



Symmetry mediates the bootstrapping of 3-D relief slant to metric slant

Xiaoye Michael Wang¹ · Mats Lind² · Geoffrey P. Bingham¹

Published online: 9 September 2019
© The Psychonomic Society, Inc. 2019

Abstract

Empirical studies have always shown 3-D slant and shape perception to be inaccurate as a result of relief scaling (an unknown scaling along the depth direction). Wang, Lind, and Bingham (*Journal of Experimental Psychology: Human Perception and Performance*, 44(10), 1508–1522, 2018) discovered that sufficient relative motion between the observer and 3-D objects in the form of continuous perspective change ($\geq 45^\circ$) could enable accurate 3-D slant perception. They attributed this to a bootstrap process (Lind, Lee, Mazanowski, Kountouriotis, & Bingham in *Journal of Experimental Psychology: Human Perception and Performance*, 40(1), 83, 2014) where the perceiver identifies right angles formed by texture elements and tracks them in the 3-D relief structure through rotation to extrapolate the unknown scaling factor, then used to convert 3-D relief structure to 3-D Euclidean structure. This study examined the nature of the bootstrap process in slant perception. In a series of four experiments, we demonstrated that (1) features of 3-D relief structure, instead of 2-D texture elements, were tracked (Experiment 1); (2) identifying right angles was not necessary, and a different implementation of the bootstrap process is more suitable for 3-D slant perception (Experiment 2); and (3) mirror symmetry is necessary to produce accurate slant estimation using the bootstrapped scaling factor (Experiments 3 and 4). Together, the results support the hypothesis that a symmetry axis is used to determine the direction of slant and that 3-D relief structure is tracked over sufficiently large perspective change to produce metric depth. Altogether, the results supported the bootstrap process.

Keywords Bootstrap process · Geographical slant perception · Affine geometry · Stereomotion · Structure from motion · Skew symmetry

Statement of significance 3-D shape and slant perception are a fundamental part of the perception of the surrounding environment necessary to support action. Studying ways through which people can achieve perceived 3-D shape and slant accurately will not only shed light on how the human visual system works but also provide ideas that could be applied to, for example, computer vision. The current study offers empirical evidence suggesting how veridical slant perception can be achieved.

✉ Xiaoye Michael Wang
wang492@indiana.edu

- ¹ Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA
² Department of Information Technology, Uppsala University, Uppsala, Sweden

Many theoretical and empirical studies have characterized the perception of 3-D objects to be intrinsically affine, or of relief structure,¹ where the frontoparallel dimension of the object is perceived accurately while there is a homogeneous stretching or compression of distances orthogonal to the frontoparallel plane by an unknown scaling factor (e.g., Domini & Caudek, 2013; Koenderink & van Doorn, 1991; Todd & Bressan, 1990; Todd & Norman, 1991, 2003; Todd, Oomes, Koenderink, & Kappers, 2001; Wagner, 1985). Originally identified by Bingham and Lind (2008), and formalized in Lind, Lee, Mazanowski, Kountouriotis, and Bingham (2014), the bootstrap process proposes that large continuous

¹ In the literature, 3-D relief structure is often referred to as affine; however, the structure is not identical to a strict affine mapping as in the geometry of Klein's hierarchy. For a more detailed discussion, see Appendix A in Wang, Lind, and Bingham (2018).

perspective change enables perception of metric or Euclidean 3-D shape (i.e., the unknown scaling factor can be recovered). Large continuous perspective change, in this context, refers to the relative motion between the observer and object, which can be accomplished by, for instance, having the observer moving around a stationary object or rotating the object around a vertical axis relative to a stationary observer. A series of experiments was conducted to test judgments of depth-to-width aspect ratios viewed in elliptical cylinders and asymmetric polyhedrons using both real objects and computer graphics (Lee & Bingham, 2010; Lee, Lind, Bingham, & Bingham, 2012; Lee, Lind, & Bingham, 2013; Lind et al., 2014). Results from these studies confirmed the essential role of large continuous perspective change in enabling the recovery of metric 3-D shape.

The logic behind the bootstrap process is rather straightforward. It first assumes the recovery of 3-D relief structures through, for instance, existing structure-from-motion (SFM) algorithms (e.g., Koenderink & van Doorn, 1991; Lind, 1996; Shapiro, Zisserman, & Brady, 1995). Subsequently, the observer identifies certain 3-D depth structures that are quantitatively equivalent in physical (which is metric) and relief space at one moment in time. With continuous perspective change in the relief space, such equivalency would no longer hold: the relief depth structure would be constantly altered by the unknown scaling factor while the physical structure would remain unchanged. Under the rigidity assumption, one can relate the updated relief structure with the original one to recover the scaling factor. With the presence of measurement noise, the key is therefore to produce sufficiently large depth variations during such transformation to reliably derive the scaling factor.

Lind et al. (2014) proposed a right-angle solution to the bootstrap process. Because the 3-D objects perceived in a natural environment commonly subtend small visual angles and under small visual angle viewing, perspective projection can be adequately approximated using scaled orthographical projection, we used this approximation in our model (Shapiro et al., 1995; Thompson & Mundy, 1987). Under this approximation, one identifies two points on the object in the relief space that are equidistant to the observer, along with a third point that is on a line parallel to the line of sight. By definition, these three points form a 90° angle in the relief space. Because depth is scaled by the same unknown scaling factor, the initial two points should also be equidistant to the observer in the physical space and therefore the three points should form a right angle. Subsequently, with continuous perspective change, the angle formed among these points would no longer be orthogonal in the relief space, albeit still so in the physical space. Equating the current angle in the relief space with that in the physical space can help one recover the relief scaling factor. Because of measurement noise, a small amount of perspective change cannot generate sufficiently large depth variations in the right angle that are due to relief scaling. The authors suggested that at

around 45° of perspective change, which was observed in many studies as the place where judgment of aspect ratios became accurate, the line of sight bisects the right angle. Because angle bisection is allowed in relief space, 45° informs observers of a sufficiently large continuous perspective change that the bootstrapped scaling factor is accurate.

Wang et al. (2018) extended application of the model from 3-D shape to 3-D slant perception. In their first experiment, they presented 2-D rectangular planar surfaces defined by random dot texture. The slant angle formed between the surfaces and the xz -plane (i.e., the ground surface) remained constant while rotating around a vertical y -axis. They also tested the effects of three different types of visual information on the model, namely those of monocular SFM, pure stereomotion (CDOT, or change of disparity over time), and the two combined, which also gives rise to a second type of stereomotion, IOVD, or interocular velocity difference (for more detailed reviews of the two different kinds of stereomotion, see Allen, Haun, Hanley, Green, & Rokers, 2015; Cumming & Parker, 1994; Nefs, O'Hare, & Harris, 2010; Shioiri, Saisho, & Yaguchi, 2000). Results yielded a strong contrast between the monocular SFM condition and the two conditions containing stereomotion, but none of the conditions yielded accurate slant judgments even with 65° of continuous perspective change (which had yielded accurate aspect ratio judgments in previous studies).

The authors suggested that this discrepancy was produced by a lack of noncoplanar points on the 2-D planar surface used in the first experiment. The SFM models require at least four noncoplanar points to reconstruct an object's 3-D relief structure, implying that the strictly 2-D planar surface failed to produce good relief structure that then can be bootstrapped to metric. The authors subsequently introduced nine cuboids as bumps in a rectangular grid on the original planar surface and tested this in the monocular SFM and combined (monocular and stereo motion) conditions. With the addition of noncoplanar structure, judgment of slant remained inaccurate with small amounts of rotation, but became accurate at about 45° of continuous perspective change and remained so as perspective change increased beyond 45° to as much as 65°. Additionally, the difference originally found between the monocular and combined conditions disappeared: both now became accurate. Results from this study demonstrated that relief structures are necessary for the bootstrap process.

Wang et al. (2018) did not test the effect of the cuboid bumps in the pure stereomotion condition (CDOT only). This omission was because the stereomotion and combined conditions had initially exhibited identical performance in that study. However, in retrospect, there is a potential issue associated with the pure stereomotion condition—namely, that the displays failed to contain trackable texture elements. To eliminate monocular information, random dots are rerandomized and regenerated in every consecutive frame (Julesz, 1971). Given the presence of the cuboid bumps arranged in a rectangular grid, pure stereomotion displays would provide 3-D

features (i.e., vertices) arranged to form right angles that could be tracked over the rotation in the displays. In other words, the displays do not contain trackable points in the optical images while retaining right angles that are trackable in the 3-D structure, where the latter is presumed by the right-angle solution of the bootstrap process. Therefore, it still remains unclear whether the bootstrap process actually requires trackable 3-D points or whether trackable 3-D relief structures that contain right angles would be sufficient for this process. In Experiment 1, we addressed this question using the same object as used in Wang et al. (2018), presented using only stereomotion information.

An additional challenge to the original right-angle implementation of the bootstrap process also emerged in the context of slant perception. The essence of the right-angle implementation is twofold. First, as a result of relief scaling, the perspective change has to produce sufficiently large changes in the 3-D right angle (i.e., the angle would increasingly become an acute or obtuse angle because of the unknown stretching or compression along the depth dimension). Second, the perspective change also has to produce sufficient variation in the relationship between the line of sight and the angle (i.e., to signal that sufficient change has occurred, the line of sight

has to bisect the right angle at some instant during the rotation). Originally, Lind et al. (2014) investigated aspect ratio judgments, using polyhedral objects with a level top surface on which lay the right angles. In this case, continuous perspective change produced both sufficiently large change to the right angle's 3-D structure and the possibility that the line of sight would bisect the angle. However, with an upright slanted surface and under the assumption of orthographic projection, the initially identified right angle would not change its value as a result of relief scaling. This violates the first requirement of the right-angle implementation. In addition, the line of sight would never bisect the angle which violates the second requirement for the right-angle implementation to work. Essentially, what this means is that there is a lack of transformation to the recovered 3-D structure to enable observers to extrapolate the invariant scaling factor.

