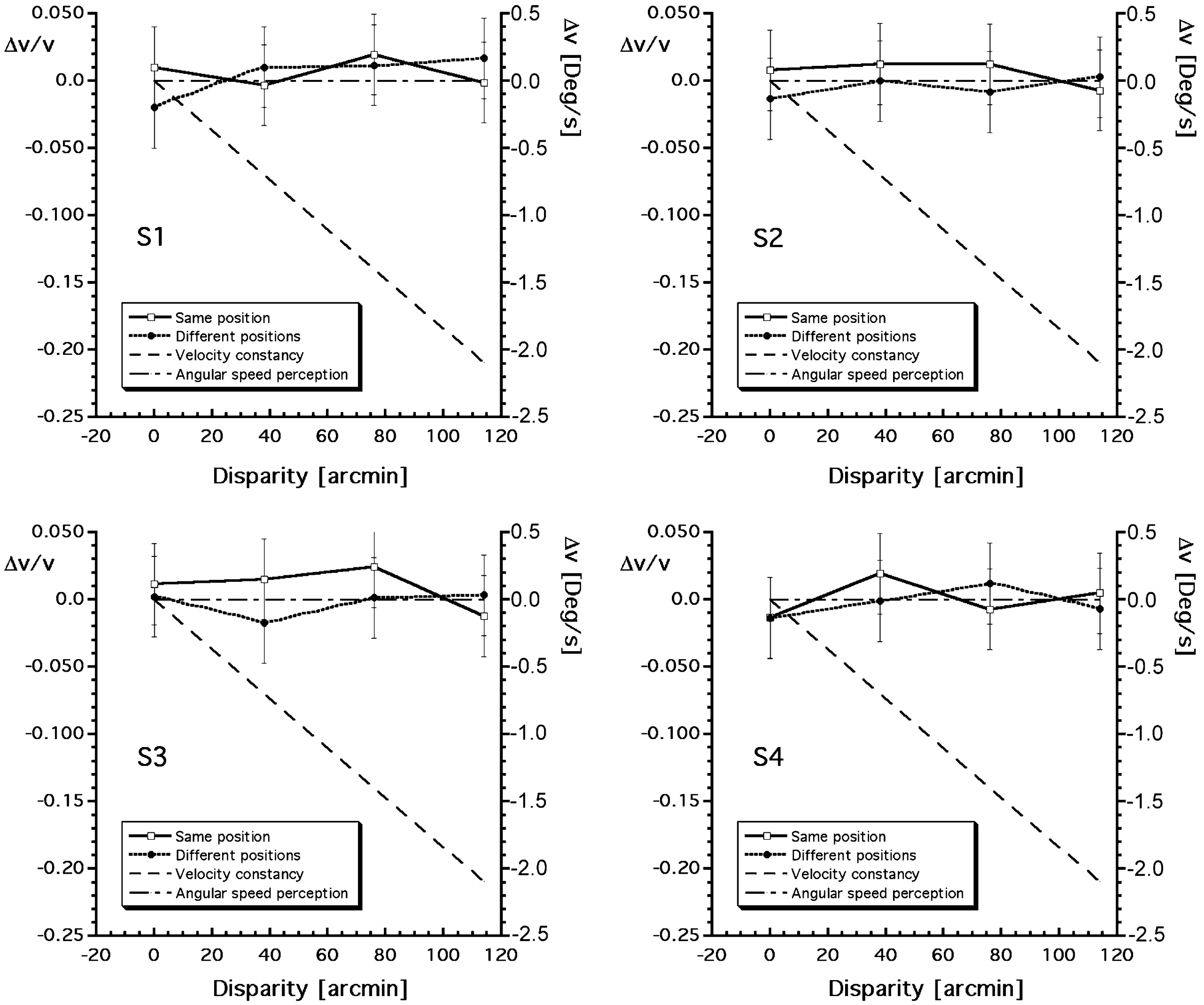


Fig. 6. Variation of relative angular speed at equally perceived speeds, as a function of the relative disparity between surfaces for the sequential-presentation condition. Squares and circles correspond to same and different retinal positions, respectively. Veridical relative angular speed perception was obtained in both cases for all four subjects. Dash-dotted lines: veridical relative angular speed perception. Dashed lines: velocity constancy prediction under the assumption of veridically perceived depth. Subject S1 is the one who showed the greatest deviation from veridical relative angular speed perception under simultaneous presentations (see Fig. 5A).

