tapraid5/zfn-xhp/zfn-xhp/zfn00413/zfn2975d13z | xppws | S=1 | 5/29/13 | 20:43 | Art: 2011-0122

APA NLM

Journal of Experimental Psychology:
Human Perception and Performance

0096-1523/13/$12.00 DOI: 10.1037/a0033245

Affine Operations Plus Symmetry Yield Perception of Metric Shape With Large Perspective Changes (≥45°): Data and Model

AQ: au
AQ: 1

Mats Lind
Uppsala University

Young Lim Lee
Dankook University

Janusz Mazanowski, Georgios K. Kountouriotis, and Geoffrey P. Bingham
Indiana University

AQ: 2

AQ: 3

AQ: 4

G. P. Bingham and M. Lind (2008, Large continuous perspective transformations are necessary and sufficient for accurate perception of metric shape, *Perception & Psychophysics*, Vol. 70, pp. 524–540) showed that observers could perceive metric shape, given perspective changes ≥45° relative to a principal axis of elliptical cylinders. In this article, we tested (a) arbitrary perspective changes of 45°, (b) whether perception gradually improves with more perspective change, (c) speed of rotation, (d) whether this works with other shapes (asymmetric polyhedrons), (e) different slants, and (f) perspective changes >45°. Experiment 1 compared 45° perspective change away from, versus centered on, a principal axis. Observers adjusted an ellipse to match the cross-section of an elliptical cylinder viewed in a stereo-motion display. Experiment 2 tested whether performance would improve gradually with increases in perspective change, or suddenly with a 45° change. We also tested speed of rotation. Experiment 3 tested (a) asymmetric polyhedrons, (b) perspective change beyond 45°, and (c) the effect of slant. The results showed (a) a particular perspective was not required, (b) judgments only improved with ≥45° change, (c) speed was not relevant, (d) it worked with asymmetric polyhedrons, (e) slant was not relevant, and (f) judgments remained accurate beyond 45° of change. A model shows how affine operations, together with a symmetry yielded by 45° perspective change, bootstrap perception of metric shape.

AQ: 5

Keywords: 3D shape perception, affine geometry, space perception, metric, structure-from-motion, stereo motion

Observers have been found to exhibit inaccurate perception of metric 3D structure in many studies investigating different sources of visual information about shape, including binocular stereopsis (e.g., Johnston, 1991; Tittle, Todd, Perotti, & Norman, 1995), monocular motion (e.g., Norman & Lappin, 1992; Norman & Todd, 1993; Perotti, Todd, Lappin, & Phillips, 1998; Tittle et al., 1995; Todd & Bressan, 1990; Todd & Norman, 1991), the combination of binocular disparity and motion (Tittle & Braunstein, 1993; Tittle et al., 1995), and other multicue conditions (Norman & Todd, 1996; Norman, Todd, & Phillips, 1995; Todd, Tittle, & Norman, 1995). Perotti et al. (1998) asked observers to view displays of surfaces and to judge two different measures of shape to investigate perception of qualitative and metric properties of object shape, respectively. The "shape index" measures qualitative variations in smooth surface shape (cylindrical, ellipsoidal, saddle, etc.). "Curvedness" measures metric variations in shape corresponding to the amount or magnitude of surface curvature (Koenderink, 1990). Perotti et al. found that the shape index was judged accurately, but that judgments of curvedness were inaccurate and highly variable (see also Experiment 4 in Norman, Todd, Norman, Clayton, & McBride, 2006).

In a similar vein, Lind, Bingham, and Forsell (2003) used actual wooden cylindrical objects to investigate whether metric structure can be perceived accurately. Objects with depth-to-width aspect ratios that varied from 0.46 and 1.8 were placed on a tabletop within reach distance, and at an eye height (≈15 cm) that allowed the tops of the objects to be seen. Observers could move their head freely in normal lighting with binocular vision. They adjusted the shape of an elliptical outline on a computer screen to be the same as the perceived shape of the horizontal cross-section of each cylinder. Similar to the results of Perotti et al. (1998), observers judged metric shape of the cylinders inaccurately and with high variably. Lind et al. also varied the viewing height and distance of the cylinders, with the result that observers only judged the metric shape correctly when looking straight down on the tops of the objects.