Figure 1 illustrates this dilemma. Right angles are found for two slanted surfaces, one with a geographical slant of 10° (flatter and closer to the ground plane) and the other at 80° (more upright), when the surface is facing the observer (0° geographical tilt). With rotation, the angles vary continuously due to relief scaling. However, the flatter 10° slant yields a much greater angle variation as a function of rotation than the

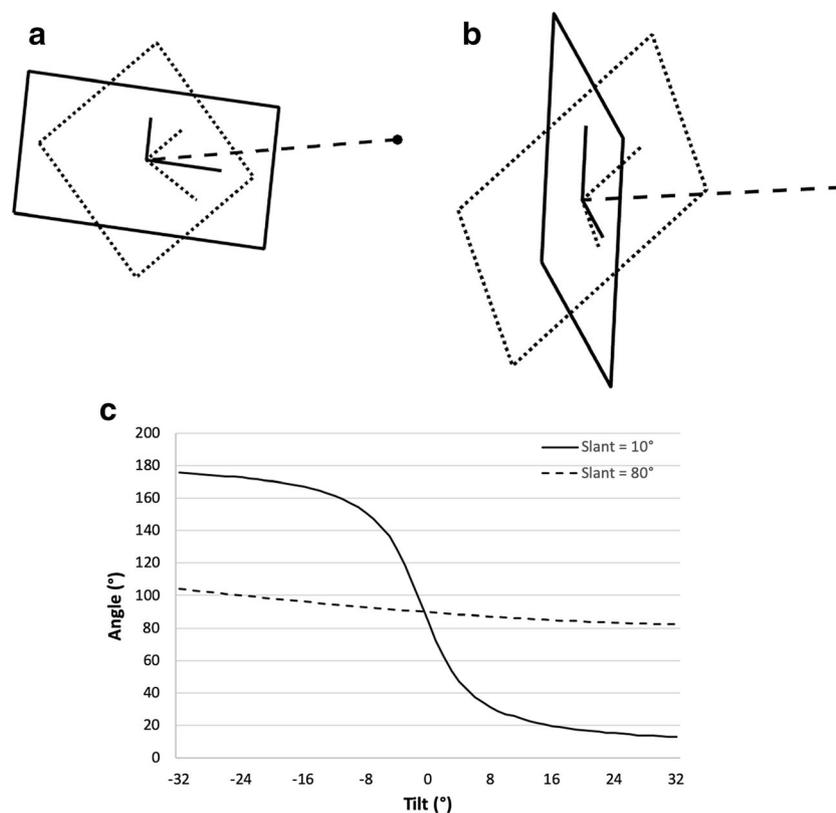


Fig. 1 Illustration of the right-angle dilemma. **a** surface at a slant of 10° with a right angle in its original orientation (solid lines) and after 22.5° of rotation around a vertical axis that goes through the center of the surface (dotted lines) with the line of sight (dashed line) congruent with one side of the right angle in its initial orientation. **b** The same setup as **a** but the

surface has a slant of 80° . **c** The actual 3-D angle of the original 3-D right angle (90° at 0° tilt) with relief scaling after rotation out to tilts of $\pm 32^\circ$ for surface slant of 10° (solid line) and 80° (dotted line). The 80° slant is more upright and closer to frontoparallel at 0° tilt

more upright 80° slant. With noise, the smaller angular variation of the more upright 80° slant is less effective in allowing the scaling factor to be recovered.

Nonetheless, as results in Wang et al. (2018) showed, 3-D slant perception still exhibits performance that is in accordance with the bootstrap prediction, where slants were judged inaccurately with small amount of rotation and became accurate when the amount of rotation was large. Therefore, this implies that there could be other bootstrap solutions that are more suitable to slant perception. Still, it is essential to explicitly rule out the possibility that using right angles is the sole method to perform the bootstrap process. In Experiment 2, we used a symmetrical hexagon as the base surface with nine randomly located tetrahedrons to provide additional noncoplanar points. The objects used in this experiment was carefully designed to not contain any intrinsic right angles in its 3-D structure. We presented these objects using either pure stereomotion or the combination of stereomotion and monocular SFM. If the right-angle solution is the only solution for the bootstrap process, performance in the stereomotion condition should be poor, even at large rotation amounts, compared with that in the combined condition, since the former contained neither trackable right angles nor trackable texture elements. Eventually, results from this experiment suggested observers could still bootstrap the correct scaling factor despite of a lack of right angles. Based on this finding, we proposed an alternative implementation of the bootstrap process, which was further illustrated in Wang, Lind, and Bingham (2019).

Finally, we also noted that the stimuli used thus far to test the bootstrap model all contained mirror symmetry. According to Pizlo and colleagues, 3-D symmetry could be a strong a priori constraint on 3-D shape perception (Bingham & Muchisky, 1993a, 1993b; Li, Sawada, Shi, Kwon, & Pizlo, 2011; Pizlo, 2010; Pizlo, Sawada, Li, Kropatsch, & Steinman, 2010). As Saunders and Knill (2001) showed, symmetry also plays a significant role in slant perception. For a symmetrical planar surface whose symmetry is described using a vertical symmetry axis and symmetry lines that connect symmetrical points around the symmetry axis, the angle between the symmetry axis and lines after projection can constrain the interpretation of slant and tilt of the surface (see Fig. 3 of Saunders & Knill, 2001). Therefore, in Experiment 3, we tested whether eliminating mirror symmetry in the slanted surface, with an asymmetrical pentagon surface, would affect slant judgments in the context of the bootstrap process.

If the elimination of symmetry was found in Experiment 3 to interfere with the ability to bootstrap 3-D metric slant with large rotation, this could be due to one of two reasons. First, symmetry may be critical to the bootstrap process and the lack of it might disrupt the recovery of the correct scaling factor with large continuous perspective change. However, this possibility seemed unlikely given that Lind et al. (2014) used asymmetrical polyhedrons in their aspect ratio judgment task

and found that observers could still recover the correct aspect ratios with sufficiently large rotation. Second, symmetry—and in particular a symmetry axis—could potentially provide information regarding the direction of slant. In previous experiments, symmetry axes coincided with the direction of slant and the lack of it might have made the slant direction ambiguous. This interpretation does not rule out the fact that observers could still recover the correct scaling factor with the bootstrap process; it simply means that the recovered scaling factor was used to recover the wrong slant. To test this alternative, in Experiment 4, we perturbed the direction of the symmetry axis of hexagonal objects used in Experiment 2 to see if this would affect performance.

Experiment 1

In Experiment 1, we examined whether large continuous perspective change would enable accurate slant perception given only pure stereomotion information about rectangular surfaces with a rectangular array of cuboid bumps. Stimuli were the same as the ones used in Wang et al. (2018), with the only difference being the type of visual information available to the perceiver. With pure stereomotion, the optical texture was not trackable over the rotational motion in the displays. Instead, only features in the perceptible 3-D structure were trackable. This experiment aimed to explore whether the bootstrap process could operate on the level of 3-D relief structures, instead of identifying and tracking 2-D texture points.

Method

Participants

Eleven adults, between 20 and 30 years of age (two males and nine females), participated in this experiment. All participants provided their informed consent prior to participating in the experiment with approval from Indiana University's Institutional Review Board (IRB). They were either paid \$10/hour or compensated with course credits. All participants had normal or corrected-to-normal eyesight and also passed a stereo fly test (Stereo Optical Co., Inc.) that measured their stereo acuity. Participants had to be able to identify the target circle indexed by a disparity of 80 seconds of arc to be included.

Stimuli and apparatus

Stimuli were presented on a Dell UltraSharp U2312HM 23-inch monitor (51 × 29 cm) with a resolution of 1,920 × 1,080 and a refresh rate of 60 Hz. We used MATLAB 2011b (the MathWorks Inc., Natick, MA, 2011) to generate and present the displays with the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli,

1997). Slant was varied randomly between 27° and 73° by 2° increments, yielding 24 different slants. We defined slant angles in terms of the angles formed between the surface normal and the axis of rotation, which was the positive y -axis. Consequently, tilt is defined as the angle between the projected surface normal on the xz -plane and the z -axis. We did not use the line of sight for this as do most slant perception studies. If slant was defined as a function of the line of sight, it would be constantly changing during continuous perspective change. Defined relative to the vertical y -axis, slant remains constant under rotation around that axis. Note, when the tilt is 0° (i.e., when the surface normal is perpendicular to the positive x -axis), the slant defined with respect to the line of sight was 90° minus the slant defined with respect to the y -axis. The range of slants was chosen to avoid extreme slants of 0° and 90° . The surface was rectangular and 10-cm wide with heights of 8, 10, or 12 cm (to control for the use of foreshortening). On the surface were nine cuboids that were 1 cm in length and width and 0.55 cm in height, arranged in a three-by-three grid. A schematic illustration of the surface is shown in Fig. 2.

Each slanted object was rotated around a horizontal x -axis at the frontoparallel position to create slant. To generate continuous perspective change, the object rotated around a vertical y -axis that passed through its center. We used five different amounts of rotation: 25° , 35° , 45° , 55° , and 65° . The object rotated at a constant speed of $20^\circ/\text{s}$. Each trial, the object started at a position slanted directly away from a frontoparallel plane and rotated to the left and right by half of the amount of rotation (e.g. for a 25° rotation, the object would first rotate to the left by 12.5° , go back to the frontoparallel position, and then rotate to the right by 12.5°). The object would continue to rotate throughout the trial until a judgment was made.