no bias was shown for the matched relative angular speeds: the average across disparities, subjects and conditions does not significantly differ from zero ($t = -1.0535$, $p = .300$, $\mathrm{df} = 31$). Also, data from none of the individual conditions or subjects were significantly different from zero. The lack of an effect of binocular disparity for sequential presentations differs from the outcome for simultaneous presentation shown in Fig. 5. It is the expected result for sequentially presented stimuli and is consistent with previous research using sequential presentations (McKee & Welch, 1989). Notice that the farther stimuli will have a larger perceived size than the nearer. Brown (1931) measured matching speeds using dot stimuli that differed in their physical sizes but were presented at different distances so that they had the same retinal size. He found that the angular speeds at the matching point were the same for large and small dots, consistent with our results for sequential presentations. Note that in Brown's experiment one stimulus was situated to the right of the observer and the other to the left, so they were never seen simultaneously.

None of our subjects showed evidence of velocity constancy for either simultaneous or sequential stimulus presentations. This result was not unexpected. Even though binocular disparity signaled relative depth, there were other cues in conflict with it (surface angular size, and dot size and density). These cues were consistent with a flat stimulus. However, for five of our six subjects, simultaneous horizontal stimulus motion was perceived with a reversed velocity constancy. For simultaneous stimulus presentations our results are at odds not only with velocity constancy, but also with veridical perception of angular speed found for sequential presentations.

4.3. Relative depth calibration

Fig. 7 shows perceived relative depth as a function of relative disparity measured in conditions similar to our

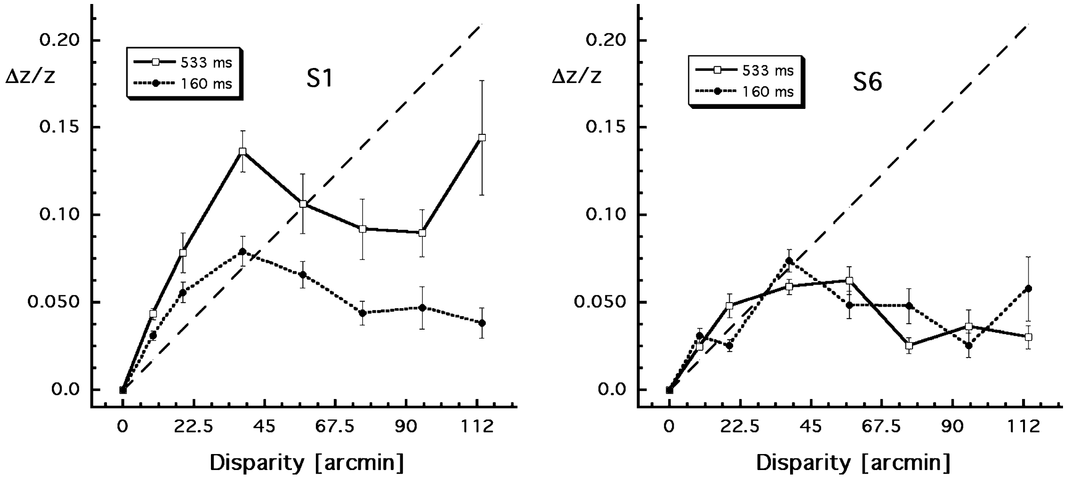


Fig. 7. Perceived relative depth as a function of the relative disparity present in the stimulus, for two different stimulus durations (circles: 160 ms, squares: 533 ms). Perceived relative depth was assessed in an adjustment task where subjects matched the relative depth of planar surfaces. Dashed lines: veridical relative depth perception. All error bars are SEM.

first experiment. We see that for small to moderate disparities (about 40′), perceived relative depth was approximately linearly related to disparity for both subjects. At larger disparities perceived depth either saturated or decreased as disparity increased. Shortening stimulus duration from 533 to 160 ms seems to have affected the estimates of relative depth above 60′ for subject S1, but seems not to have affected the estimates of subject S6.