Contrary to these previous results, Bingham and Lind (2008) showed that it is possible for observers to perceive metric structure

AQ: 10 Mats Lind, Department of Information Technology, Division of Visual Information and Interaction, Uppsala University, Uppsala, Sweden; Young Lim Lee, Department of Psychology, Dankook University, Cheonan, South Korea; Janusz Mazanowski, Georgios K. Kountouriotis, and Geoffrey P. Bingham, Department of Psychological and Brain Sciences, Indiana University.

Correspondence concerning this article should be addressed to Geoffrey P. Bingham, Department of Psychological and Brain Sciences, Indiana University, 1101 East Tenth Street, Bloomington, IN 47405-7007. E-mail: gbingham@indiana.edu

accurately, given sufficiently large perspective changes, namely, a continuous rotation of 45° or more. Using stereo vision in a virtual environment, observers viewed an elliptical cylinder while it was rotated back and forth by 30°, 45°, 60°, or 90°. Then they judged object shape by touching the locations of the front, back, and sides of the virtual object with a nonvisible stylus.[1] Judgments of metric shape were incorrect with only 30° rotation, but with rotations of 45° or greater, they became accurate. Observers also judged metric shape from two discrete views separated by 90°, but this did not allow accurate judgments. Bingham and Lind concluded that a continuous 45° perspective change is both necessary and sufficient to allow accurate perception of metric 3D shape. Relatedly, Brenner and van Damme (1999) had observers use their right hand to move a computer mouse to adjust the depth of a simulated ellipsoid to match a tennis ball that they were holding in their left hand (but were unable to see). Observers overestimated the depth of the ellipsoid when a static object was presented, but their judgments became accurate when the object was rotated by 60°.

These results show that a large continuous perspective change (that is, at least 45°) can yield accurate perception of metric shape. A number of questions remain. First, in Bingham and Lind (2008), elliptical cylinders were rotated away from and back to an initial perspective in which the gaze axis was aligned with a principal axis of the elliptical objects. Thus, it remains possible that a perspective rotated 45° from a principal axis of the object yielded accurate judgment of metric shape, rather than a generic perspective change of 45°. We tested this in Experiment 1.

Second, if it is simply the case that more perspective change yields more information, then perception of metric shape might improve gradually as the degree of perspective change increases. Alternatively, the existing results suggest that the requisite information only becomes available once perspective change reaches 45°. We tested this in Experiment 2. Third, it is well known that speed of object rotation can be confused with object depth and, thus, 3D metric shape (e.g., Koenderink & van Doorn, 1991; Lind, 1996). In Experiment 2, we also tested the effect of different speeds of rotation.

Fourth, the depth of objects viewed with less than 45° of perspective change tends to be underestimated. With 45° of perspective change, metric shape was found to be judged accurately, meaning that the perceived depth was greater. If perspective changes were increased beyond 45°, perceived depth might continue to increase or, instead, the perceived metric shape might remain accurate. We tested this in Experiment 3. Bingham and Lind (2008) tested the perception of elliptical cylinders that varied in aspect ratio, that is, in the ratio of width to depth (see also Lee & Bingham, 2010). Ellipses exhibit reflective (or mirror) symmetry in respect to their principal axes (or principal planes in the case of elliptical cylinders). This symmetry might be essential for good perception of metric shape, in which case the findings of Bingham and Lind (2008) may not generalize to other shapes, in particular, asymmetric ones. We tested this in Experiment 3 using asymmetric polyhedrons. Finally, because the top surfaces of these objects have been visible in all the experiments, it is important to test the effect of variations in the perspective in respect to the top surface, that is, in the slant. We also tested this in Experiment 3.

Experiment 1

We compared 45° of continuous perspective change away from, versus centered on, a principal axis of an elliptical cylinder to determine whether a generic set of perspective changes would enable observers to perceive metric shape or whether the particular perspective at 45° to the principal axis is required. Observers adjusted an ellipse in a display to match the shape of the horizontal cross-section of an elliptical cylinder viewed in a 3D display containing binocular disparity and motion.

Method

Participants. Ten adults (five females and five males) at Indiana University participated in this experiment. All had normal or corrected-to-normal vision and passed a stereo fly test (Stereo Optical Co., Inc.) that was used to check stereoscopic depth perception. All of the participants were naïve as to the purpose of the study and were paid $7 per hour. All procedures were approved by and conform to the standards of the Indiana University Human Subjects Committee.