We used red–blue random dot anaglyphs to generate the displays with pure stereomotion information available only. To do this, we first set up a 3-D environment with the slanted object placed 9 cm behind the screen (i.e., the projection surface) and a background 18 cm behind the screen. The point of observation was 76.2 cm in front of the screen, as was the viewing distance for participants in the actual experiment. Given this viewing geometry, the background surface was approximately 9.1° of visual angle both horizontally and vertically. The horizontal size of the projected object was 6.7° , whereas its vertical size varied as a function of slant angle and

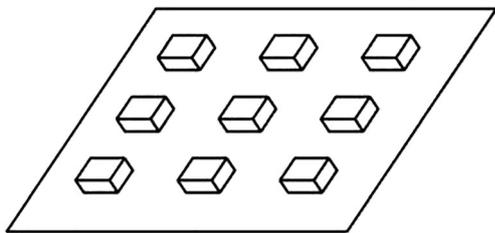


Fig. 2 Schematic of the noncoplanar object used in Experiment 1. Note that the actual display used random dots

surface length, approximately ranging from 2.1° (slant angle 27° and surface length 8 cm) to 7.7° (slant angle 73° and surface length 12 cm).

To construct the display, we first populated the screen with 6,000 single pixel texture elements. The distance between the left and right projection points was set at 6 cm as the inter-pupillary distance (IPD). Using perspective projection, each point on the screen was projected through these projections onto different surfaces, and subsequently back-projected onto the screen to yield disparity. Points from the left eye were marked as red, whereas those from the right were marked as blue. Each frame of the display contained a uniform distribution of red and blue dots without any monocular information. Additionally, dots were rerandomized between every frame to further eliminate any monocular information. Therefore, only the evolving disparity information was preserved across frames. Hidden line removal was used to treat the occlusion of the background by the object as well as occlusion of the slanted surface by the cuboids.

Similar to Wang et al. (2018), to avoid too many trials, we did not fully cross the surface height manipulation with slant angles and rotation amounts. Instead, we randomly assigned the three possible surface heights to three consecutive slant angles in the range of tested slants (e.g., 27° , 29° , and 31°). Consequently, there were a total of 120 trials in this experiment with 24 different slant angles crossed with five different amounts of rotation.

Procedure

After participants provided informed consent and passed the stereo fly test, they sat in front of the computer screen wearing a pair of red–blue filter glasses. The experimenter instructed the participants how to perform the task. Each trial, they would first observe a rotating object for one full cycle (i.e., tilt 0° , with surface normal pointed toward the observer, to the left extreme, and to the right extreme, and back to 0°). Then, a 2-D response line appeared on the screen to the right of the display. Using the left and right arrow keys on the keyboard, participants adjusted the orientation of the line so that it matched the slant of the object (as in Todd, Christensen, & Guckes, 2010; Wang et al., 2018). Although studies have demonstrated that perception of 2-D lines tends to be biased (Dick & Hochstein, 1989; Durgin & Li, 2011), Cherry and Bingham (2018) tested the effectiveness of using a 2-D line as a response measure for 3-D slant judgment. They compared slant judgment using the line response figure with adjusting an actual surface, and they did not find a difference between the two methods. The object continued to rotate back and forth throughout the entire trial while participants were making their judgments. They were asked to press the space bar to confirm their answer and to continue to the next trial. We presented the rotation amounts sequentially from 25° to 65° to avoid the

possibility that large perspective change would calibrate the judgment performed with smaller perspective changes. Within each rotation condition, objects with slant angles were randomly displayed.

Data analysis

We performed multiple regression to examine the effects of variations of surface height on slant judgment. We found that the effect of surface height is extremely small and therefore we ignored it in the subsequent analyses as we had in the previous study. We performed linear regressions with the actual slant as the independent variable and the perceived slant as the dependent variable for every rotation amount for each participant. We subsequently used the regression slopes, intercepts, and r^2 as our measures of performance and entered them in repeated-measure analysis of variance (ANOVA) with the rotation amount (five levels) as repeated-measures factor. Additionally, we were also interested in examining the actual values of the regression slopes and intercepts. For each rotation amount, a slope of 1 and an intercept of 0 represented veridical performance in that block. We used 95% within-subject confidence interval around the mean, which was specifically designed for repeated-measure designs (Cousineau, 2005; with correction by Morey, 2008) to compare regression slopes with 1 and intercepts with 0.

Results and discussion

Using multiple regression, we evaluated the effects of surface height on judgments. The independent variables were slant angle, rotation amount, and surface height. Surface height accounted for a very small portion of the total variance (3.7%, with Cohen's $f^2 = 0.04$). So, we did not include it as an additional variable in the subsequent analyses. We performed linear regressions for each participant within each rotation amount. Figure 3 show the average regression slopes, intercepts, and r^2 for each rotation amount.

For regression slopes, ANOVA showed that there was a significant main effect of rotation amount, $F(4, 40) = 7.63$, $p < .001$, $\eta_p^2 = 0.43$. Pairwise comparison showed that only the 25° rotation condition was significantly different from all other rotation conditions ($p < .05$), whereas the others were not significantly different from one another. For regression intercepts, Mauchly's test of sphericity showed that the sphericity assumption was violated, $\chi^2(9) = 19.11$, $p < .05$. With Greenhouse–Geisser correction, ANOVA on regression intercepts showed a significant main effect of rotation amount, $F(2.02, 20.18) = 6.03$, $p < .01$, $\eta_p^2 = 0.38$. Post hoc pairwise comparison showed that, similar to regression slopes,

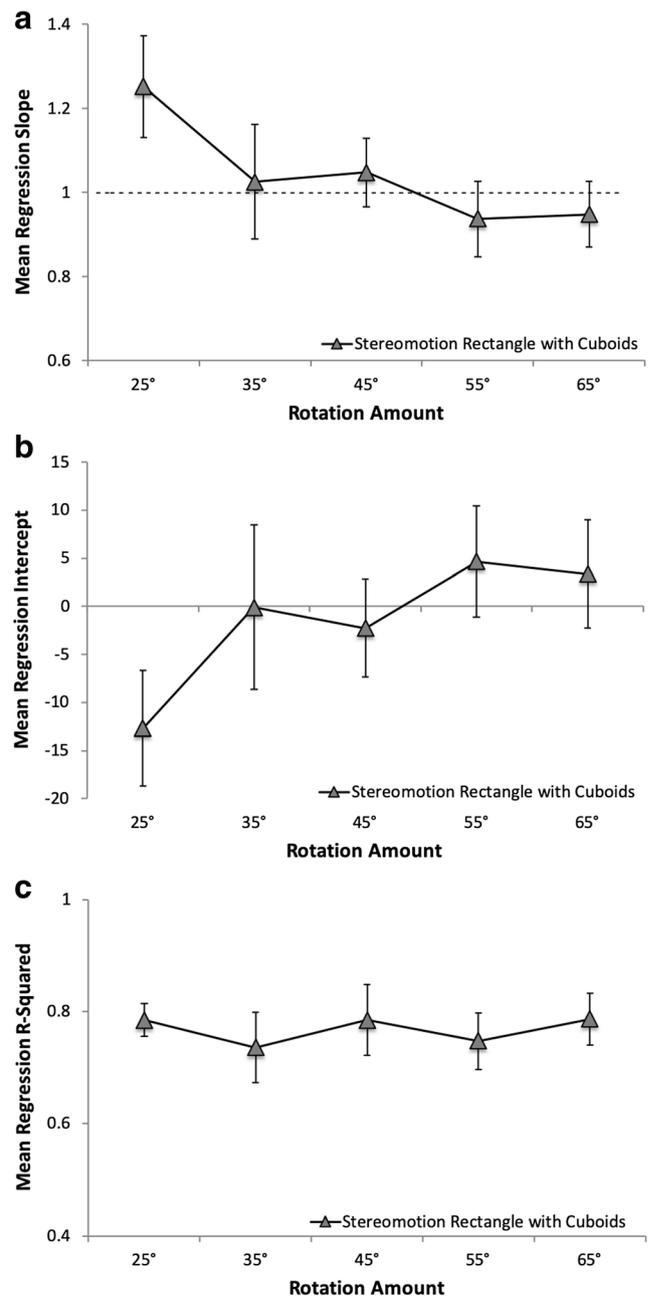


Fig. 3 Mean regression (a) slopes, (b) intercepts, and (c) r^2 in Experiment 1 for pure stereomotion condition of the slanted surface with 3-D cuboids laying on a grid, plotted as a function of rotation amount. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008)

intercepts in the 25° rotation condition were significantly different from those in other conditions ($p < .05$), while those in other conditions were not significantly different from one another.

Subsequently, we looked at the actual values of the regression slopes and intercepts using the 95% within-subject

confidence interval around the mean. As Fig. 3a–b show, at 25°, the confidence interval for regression slopes did not include 1, nor did the regression intercepts include 0. This suggested that performance was inaccurate with a small amount of rotation. Based on the actual values of the mean regression slope and intercept, we can see that participants tended to judge large slants to be more upright while small slants to be flatter. However, as rotation increased to 35° and beyond, regression slopes became indistinguishable from 1 and intercepts from 0. This suggested that participants achieved veridical performance at 35° of rotation and such performance remained relatively stable as the amount of rotation further increased. Such patterns of results confirmed predictions from the bootstrap process (Wang et al., 2019) and matched the slant judgment results from Experiment 2 in Wang et al. (2018) where the authors presented the same objects using either monocular SFM or the combination of SFM and stereomotion.