In Fig. 8, we have replotted the data from Fig. 5 for these two subjects, but this time as a function of perceived relative depth rather than relative disparity. Data are shown separately for the two presentation durations. We can see that now the predictions of reversed SFM show a reasonably accurate match to the subjects' velocity perception and capture the variation of the matches as a function of perceived depth. In no case does perceived velocity follow the flat function predicted by angular-speed-based perception, or the negative values predicted by velocity constancy. Linear regression provided the fit to the data. In all four cases the value of the correlation obtained was significant, with all r^2 values exceeding 0.6: S1 (160 ms): slope $= 1.48 \pm 0.03$ ($r^2 = 0.68$, $t = 3.29$, df $= 5$, $p = 0.011$), S1 (533 ms): slope $= 1.24 \pm 0.04$ ($r^2 = 0.71$, $t = 3.11$, df $= 4$, $p = 0.018$), S6 (160 ms): slope $= 1.08 \pm 0.02$ ($r^2 = 0.68$, $t = 3.3$, df $= 5$, $p = .011$), S6 (533 ms): slope $= 0.85 \pm 0.02$ ($r^2 = 0.63$, $t = 2.93$, df $= 5$, $p = 0.016$).

From these slope values, we can obtain the values for ω, the weight factor of the depth-consistent speeds for computing perceived speeds in Eq. (8). For S1 these are $\omega = 0.59 \pm 0.01$ (160 ms) and $\omega = 0.55 \pm 0.01$ (533 ms); and for S6 they are $\omega = 0.52 \pm 0.01$ (160 ms) and $\omega = 0.46 \pm 0.01$ (533 ms). Hence, these subjects attached approximately equal weight to the two terms, angular speed and the depth consistent speed, of Eq. (6).

4.4. Parametric studies

Spurious motion matches could possibly produce artifacts that might account for the results. Mismatches might have arisen because of the fast dot speeds and small random dot surfaces we used. We also wondered if a spatial separation between the surfaces might mimic the temporal separation found in sequential presentations and would weaken or eliminate the MFS effect. Therefore, to further characterize the phenomenon and assess its dependence on particular values of dot speed and proximity, we examined the range of speeds and separations between the surfaces over which the effect holds.

In studying the effect for a range of speeds, we limited the study to a fixed disparity difference (38′) between surfaces. The size of the effect for this disparity difference will be taken as a measure of the strength of the MFS phenomenon. This value was chosen because it resulted in near maxima for both perceived depth (Fig. 7) and the size of the MFS effect. Fig. 9A shows the results for the two subjects tested in the parametric studies. The effect of speed on MFS strength was small for subject S6, unsystematic for subject S1, and not statistically significant for either (S1: $F(3,15) = 2.76$, $p = 0.1037$; S6: $F(2,11) = 0.656$, $p = 0.55$). The range of speeds tested was large (2–40°/s) and extended beyond the upper limit that could support task performance; at speeds faster than those shown in Fig. 9A, subjects had problems in perceiving coherent motion and in matching speed, due in part, presumably, to the small size of the test surfaces. Subject S1 could not do the task at speeds of 30°/s and above, and subject S6 at speeds of 15°/s and above.