Stimuli. Target objects were generated in a computer display using anaglyphs (red–blue) for stereo (see Figure 1). A calibration procedure was used to minimize any cross talk between the channels.[2] Five cylindrical objects were generated by varying the depth dimension to produce different depth-to-width aspect ratios as follows: 0.70, 0.85, 1.00, 1.15, and 1.30. Objects were 4 cm in height and 10 cm in width, with a depth determined by the depth-to-width ratio. As shown in Figure 1, objects were covered by texture consisting of randomly oriented line segments that were randomly distributed on the object surface and then selectively drawn after hidden line removal. Objects were displayed on a Mitsubishi Diamond Plus 74SB CRT computer screen with a resolution of 1280 × 1024 and a frame rate of 60 Hz. The cyclopean eye of the observer was 60 cm from the screen and 15 cm above the lowest visible line of the object on the screen. IPD was 6 cm. The slant of the visible top of the object was 8°. Rotational speed was 15°/s. The projection of the 3D stimuli onto the computer monitor was perspective.

Procedure. Each participant's eye height and distance from the display were set by adjusting chair height and position and having participants rest their head on a chin rest attached to the chair. A 3D cylindrical object was shown with stereo and motion, rotating by either 45° or 22.5°. Each rotation amount occurred in two ways: full and half. In the full rotation, the object was rotated continuously by either 45° or 22.5° to one side from a canonical view so that, at the end of the rotation, the perspective was at either 45° or 22.5° from the canonical view. A canonical view is looking straight down a principal axis of the elliptical object. Each object was rotated from 0 to 45° (or 0 to 22.5°) and then back again. In

[1] A previous study using the same response measure and similar viewing conditions, but with free head movements yielding perspective changes of about 10°, had found poor performance (Bingham, Bradley, Bailey, & Vinner, 2001). In their first experiment, Bingham and Lind (2008) replicated this result using a variety of elliptical shapes.

[2] Screen color was adjusted to minimize visibility through the opposite colored lens of the glasses, so that the red image elements would be invisible through the blue lens and the blue image elements would be invisible through the red lens.

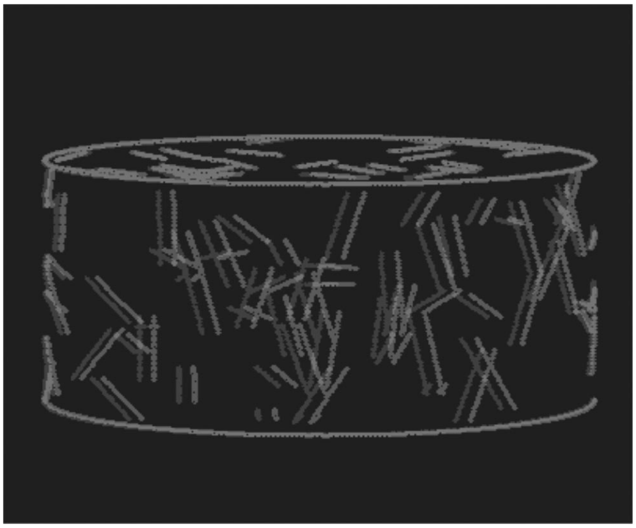

Figure 1. A sample anaglyph display showing a stereoscopic elliptical cylinder that would be viewed by a participant wearing red–blue glasses. A gray-scale version of the display is shown. The actual display was in red and blue.

the half rotation, the object was rotated by either 22.5° or 11.25°, first to one side from a canonical view, and then to the other side of the canonical view, so that the perspective was changed by a total of either 45° or 22.5°, respectively (Pedhazur, 1982). The object was rotated from 0 to −22.5° (or −11.25°) and back, and then from 0 to + 22.5° (or +11.25°) and back. For each trial, a 3D cylinder was randomly selected from among five different aspect ratios and displayed on the bottom half of the screen, rotated in a fashion randomly selected from among the four different conditions (2 amounts × 2 ways). The cylinder remained displayed on the screen after rotation (with stereo), and then a 2D ellipse was shown on the top half of the computer screen. The initial height of 2D ellipse was randomly selected. Participants were asked to adjust the height of the ellipse by pressing computer keys so as to match the eccentricity (or aspect ratio) of the ellipse to the eccentricity (or aspect ratio) of horizontal cross section of the 3D object. Because the aspect ratio of 3D object was changed only by varying the depth dimension, participants altered only the height of the 2D ellipse by pressing arrow keys. Time to make this judgment was not limited. When the participant was satisfied with the ellipse, he or she hit the space bar to finish the judgment and perform the next trial. A total of 80 judgments was performed by each participant (4 rotation conditions × 5 objects × 4 repetitions).