Finally, an ANOVA on r^2 showed that there was no significant effect of rotation amount ($p > .30$, $\eta_p^2 = 0.10$). As shown in Fig. 3c, regression r^2 remained relatively stable across rotation amounts. This suggested that although performance was inaccurate at small rotation (i.e., 25°), participants were consistent in their judgment.

Results from Experiment 1 showed that, in a condition devoid of trackable texture elements (that is, the pure stereomotion condition), performance was equivalent to that in conditions with trackable texture elements (that is, the monocular SFM condition and the combined monocular SFM and stereomotion condition). This demonstrated that the bootstrap process does not require trackable texture elements, and instead functions using trackable 3-D relief structure. This finding implies that the bootstrap process generally works only with 3-D relief structure, not 2-D optical inputs. This is in accordance with the findings in Wang et al. (2018), where failures in obtaining relief structure impeded the bootstrap process.

If the bootstrap process is truly relevant, we should test the role of available 3-D right angles. The object tested in Experiment 1 as well as that used in Experiment 2 of Wang et al. (2018) consisted of cuboids arranged in a rectangular grid on a rectangular surface, all providing 3-D right angles (meaning right angles are inherent to the 3-D relief structure of the object). The original implementation of the bootstrap process proposed in Lind et al. (2014) relies on the identification and tracking of right angles on the object. However, as mentioned earlier, this right-angle solution is merely one potential way that the bootstrap algorithm could be implemented. In addition to its stringent requirement (i.e., identifying three points that form a right angle or having inherent right angles in the object), more upright surfaces could also pose a challenge to this method. Nonetheless, it is essential to investigate whether eliminating both right angles and trackable surface texture would render the large continuous perspective change ineffective in enabling the perception of 3-D metric slant.

Experiment 2

In Experiment 2, we eliminated intrinsic right angles formed among vertices in the object structure by substituting a hexagonal surface for the rectangular one, and tetrahedrons at random locations on the surface instead of cuboids in a rectangular grid. To further explore the effects of trackable 3-D features versus trackable texture elements, we tested this setup using both pure stereomotion and the combination of stereomotion and monocular SFM.

Method

Participants

Twenty-two adults, between the ages of 20 and 30 years participated in this experiment. There were four males and seven females in the combined condition, and four males and seven females in the stereomotion condition. All participants provided their informed consent prior to participating in the experiment with the approval from Indiana University's Institutional Review Board (IRB). They were compensated with course credits. All participants had normal or corrected-to-normal eyesight and also passed a stereo fly test, as in Experiment 1.

Stimuli and apparatus

Experimental apparatus and setup were the same as in the previous experiment. For the stimuli, we used a symmetrical hexagon instead of a rectangle (see Fig. 4). To control for the potential effects of differences in size, we constructed the hexagon so that its maximum width was 10 cm, and its maximum height varied among 8 cm, 10 cm, and 12 cm. Given the viewing geometry, the horizontal dimension of the projected object was 6.7°, and its vertical dimension ranged from 2.1° to 7.7°, which was equivalent to the size of the stimuli used in Experiment 1 and those in Wang et al. (2018). Additionally, we introduced nine tetrahedrons at random locations on the pentagon to avoid any explicit 90° angles among them. Each tetrahedron had a height of 0.55 cm and each triangular side had a length of 1 cm. Thus, the projected dimension of the tetrahedrons was approximately 0.37°. Additionally, the locations of the tetrahedrons were selected so that they did not block one another throughout rotation for different slant angles. We presented the stimuli with pure stereomotion generated using the same method as in Experiment 1. In addition, for the combined display, we constructed each surface using the appropriate dot density based on surface area, instead of back projecting dots on the screen, and did not rerandomize dots for each frame.

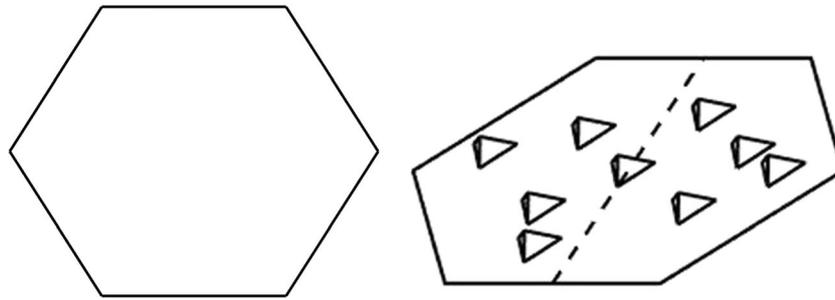


Fig. 4 Schematic of the shape of the hexagon surface used in Experiment 2 (left) and layout of the 3-D object (right). The surface displayed here has a width of 10 cm and a height of 10 cm. The dashed line on the actual

object represents the object's symmetry axis. Note that the actual displays were presented using random dots, and the dashed line was not displayed

Procedure

We used the same experimental procedure as in Experiment 1.

Results and discussion

We first evaluated the effects of surface height using multiple regression. Surface height again only accounted for a small portion of the total variance (5.6%, with Cohen's $f^2 = 0.059$) and was not included in subsequent analysis. For each visual information condition, we regressed judged slant onto actual slant for each participant for each rotation amount. Figure 5 show the mean regression slopes, intercepts, and r^2 for the combined and stereomotion conditions. Subsequently we ran mixed-design ANOVA for the slopes, intercepts, and r^2 with rotation amount as a within-subject factor (five levels) and visual information as a between-subject factor (two levels).

An ANOVA on regression slopes showed that there was a significant main effect of rotation amount, $F(4, 80) = 5.75, p < .001, \eta_p^2 = 0.22$, but no significant main effect of visual information ($p > .90, \eta_p^2 = 0.00$) or a significant interaction effect ($p > .80, \eta_p^2 = 0.018$). Post hoc analysis with LSD adjustment showed that regression slopes at 25° and 35° were not significantly different from each other, but they were both significantly different from those in 45°, 55°, and 65° rotation ($p < .05$). An ANOVA on regression intercepts showed there was a significant main effect of rotation amount, $F(4, 80) = 10.06, p < .001, \eta_p^2 = 0.34$, but no significant main effect of visual information ($p > .60, \eta_p^2 = 0.01$) or a significant interaction effect between the two ($p > .30, \eta_p^2 = 0.06$). Post hoc LSD correction showed that regression slopes did not differ between 25° and 35° conditions, but both rotation conditions were significantly smaller than the other three conditions ($p < .001$) while there was no significant difference among the latter three conditions.

Next, we evaluated the quality of the performance using 95% within-subject confidence intervals around the mean for regression slopes and intercepts. As can be seen from Figure 5 (a) and (b), confidence intervals for regression slopes did not include 1

for both combined and stereomotion conditions at 25° and 35° of rotation, corresponding to intercepts not including 0 in those two conditions. This suggested that slant judgments were poor with small amounts of rotation for both stereomotion and combined conditions. However, at 45° of rotation and beyond, the confidence intervals for slopes included 1 and those for intercepts included 0. This indicated that slant judgment became veridical starting at 45° of rotation and remained relatively stable as rotation further increased. Again, the results confirmed the prediction from the bootstrap model (Wang et al., 2019).

Finally, an ANOVA on regression r^2 did not show a significant main effect for either rotation amount ($p > .50, \eta_p^2 = 0.04$) or visual information ($p > .10, \eta_p^2 = 0.10$) or a significant interaction effect ($p > .60, \eta_p^2 = 0.03$). As Fig. 5c shows, there was no pronounced trend of change in r^2 as a function of rotation, where they all remained relatively stable. However, there did appear to be a trend for difference between the two visual information conditions, where r^2 for the combined condition was slightly higher than that for the pure stereomotion condition, even though statistical analysis did not return a significant difference between them. Because both conditions had comparable regression slopes, r^2 in this context could be considered as a measure of relative judgment of consistency within a rotation amount (i.e., how noisy the actual judgment for a participant was). Since the combined condition introduced additional visual information to the display, it is reasonable to expect a difference in judgment consistency between the two conditions.

Results from Experiment 2 suggested that right angles were not necessary for the bootstrap process. Performance in both visual conditions remained poor with small amounts of rotation but became veridical at 45° of rotation and beyond. In the pure stereomotion condition in particular, there was neither trackable 3-D texture elements that could potentially form right angles or any right angles inherent to the object and participants could still recover metric slant with large continuous perspective change. This indicated that the bootstrap process was performed with this set of stimuli but not necessarily with the right-angle solution. So, what could the bootstrap process be here?