Fig. 9B shows the effect of varying the gap between surfaces for the same two subjects. We can see that the effect decays quickly with separation (although only data for S1 reached statistical significance; S1: $F(3,15) = 11.21$,

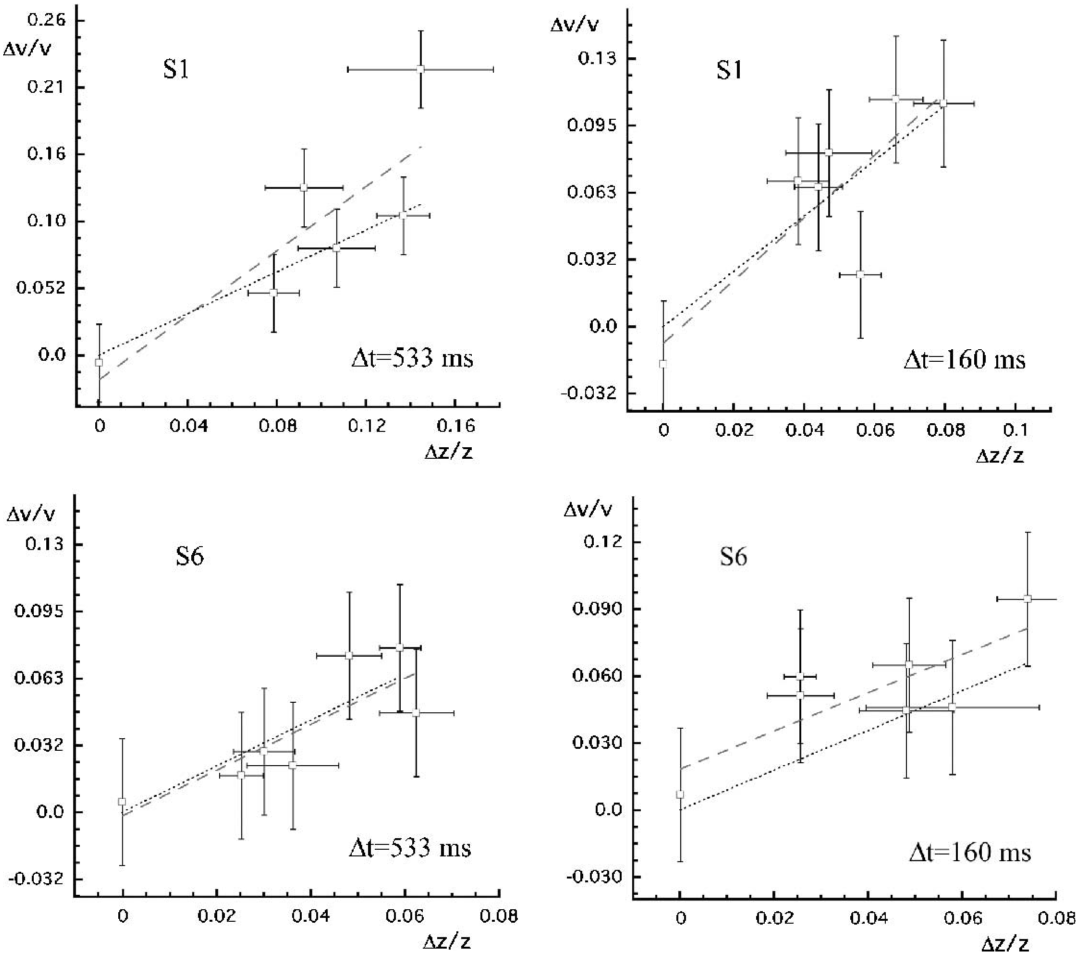


Fig. 8. Variation of relative angular speed at equally perceived speeds (squares), as a function of perceived relative depth. Dashed line represents linear regression fitting. Dotted line, shown for comparison purposes, represents an effect of exactly the same magnitude but of opposite polarity to velocity constancy (i.e., corresponding to a weight $\omega = 0.5$ (see Eq. (8))). Error bars are SEM.

$p = 0.0021$; S6: $F(3, 15) = 1.416$, $p = 0.3006$), and became negligibly small at a separation of 4°.

The MFS effect endured across a wide range of speeds, showing that the phenomenon is not the result of spurious matches. On the other hand, spatial proximity, like temporal coincidence, is necessary for the effect.

5. Discussion

Our data show that as a cue to relative depth, retinal disparity is not by itself sufficient to support velocity constancy. This holds true whether the two stimuli appear in the same or different retinal locations. In the case of non-simultaneous motion these results agree with those of McKee and Welch (1989), who also found no evidence that disparity alone supports velocity constancy.