Results and Discussion

We addressed two questions in this first experiment. First, we compared judgments of aspect ratios with either 45° of rotation or AQ: 6 only 22.5° of rotation. Assuming that a continuous 45° rotation is both necessary and sufficient for accurate perception of metric shape, as found by Bingham and Lind (2008), then the second question was whether a generic 45° continuous perspective change would allow accurate perception of metric shape or instead, a change to a 45° perspective with respect to a principal axis would be required. To test the second question, we used the full- and the half-rotation conditions. We regressed judged aspect ratios on actual aspect ratios. We performed a multiple regression to test differences in slopes of judged aspect ratios between the full- and the half-rotation condition and between the two rotation amounts, 45° and 22.5°, respectively. There were five independent variables: actual aspect ratio (a continuous variable), full versus half rotation (coded as ±1), an interaction vector (computed as the product of the first two vectors), 45° versus 22.5° rotation amount (coded as ±1), and an interaction vector (computed as the product of the aspect ratio and rotation amount vectors). The overall regression was significant, $F(5, 757) = 348.2$, $p < .001$, and accounted for 72% of the variance. There was no main effect or significant interaction for the full- and the half-rotation conditions. However, there was a main effect of 45° versus 22.5° rotation amount, $t(757) = 3.55$, $p < .001$, as well as a significant interaction, $t(757) = 3.85$, $p < .001$, showing that there was a significant difference in the slopes and intercepts for judged aspect ratios between the two rotation amounts. As shown in Figure 2, we found F2 that 45° of continuous rotation yielded good perception of metric shape (slope = 1.00 and intercept = −0.01 ≈ 0), whereas less than this (that is, 22.5°) yielded low slope = .84 and high intercept = 0.14.

Experiment 2

In Experiment 1, we found that 45° of perspective change is necessary and sufficient for good metric shape perception. The next question we addressed was whether progressively increasing amounts of rotation would yield a corresponding gradual improve-

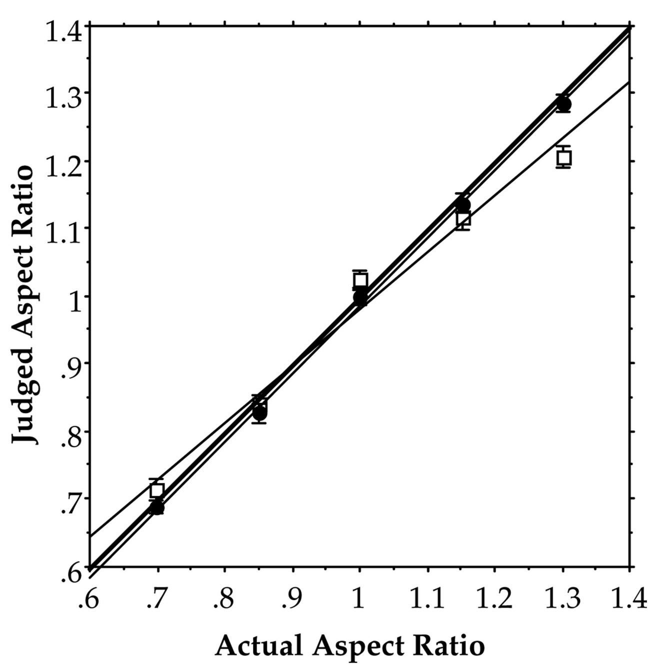

Figure 2. Mean judged aspect ratios plotted as a function of actual aspect ratios with standard error bars representing between-subjects variability. Filled circles: 45° rotation. Open squares: 22.5° rotation. The dark line represents the correct target aspect ratio, that is, a line of slope = 1 and intercept = 0.

APA NLM
tapraid5/zfn-xhp/zfn-xhp/zfn00413/zfn2975d13z | xppws | S=1 | 5/29/13 | 20:43 | Art: 2011-0122

ment in judgments of metric shape, or instead whether a sudden improvement would occur only with 45° of perspective change. In this experiment, we tested increments in the amount of perspective change from 11.25° to 22.5°, 30°, 37.5°, and 45° to determine if they yielded incremental improvements in judgments of the aspect ratios, or instead whether a single discrete improvement occurred with an increase to 45° of perspective change. We also tested the potential effect of different speeds of rotation (or perspective change). We manipulated rotational speed because judgments of aspect ratios could potentially vary depending on differential speeds of texture elements.