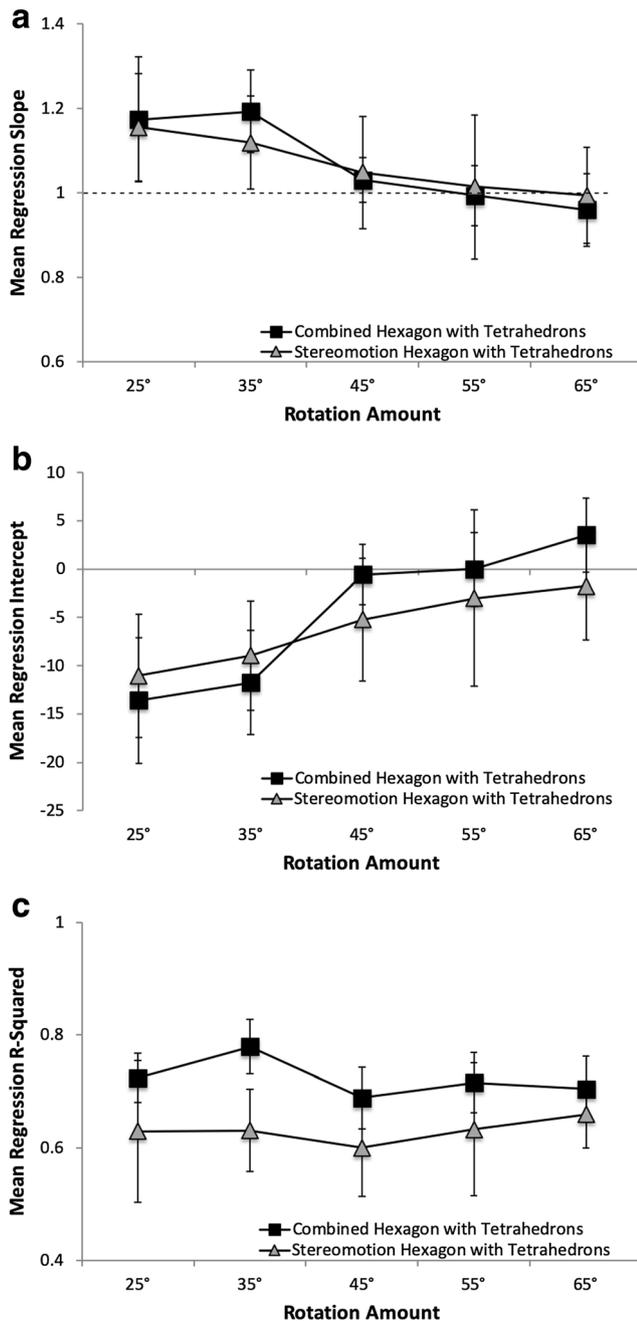


Fig. 5 Mean regression (a) slopes, (b) intercepts, and (c) r^2 for the hexagon display with tetrahedrons in Experiment 2, presented using stereomotion or combined visual information. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008)

Revision of the bootstrap process

Upon a closer examination of the original right-angle solution, we noted that the equidistant points that made up one leg of the right angle could potentially be used to extrapolate the relief scaling factor. By definition, each pair of equidistant points are in a frontoparallel plane and the distance between

them can be accurately perceived. As the object rotates, the two points would no longer be in a frontoparallel plane and their distance would be subject to relief scaling. Following a similar rationale as presented in the original right-angle solution, observers can track the 3-D distance between these points to bootstrap the correct scaling factor.

Specifically, assume two points on a slanted object that are equidistant to the observer, at t_0 they have the coordinates: $P_{1t_0}, (x_{1t_0}, y_{1t_0}, z_{1t_0})$, and $P_{2t_0}, (x_{2t_0}, y_{2t_0}, z_{2t_0})$. By definition, the length of the line formed between P_1 and P_2 is simply the distance between two points in the frontoparallel plane:

$$L = \sqrt{(x_{1t_0} - x_{2t_0})^2 + (y_{1t_0} - y_{2t_0})^2}. \quad (1)$$

With relative motion between the observer and the object, e.g. motion from rotation around an axis in the xy -plane, P_1 and P_2 will no longer be equidistant to each other or be in the same frontoparallel plane. Let $l(t)$ be the line's orthographically projected 2-D length in the image plane at any given time t_1 :

$$l(t) = \sqrt{(x_{1t_1} - x_{2t_1})^2 + (y_{1t_1} - y_{2t_1})^2}. \quad (2)$$

We know that at t_0 , the orthographically projected 2-D length is the same as the actual length of the line: $l(t) = L$. The change in l between two consecutive times can be expressed as:

$$\dot{l} = \frac{d(l(t))}{dt}. \quad (3)$$

Then, the unknown scaling factor, i.e. the angular velocity of rotation, at any given time can be expressed as:

$$\dot{\alpha}(t) = \frac{\dot{l}}{\sqrt{L^2 - l(t)^2}}, \quad (4)$$

which can be computed using only x and y coordinates of the two equidistant points. When noise is added to the system, measuring the projected length of the line could be variable. However, with an increasing amount of perspective change, the projected length would become increasingly smaller, and therefore is increasingly less affected by the addition of noise. Therefore, with larger amounts of rotation, $\dot{\alpha}$ estimations should also become more accurate. As shown in Fig. 6, starting at a rotation of 0° , which is when the object is facing the observer, the bootstrap process produces highly variable and inaccurate predictions of the scaling factor. As the amount of rotation increases, scaling factor estimation becomes more accurate and less variable.

In addition, as observed for the right-angle solution, because of the relief scaling, observers are uncertain about the amount of rotation that has taken place. The right-angle solution used angle bisection as the way to identify when sufficient

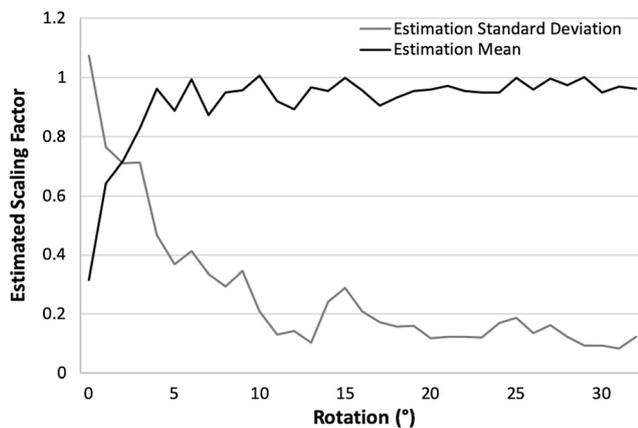


Fig. 6 Estimated scaling factor and standard deviations of estimations as a function of rotation with 5% Gaussian motion noise added. 0° rotation corresponds to when the surface is facing the observer. The accurate scaling factor is 1

rotation has occurred (i.e., when the line of sight bisects the right angle at 45° of rotation). For the equidistant-point solution, observers could use the change in 3-D distance between the equidistant points as a criterion to evaluate the amount of perspective change. Norman, Todd, Perotti, and Tittle (1996) found that the Weber fractions for discriminating 3-D distance was approximately between 19% to 26%, which, in the current context, translate to a continuous perspective change of between 36° to 42°. This range corresponds to the originally identified 45° with the distinction that it is not a set number, but a range of values that might be dependent upon individual observer and the object. This allows for performance to potentially improve by as little as 35° of perspective change, as found for the rectangular surface with cuboids. The transformation used in this method embodies the essence of the bootstrap process. It tracks depth structure, in this case the equivalence in depth of two points, and, with rotation, bootstraps the correct scaling factor based on the transformation of such structure due to perspective change. Wang et al. (2019) presented a more detailed account on this new solution, along with simulation results that demonstrated its effectiveness in accounting for behavioral results.

With a new implementation of the bootstrap process, the next question to address was whether symmetry plays a role in bootstrapping metric slant. We noted that both types of stimuli in Experiments 1 and 2 contained mirror symmetry. Based on work by Pizlo and colleagues, we know that symmetry could be a constraint on perceiving 3-D objects (Li et al., 2011; Pizlo, 2010; Pizlo et al., 2010). In addition, Saunders and Knill (2001) demonstrated that symmetry and the resulting skew symmetry after projection play a role in perceiving local slants when given 2-D images of the surface. Therefore, in Experiment 3, we eliminated mirror symmetry in the stimuli and explored whether slant judgments would change accordingly.

Experiment 3

In Experiment 3, we removed mirror symmetry in the object by using an asymmetrical pentagon surface, scattered with the same tetrahedrons at random locations as in Experiment 2. We again presented the stimuli using either pure stereomotion or the combination of stereomotion and monocular SFM.

Method

Participants

Twenty adults participated in this experiment. There were four males and six females in the combined condition, and four males and six females in the stereomotion condition. All participants provided informed consent prior to participating in the experiment with the approval from Indiana University's Institutional Review Board (IRB). They were either paid \$10/hour or compensated with course credits. All participants had normal or corrected-to-normal eyesight, and also passed a stereo fly test as in previous experiments.

Stimuli and apparatus

Experimental apparatus and setup were the same as in the previous experiment. For the stimuli, we used an asymmetric pentagon as the base of the slanted object. The pentagon had a maximum width of 10 cm and maximum height varied among 8 cm, 10 cm, and 12 cm. Given the viewing geometry, the horizontal dimension of the projected object was 6.7°, and its vertical dimension ranged from 2.1° to 7.7°. The same tetrahedrons as in Experiment 2 were added to the surface as a means to introduce noncoplanar points. They were set at random locations so that to avoid right angles in the stimuli. Figure 7 offers a schematic illustration of the shape of the pentagon used in this experiment. We presented the stimuli either with pure stereomotion or with the combination of stereomotion and monocular SFM generated using the same method as in Experiment 2.

Procedure

We used the same experimental procedure as in Experiments 1 and 2.

Results and discussion

The initial multiple regression showed that the surface height accounted for a small portion of the total variance (1.4%, with Cohen's $f^2 = 0.01$), and consequently was not included in subsequent analysis. We performed linear regressions for each participant within each rotation amount and each visual

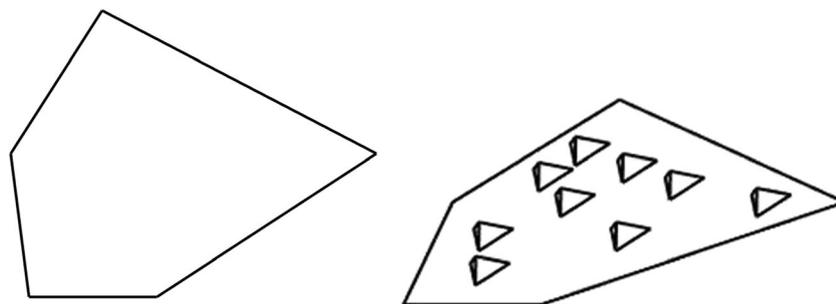


Fig. 7 Schematic of the shape of the pentagon surface used in Experiment 2 (left) and layout of the 3-D object (right). The surface displayed here has a width of 10 cm and a height of 10 cm. Note that the actual displays were presented using random dots

information condition, and evaluated performance using regression slopes, intercepts, and r^2 (see Fig. 8).