The main focus of our work was to assess what happens in the case of simultaneous motion. Here, we found that perceived speed varies with stimulus depth in the direction *opposite* the velocity constancy direction. Thus a near surface must have a *faster* retinal angular velocity than a farther surface for the two surfaces to appear to move at the same speed. It is not possible to explain these results as an artifact of dot tracking, moving vernier judgments, or similar strategies. At most, these strategies would produce veridical angular speed perception; they would not produce the reverse of velocity constancy. Thus, our results conform nicely with the motion-from-structure hypothesis.

Our stimuli were random dots depicting frontoparallel surfaces. Testing for motion from structure with stimuli like the cylindrical surfaces shown in Fig. 1 would offer some advantages, for here the surface over which velocity varies is unambiguously that of a single object. However, using such stimuli would also present difficulties. A uniform angular velocity field with a cylindrical depth structure corresponds to a rotating cylinder in which the 3D speed is larger at the edges than at the borders. This is the opposite pattern of angular velocities from what is expected from MFS. Thus, a test that measures whether observers perceive a faster speed at the cylinder's center or at its edges will give opposite results depending on whether subjects base their judgment on angular speed or on 3D speed. It is unclear how subjects could be induced to use angular speeds without biasing the experimental out-

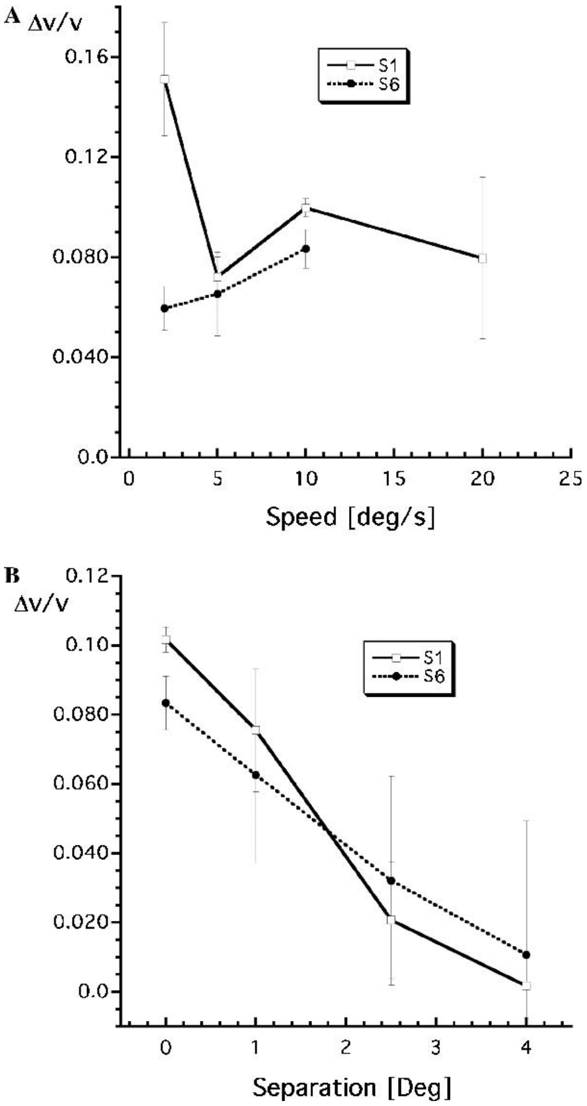


Fig. 9. (A) Effect of dot speed. The strength of the phenomenon does not vary strongly or systematically with speed over the range tested. (B) Effect of surface separation. The strength of the phenomenon decays as the gap between the surfaces increases. In both panels, error bars are ±1 SEM.

come. Using two cylinders and the task of selecting the one with the faster center speed also presents methodological problems, although we will not enter into detail here. To avoid the complication introduced by the in-depth component of the 3D speeds arising from the changing disparity cue, we considered it preferable to use a stimulus in which all speeds were frontoparallel.[1]