An ANOVA on regression slopes showed that there was a significant main effect of rotation amount, $F(4, 72) = 10.57$, $p < .001$, $\eta_p^2 = 0.37$, but not of visual information, $F(1, 18) = 3.07$, $p = .097$, $\eta_p^2 = 0.15$, or of the interaction ($p > .90$, $\eta_p^2 = 0.01$). Pairwise comparison showed that slopes in the 25° condition were significantly greater than those in other conditions except for at 35° ($p < .05$), while 65° rotation condition were significantly smaller than those in other conditions ($p < .05$). An ANOVA on regression intercepts showed that there was a significant main effect of rotation amount, $F(4, 72) = 13.26$, $p < .001$, $\eta_p^2 = 0.42$. However, there was no main effect of visual information ($p > .40$, $\eta_p^2 = 0.04$) or a significant interaction ($p > .90$, $\eta_p^2 = 0.01$). Pairwise comparisons showed that intercepts at 25° of rotation were significantly smaller than those in the other conditions ($p < .05$) and intercepts at 65° rotation were significantly greater than those in other conditions ($p < .05$).

To evaluate the accuracy of performance, we again looked at the 95% within-subject confidence interval around the mean for regression slopes and intercepts. As shown in Fig. 8a–b, regression slopes did not reach 1 and regression intercepts did not reach 0 with as much as 55° of rotation for both stereomotion and combined conditions. At 65° rotation, regression slopes reached 1 and intercepts reached 0 only for the combined condition, but not for the stereomotion condition. This was in stark contrast with the pattern of results in the previous experiments, where performance became veridical (i.e., regression slope equaled 1 and intercept equaled 0) at either 35° or 45° rotation and remained so with subsequent increase in rotation. In other words, slant judgment with the pentagonal objects remained relatively poor throughout rotation. However, there were some noticeable changes in performance as a function of increasing rotation amount, as seen from the ANOVA results. Evaluating regression slopes and intercepts based on the within-subject 95% confidence interval did suggest a trend for a difference between the two visual

information conditions (nonsignificant in the ANOVA), where the combined condition provide slightly better performance as might be expected given the additional information available.

Finally, the ANOVA on r^2 yielded no significant main effect of rotation amount ($p > .05$, $\eta_p^2 = 0.12$), of visual information ($p > .70$, $\eta_p^2 = 0.005$), or an interaction ($p > .90$, $\eta_p^2 = 0.01$). As shown in Fig. 8c, regression r^2 remained relatively stable for both visual information conditions.

Results from Experiment 3 did not replicate the pattern that was expected from the bootstrap process. This failed for both the stereomotion and the combined conditions. Such a failure could be due to one of two reasons. First, it could be because the bootstrap process failed to extrapolate the correct scaling factor. With everything being equal, the only difference between Experiments 2 and 3 was the presence of mirror symmetry in the object. In other words, this interpretation would suggest a necessary role of mirror symmetry in the bootstrap process. However, Lind et al. (2014) used asymmetric polyhedrons to test aspect ratio judgments, where the polyhedrons' top cross section was the same as the pentagons used in the current experiment. Judgment in their study still became veridical when there was sufficiently large continuous perspective change. This indicated that the first interpretation was unlikely. The alternative explanation is that the direction of slant was based on the symmetry axis and the lack of symmetry in the current experiment led to ambiguity in the direction of slant. Although observers could potentially recover the correct scaling factor with large continuous perspective change, this factor was paired with some random tilt that was different from the actual tilt. To further explore the role of a symmetry axis, in Experiment 4, we perturbed the symmetry axis of the hexagonal object used in Experiment 2.

Experiment 4

In Experiment 4, we used the hexagonal objects in Experiment 2 and perturbed the direction of their symmetry

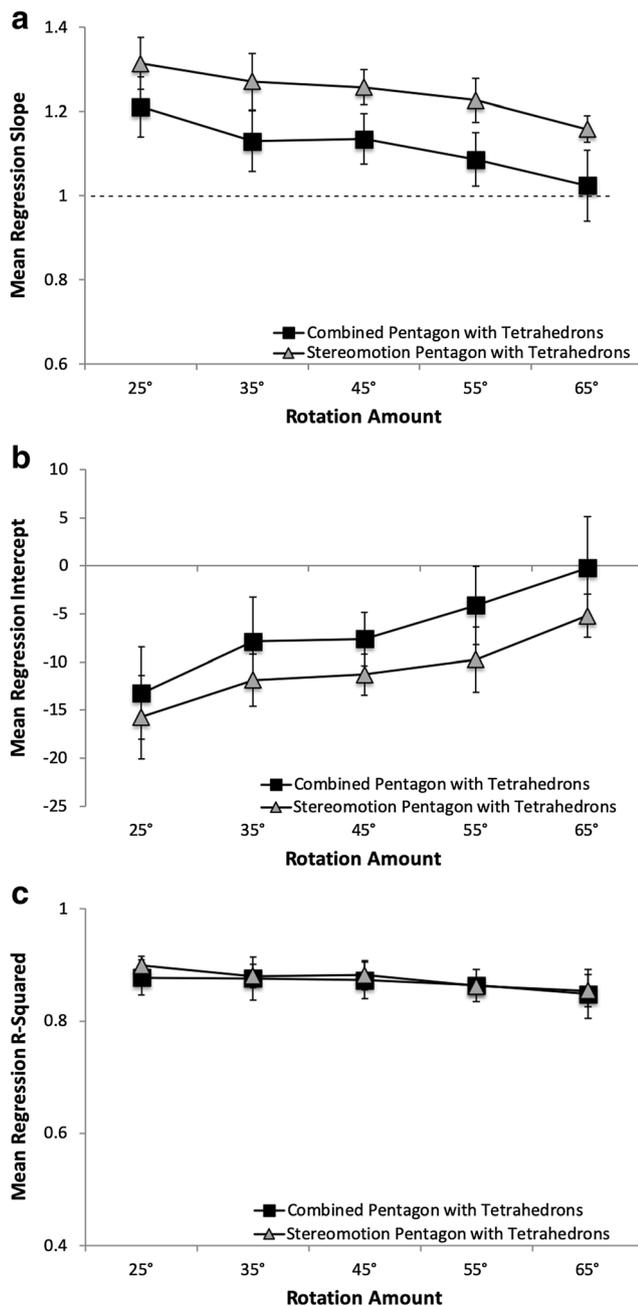


Fig. 8 Mean regression (a) slopes, (b) intercepts, and (c) r^2 in Experiment 3 for the slanted pentagonal surface with nine tetrahedrons at random locations on the surface displayed with either stereomotion or the combination of monocular SFM and stereomotion information, plotted as a function of rotation amount. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008)

axis before rotating them around a vertical axis to produce large continuous perspective change. Perturbation to the direction of the symmetry axis would not change the direction of slant. This is equivalent to changing the roll of the object (or spin as described by Saunders & Knill, 2001). If

the symmetry axis was used to identify the direction of slant, regression slopes and intercepts derived by regressing judged slants onto actual slant should first exhibit a similar pattern as one would expect from the bootstrap process (i.e., regression slopes would decrease and intercepts would increase as the amount of rotation increased) and remained relatively stable once rotation reached and went beyond 45°. However, because the judged slant was based on the direction of the symmetry axis, the actual values of slopes and intercepts should be off. Alternatively, if regression was based on slant specified by the symmetry axis, regression slopes should become 1 and intercepts, 0, at 45° of rotation and beyond. Because of the lack of difference between stereomotion and combined information in Experiment 2, we only presented the stimuli using combined visual information.

Method

Participants

Ten adults participated in this experiment, with four males and six females. All participants provided informed consent prior to participation with the approval from Indiana University's Institutional Review Board (IRB). They were compensated with course credits. All participants had normal or corrected-to-normal eyesight, and also passed a stereo fly test as in previous experiments.

Stimuli and apparatus

Experimental apparatus and setup were the same as in Experiment 2. We perturbed the object's symmetry axis (see Fig. 9). To do this, we first rotated the flat object around its surface normal by 15° (compare Fig. 9 left with Fig. 4 left). Then we rotated the object around a horizontal axis to produce a certain slant angle. Consequently, the angle between the direction of slant and the symmetry axis was 15° (compare the solid and dotted lines of Fig. 9 left). The continuous perspective change is achieved by rotating the object around a vertical axis. The object in Fig. 9 (right) had the same slant and tilt as that presented in Fig. 4 (right).

Procedure

We used the same experimental procedure as in previous experiments.

Data analysis

We used the same analysis protocol as in previous experiments. In addition, we also “corrected” the actual slant by

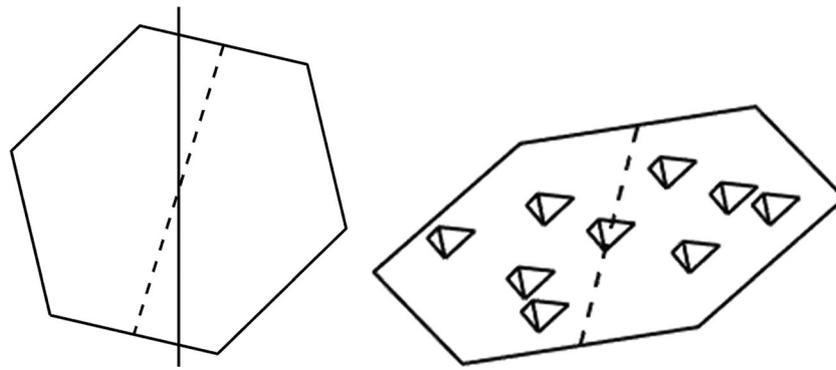


Fig. 9 Schematic of the hexagonal shape rotated around its normal by 15° (left) and the layout of the 3-D object (right). The dashed line shows the symmetry axis of the surface and the solid line shows the actual direction of slant. In the right figure, the 3-D object had the same slant and tilt as that in Fig. 4 (right). The only difference was that, in the current case, we

rotated the object by 15° around an axis that is perpendicular to the surface (its normal) and through its center (not the same as rotating the entire object around the y -axis so as to produce continuous perspective change). Note that the actual displays were presented using random dots and the dashed line was not displayed

computing slant in the direction of the symmetry axis and performed the regressions again. We conducted separate linear regressions using the uncorrected and corrected slant. Subsequently, we used mixed-design ANOVA's to analyze the resulting regression slopes and intercepts with five levels of within-subject factor of rotation amount and two levels of type of slant, either corrected based on the direction of the symmetry axis or uncorrected based on the direction of surface normal.

Results and discussion

Figure 10 shows the mean regression slopes and intercepts based on uncorrected and corrected slants. For regression slopes, Mauchly's test of sphericity showed that the sphericity assumption was violated, $\chi^2(9) = 25.20$, $p < .01$. With Greenhouse–Geisser correction, ANOVA on regression slopes showed a significant main effect of rotation amount, $F(2.54, 45.66) = 18.36$, $p < .001$, $\eta_p^2 = 0.51$, but not of the type of slant ($p > .10$, $\eta_p^2 = 0.13$) or a significant interaction effect ($p > .90$, $\eta_p^2 = 0.002$). As shown in Fig. 10a, before correction, although regression slopes did exhibit similar patterns as one would expect with an increasing amount of rotation, namely, the slopes dropped as rotation increased to 45° and remained relatively steady with further increase of rotation, the actual values of the regression slopes plateaued at a level below 1 (the upper bound of 95% confidence interval was below 1 at 55° and 65° rotations). With the slant determined along the symmetry axis (that is, “corrected”), the regression slopes actually became 1 at 45° of rotation and remained steady. Regression intercepts showed a similar pattern, where an ANOVA showed that there was a significant main effect of rotation amount, $F(4, 72) = 24.94$, $p < .001$, $\eta_p^2 = 0.58$, but not a significant main effect of type of slant ($p > .50$, $\eta_p^2 = 0.023$)

or a significant interaction effect ($p = 1.0$, $\eta_p^2 = 0$). As shown in Fig. 10b regression intercepts from comparison of judgments

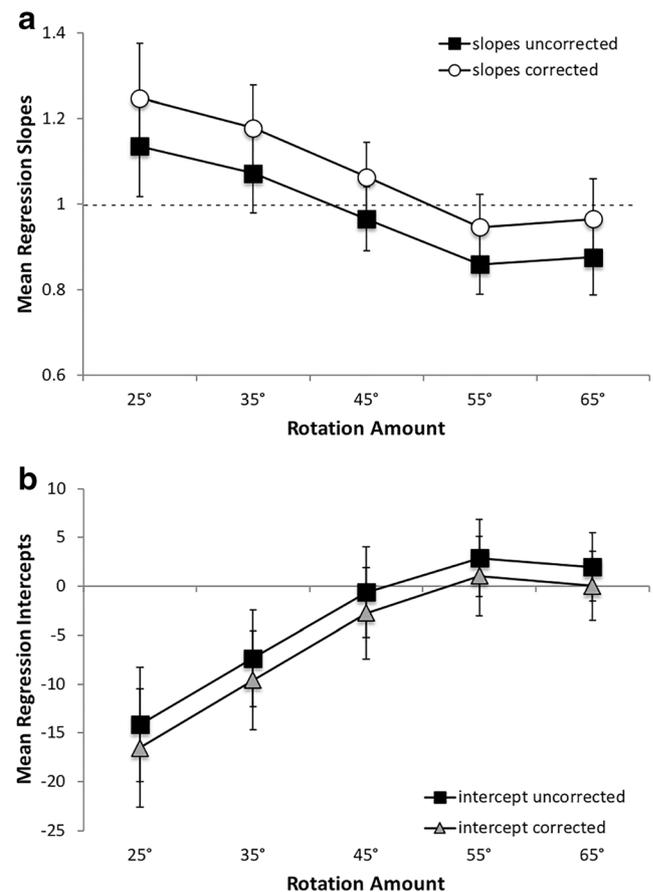


Fig. 10 Mean regression (a) slopes and (b) intercepts computed using uncorrected actual slant and corrected slant based on the direction of the symmetry axis for the hexagonal objects with a roll of 15° presented using combined information. Error bars represent 95% confidence intervals around the mean, calculated for repeated-measures designs (Cousineau, 2005; with correction by Morey, 2008)

with the uncorrected slant were slightly higher than the corrected. Comparing confidence intervals at each rotation amount, regression intercepts for both types of slant were indistinguishable from 0 at 45° of rotation and remained there as rotation increased to up to 65°.

This experiment confirmed the second postulate regarding the role of symmetry—namely that the symmetry axis is used to determine the direction of slant. This also implies that the lack of symmetry of the pentagonal object does not necessarily mean that people cannot obtain the correct scaling factor with large amounts of perspective change. In fact, Wang et al. (2019) simulated performance using the pentagonal objects in Experiment 3 and computed slant estimates based on a random direction of slant (the direction of a line formed between the top vertex and a random point along the bottom edge of the pentagonal object). They found that with a correct scaling factor at large rotation, performance as measured through regression slopes and intercepts was still poor and there was no significant difference between simulation and behavioral results.

General discussion

In this study, we examined three questions regarding the use of a bootstrap process for 3-D slant perception. The first question was whether eliminating trackable texture elements while only providing trackable 3-D relief structures could still allow the bootstrap process to work. In Experiment 1, we tested this with the same rectangular surface with nine cuboids as in Wang et al. (2018) presented using pure stereomotion information. Results showed that observers could indeed perform the bootstrap process using only 3-D relief structures without the presence of trackable texture elements, suggesting that the bootstrap process could operate on 3-D relief structures.

The second question was whether right angles were necessary for the bootstrap process, since this form of the solution becomes relatively ineffective for more upright surfaces. In Experiment 2, we tested this by eliminating explicit right angles in the stimuli using a symmetrical hexagonal surface with nine tetrahedrons at random locations, presented with either pure stereomotion or the combination of stereomotion and monocular SFM information. Results in both visual conditions matched predictions from the bootstrap model (Wang et al., 2019), where performance was poor at small rotation but became accurate at 45° of rotation and remained relatively stable as rotation further increased. Such results suggested that right angles might not be necessary for observers to perform the bootstrap process in the context of slant perception.

We subsequently developed an alternative implementation of the bootstrap process that took advantage of the equidistant points used to form one leg of the right angle in the original implementation. The line formed between these two points lies

in a frontoparallel plane allowing observers to know its length. With rotation, this line starts to be subject to the relief scaling and its length fails to remain constant. Thus, relating its original length and that after rotation could enable observers to recover the unknown scaling factor. Due to measurement noise, estimations with small rotations would be inaccurate and unstable. However, with larger rotation, estimations should start to become more accurate. This method replicates the right-angle method as proposed in Lind et al. (2014; see Fig. 10) in a way that is more effective for slant perception.

Finally, the last question was whether symmetry plays a role in metric slant perception enabled by large continuous perspective change. In Experiment 3, we eliminated mirror symmetry using an asymmetrical pentagonal surface with tetrahedrons. Results in both stereomotion and combined conditions failed to replicate previous experiments: neither conditions produced veridical performance at 45° of continuous perspective change or greater, although the combined condition did achieve veridical performance at 65° of continuous perspective change. In Experiment 4, we further tested the role of symmetry by using the hexagonal objects from Experiment 2, but with changed orientation in the plane containing the surface (that is, the roll) so that the symmetry axis was no longer aligned with the direction of slant. Results confirmed that observers could recover the correct scaling factor with a sufficiently large amount of rotation, but slant judgment was based on the direction of the symmetry axis. This finding indicated that poor performance in Experiment 3 was due to the ambiguity of the direction of slant, as confirmed in Wang et al. (2019).

This study yields an interesting implication regarding the debate on optical and geographical slant (Gibson & Cornsweet, 1952). There has been a long history of studies that defined slant as the angle formed between a surface normal and the line of sight (e.g., Gibson, 1950; Norman et al., 2006; Saunders & Knill, 2001; Sawada & Pizlo, 2008; Stevens, 1983a; Todd & Perotti, 1999). However, as Sedgwick and Levy (1985) showed, observers are not adept at judging the slant of a surface based on the line of sight. They argued that it would be more efficient to define slant using an environmental frame of reference, relative to the direction of gravity. They called this environment-centered slant. Cherry and Bingham (2018) explicitly manipulated optical slant by changing the viewing perspective while maintaining the geographical slant. They did not find an effect of perspective, demonstrating that observers might not perceive optical slant.