Because of this, our experiments used stimuli that allowed independent control of the depth and speed of each of the two surfaces and avoided judgments based on 3D speed. The surfaces were also kept adjacent and otherwise similar, to support, as far as possible, a perceptual interpretation of the surfaces as belonging to a single rigid object. As mentioned in the introduction, it is possible that perception of a single rigid object might be necessary for linking the depth and velocity structures of a stimulus to each other. However, it is unclear a priori whether this is necessary, or if the subjective appearance of the stimuli was sufficient to mimic the visual system's response to a single object undergoing translation. In any case, we obtained a rather good match with these stimuli to the predictions of MFS.

The difference in outcome between simultaneous and sequential presentations is important. As previously noted, the departure from veridical relative angular speed perception was hypothesized to appear only during simultaneous presentations. That is what we found. By analogy with SFM, MFS is expected to occur for contrasting speeds within an object viewed at a given time, and not across objects or times.

Summing up, we have obtained experimental support for a new phenomenon that we call motion from structure. This phenomenon inverts the cue dependence seen in structure from motion. In structure from motion, motion cues participate in computing the perceived depth structure. In motion from structure, depth cues participate in computing the perceived motion. In both SFM and MFS the final percept is a compromise between the cues present in the stimulus and inferences based on known properties of rigid objects.

Acknowledgments

We thank Fred Kingdom and the two anonymous reviewers for their very helpful comments and suggestions. This research was supported by NEI Grants F32 EY015673 (J.M.F.) and EY012286 (B.F.).

References

Brown, J. F. (1931). The visual perception of velocity. *Psychologische Forschung, 14*, 199–232.

De Bruyn, B., & Orban, G. A. (1988). Human velocity and direction discrimination measured with random dot patterns. *Vision Research, 28*, 1323–1335.

Epstein, W. (1978). Two factors in the perception of velocity at a distance. *Perception and Psychophysics, 24*, 105–114.

Fernandez, J. M., Watson, B., & Qian, N. (2002). Computing relief structure from motion with a distributed velocity and disparity representation. *Vision Research, 42*, 883–898.

Foley, J. M., Ribeiro-Filho, N. P., & Da Silva, J. A. (2004). Visual perception of extent and the geometry of visual space. *Vision Research, 44*, 147–156.

Howard, I. P., & Rogers, B. J. (1995). *Binocular vision and stereopsis*. New York: Oxford University Press.

Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. *Vision Research, 31*, 1351–1360.

---

[1] It is interesting to note that stimuli like those shown in Fig. 1 appear to observers to have a 3D speed at the edges that is almost parallel to the line of view (Fernandez and Farell, 2002, unpublished data). This suggests a perceptual adjustment that diminishes the frontoparallel component of the velocity near the edges relative to that at the center. Perhaps the adjustment is to the angular velocities and by extension to the horizontal component of the 3D velocities. But as we mentioned, the existence of in-depth components of motion makes this problematic to measure.

McKee, S. P., & Welch, L. (1989). Is there a constancy for velocity? *Vision Research, 29*, 553–561.

Ogle, K. N. (1952). On the limits of stereoscopic vision. *Journal of Experimental Psychology, 44*, 253–259.

Patterson, R., Moe, L., & Hewitt, T. (1992). Factors that affect depth perception in stereoscopic displays. *Human Factors, 34*, 655–667.

Pylyshyn, Z. W. (2003). *Seeing and visualizing*. Cambridge, MA: MIT Press.

Ramachandran, V. S., Cobb, S., & Rogers-Ramachandran, D. (1988). Perception of 3-d structure from motion: The role of velocity gradients and segmentation boundaries. *Perception and Psychophysics, 44*, 390–393.

Rock, I., Hill, A. L., & Fineman, M. (1968). Speed constancy as a function of size constancy. *Perception and Psychophysics, 4*, 37–40.

Zohary, E., & Sittig, A. C. (1993). Mechanisms of velocity constancy. *Vision Research, 33*, 2467–2478.