For a slanted surface rotating to yield structure-from-motion, optical slant would constantly change. If observers truly perceived optical slant, slant judgments should be highly variable in this circumstance. This was not what we found, and, instead, slant judgments were predicted well by geographical slant. In the context of SFM, the perception of slant in terms of geographical slant means that the momentary direction of the

slant cannot be determined in terms of the direction of maximum increasing depth (e.g., Stevens, 1983a, 1983b). Because of this, observers may be somewhat imprecise in determining the direction of slant. If so, then observers might rely on the direction of major axes of symmetry of a surface contour shape as a constraint used to estimate the slant direction. Indeed, we found that slant judgments were relatively poor in the absence of symmetry and good with symmetry, and that a change in the direction of the symmetry axis impacted slant judgments as predicted by this hypothesis. Finally, most studies on optical slant and tilt perception suggested that tilt is perceived reliably and accurately whereas slant is not (Koenderink & van Doorn, 1995; Koenderink, van Doorn, & Kappers, 1992, 1994, 1995; Norman, Todd, & Phillips, 1995; Stevens, 1983b; Todd, Koenderink, van Doorn, & Kappers, 1996). Now, with a geographical frame of reference and continuous perspective change, slant is perceived accurately but perception of tilt (or the direction of slant) is a function of the symmetry of the surface contour.

Open practice statement Data and materials for all experiments are available upon request to the corresponding author, and none of the experiments were preregistered.

References

- Allen, B., Haun, A. M., Hanley, T., Green, C. S., & Rokers, B. (2015). The optimal combination of the binocular cues to 3-D motion. *Investigative Ophthalmology & Visual Science*, *56*(12), 7589–7596.
- Bingham, G. P., & Lind, M. (2008). Large continuous perspective transformations are necessary and sufficient for accurate perception of metric shape. *Perception & Psychophysics*, *70*(3), 524–540.
- Bingham, G. P., & Muchisky, M. M. (1993a). Center of mass perception and inertial frames of reference. *Perception & Psychophysics*, *54*(5), 617–632.
- Bingham, G. P., & Muchisky, M. M. (1993b). Center of mass perception: Perturbation of symmetry. *Perception & Psychophysics*, *54*(5), 633–639.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.
- Cherry, O. C., & Bingham, G. P. (2018). Searching for invariance: Geographical and optical slant. *Vision Research*, *149*, 30–39.
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45.
- Cumming, B. G., & Parker, A. J. (1994). Binocular mechanisms for detecting motion-in-depth. *Vision Research*, *34*(4), 483–495.
- Dick, M., & Hochstein, S. (1989). Visual orientation estimation. *Perception & Psychophysics*, *46*(3), 227–234.
- Domini, F., & Caudek, C. (2013). Perception and action without veridical metric reconstruction: An affine approach. In S. Dickinson & Z. Pizlo (Eds.), *Shape perception in human and computer vision* (pp. 285–298). London, UK: Springer.
- Durgin, F. H., & Li, Z. (2011). The perception of 2-D orientation is categorically biased. *Journal of Vision*, *11*(8), 13–13.
- Gibson, J.J. (1950). The perception of visual surfaces. *The American Journal of Psychology*, *63*(3), 367–384.
- Gibson, J. J., & Cornsweet, J. (1952). The perceived slant of visual surfaces—Optical and geographical. *Journal of Experimental Psychology*, *44*(1), 11–15.
- Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago, IL: University of Chicago Press.
- Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3?. *Perception*, *36*(14) 1–16.
- Koenderink, J. J., & van Doorn, A. J. (1991). Affine structure from motion. *Journal of the Ophthalmic Society of America*, *8*(2), 377–385.
- Koenderink, J. J., & van Doorn, A. J. (1995). Relief: Pictorial and otherwise. *Image and Vision Computing*, *13*(5), 321–334.
- Koenderink, J. J., Van Doorn, A. J., & Kappers, A. M. (1992). Surface perception in pictures. *Perception & Psychophysics*, *52*(5), 487–496.
- Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. (1994). On so-called paradoxical monocular stereoscopy. *Perception*, *23*(5), 583–594.
- Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. (1995). Depth relief. *Perception*, *24*(1), 115–126.
- Lee, Y., & Bingham, G. P. (2010). Large perspective changes (>45°) yield perception of metric shape that allows accurate feedforward reaches-to-grasp and it persists after the optic flow has stopped! *Experimental Brain Research*, *204*, 559–573.
- Lee, Y. L., Lind, M., & Bingham, G. P. (2013). Perceived 3-D metric (or euclidean) shape is merely ambiguous, not systematically distorted. *Experimental Brain Research*, *224*, 551–555.
- Lee, Y. L., Lind, M., Bingham, N., & Bingham, G. P. (2012). Object recognition using metric shape. *Vision Research*, *69*, 23–31.
- Li, Y., Sawada, T., Shi, Y., Kwon, T., & Pizlo, Z. (2011). A Bayesian model of binocular perception of 3-D mirror symmetrical polyhedra. *Journal of Vision*, *11*(4), 1–20.
- Lind, M. (1996). Perceiving motion and rigid structure from optic flow: A combined weak-perspective and polar-perspective approach. *Perception & Psychophysics*, *58*(7), 1085–1102.
- Lind, M., Lee, Y. L., Mazanowski, J., Kountouriotis, G. K., & Bingham, G. P. (2014). Affine operations plus symmetry yield perception of metric shape with large perspective changes ($\geq 45^\circ$): Data and model. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(1), 83.
- Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64.
- Nefs, H. T., O'Hare, L., & Harris, J. M. (2010). Two independent mechanisms for motion-in-depth perception: Evidence from individual differences. *Frontiers in Psychology*, *155*. <https://doi.org/10.3389/fpsyg.2010.00155>
- Norman, J. F., Todd, J. T., Norman, H. F., Clayton, A. M., & McBride, T. R. (2006). Visual discrimination of local surface structure: Slant, tilt, and curvedness. *Vision research*, *46*(6–7), 1057–1069.
- Norman, J. F., Todd, J. T., Perotti, V. J., & Tittle, J. S. (1996). The visual perception of three-dimensional length. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(1), 173–186.
- Norman, J. F., Todd, J. T., & Phillips, F. (1995). The perception of surface orientation from multiple sources of optical information. *Perception & Psychophysics*, *57*(5), 629–636.
- Pelli, D. G. (1997) The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Pizlo, Z. (2010). *3-D shape: Its unique place in visual perception*. Cambridge, MA: MIT Press.
- Pizlo, Z., Sawada, T., Li, Y., Kropatsch, W. G., & Steinman, R. M. (2010). New approach to the perception of 3-D shape based on veridicality, complexity, symmetry and volume. *Vision Research*, *50*(1), 1–11.
- Saunders, J. A., & Knill, D. C. (2001). Perception of 3-D surface orientation from skew symmetry. *Vision Research*, *41*(24), 3163–3183.

- Sawada, T., & Pizlo, Z. (2008). Detection of skewed symmetry. *Journal of Vision*, *8*(5), 14, 1–18.
- Sedgwick, H. A., & Levy, S. (1985). Environment-centered and viewer-centered perception of surface orientation. *Computer Vision, Graphics, and Image Processing*, *31*, 248–260.
- Shapiro, L. S., Zisserman, A., & Brady, M. (1995). 3-D motion recovery via affine epipolar geometry. *International Journal of Computer Vision*, *16*(2), 147–182.
- Shioiri, S., Saisho, H., & Yaguchi, H. (2000). Motion in depth based on inter-ocular velocity differences. *Vision Research*, *40*, 2565–2572.
- Stevens, K. A. (1983a). Slant-tilt: The visual encoding of surface orientation. *Biological Cybernetics*, *46*, 183–195.
- Stevens, K. A. (1983b). Surface tilt (the direction of slant): A neglected psychophysical variable. *Perception & Psychophysics*, *33*(3), 241–250.
- Thompson, D., & Mundy, J. (1987, March). Three-dimensional model matching from an unconstrained viewpoint. *Proceedings. 1987 IEEE International Conference on Robotics and Automation* (Vol. 4, pp. 208–220). New York, NY: IEEE.
- Todd, J. T., & Bressan, P. (1990). The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Perception & Psychophysics*, *48*(5), 419–430.
- Todd, J. T., Christensen, J. C., & Guckes, K. M. (2010). Are discrimination thresholds a valid measure of variance for judgments of slant from texture?. *Journal of Vision*, *10*(2), 1–18.
- Todd, J. T., Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. (1996). Effects of changing viewing conditions on the perceived structure of smoothly curved surfaces. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(3), 695.
- Todd, J. T., & Norman, J. F. (1991). The visual perception of smoothly curved surfaces from minimal apparent motion sequences. *Perception & Psychophysics*, *50*(6), 509–523.
- Todd, J. T., & Norman, J. F. (2003). The visual perception of 3-D shape from multiple cues: Are observers capable of perceiving metric structure?. *Perception & Psychophysics*, *65*(1), 31–47.
- Todd, J. T., Oomes, A. H., Koenderink, J. J., & Kappers, A. M. (2001). On the affine structure of perceptual space. *Psychological Science*, *12*(3), 191–196.
- Todd, J. T., & Perotti, V. J. (1999). The visual perception of surface orientation from optical information. *Perception & Psychophysics*, *61*(8), 1577–1589.
- Wagner, M. (1985). The metric of visual space. *Perception and Psychophysics*, *38*(6), 483–495.
- Wang, X. M., Lind, M., & Bingham, G. P. (2018). Large continuous perspective change with noncoplanar points enables accurate slant perception. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(10), 1508–1522.
- Wang, X. M., Lind, M., & Bingham, G. P. (2019). Bootstrapping a better slant: A stratified process for recovering 3-D metric slant. *Attention, Perception, & Psychophysics*. <https://doi.org/10.3758/s13414-019-01860-y>.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.