Method

Participants. Nine adults (three females and six males) at Indiana University participated in this experiment. All had normal or corrected-to-normal vision and passed a stereo fly test (Stereo Optical Co., Inc.) that was used to check stereoscopic depth perception. All of the participants were naïve as to the purpose of the study and were paid $7 per hour. All procedures were approved by and conform to the standards of the Indiana University Human Subjects Committee.

Stimuli and procedure. The stimuli and procedure were similar to Experiment 1, with the following differences. First, we used four elliptical cylinders with aspect ratios as follows: 0.7, 0.9, 1.1, and 1.3. Second, because there was no difference between the full- and the half-rotation conditions in Experiment 1, we used only the half-rotation condition, so the 3D object was rotated to both sides from a canonical view. Third, we tested five different rotation amounts: 11.25°, 22.5°, 30°, 37.5°, and 45°. Fourth, we varied rotational speed using three speeds: 10.5°/s, 14.8°/s, and 19°/s. Rotation was about a vertical axis through the center of the object. Lastly, although the static 3D object remained after rotation in the first experiment, in this experiment, the 3D object disappeared after rotation and before the 2D ellipse was displayed (to be adjusted by the observer). The three factors—rotation amount, object aspect ratio, and rotation speed—were ordered randomly. A total of 60 judgments were performed by each participant (5 rotation amounts × 4 objects × 3 speed variations).

Results and Discussion

In Experiment 2, we investigated whether performance would improve in proportion as more perspective change was provided, or instead whether 45° of perspective change would be special in allowing accurate metric shape perception. First, we performed linear regressions of judged aspect ratios on actual aspect ratios and built a data set of slopes and another of r^2 for each Rotation Amount × Rotational Speed cell for each participant. The two data sets, slopes and r^2, were each analyzed by means of ANOVA with two repeated measures factors: rotation amount (five levels) and speed (three levels). The ANOVA on slopes yielded a main effect of rotation amount, $F(4, 32) = 6.7$, $p < .01$, and a main effect of rotational speed, $F(2, 16) = 6.4$, $p < .01$, but no interaction (see
F3 Figures 3a and b.) The ANOVA on r^2 was similar, yielding a main effect of rotation amount, $F(4, 32) = 3.1$, $p < .05$, as well as a main effect of rotational speed, $F(2, 16) = 3.7$, $p < .05$, but no interaction (see Figures 3a and b).

Because there was no significant interaction either in slopes or in r^2, we proceeded by using nonparametric analyses separately for

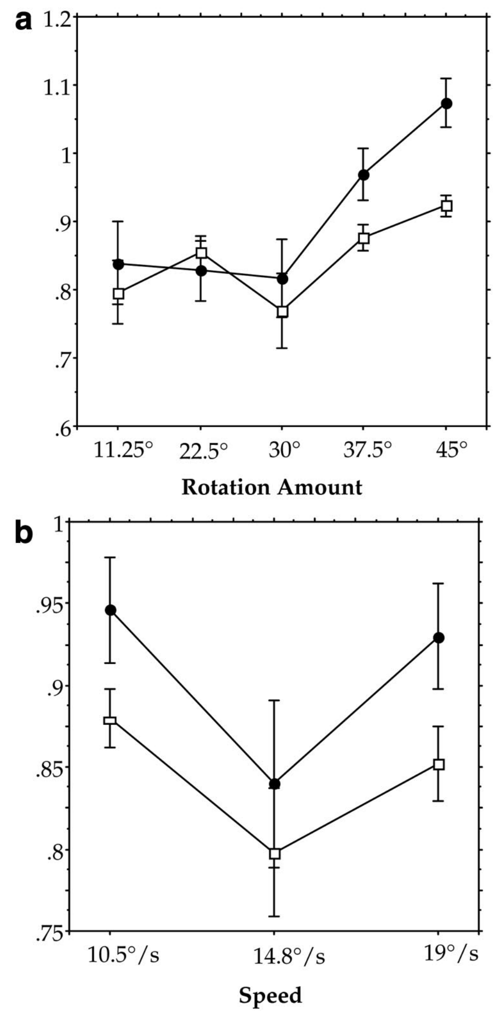

Figure 3. (a) Mean slopes (filled circles) and mean r^2 (open squares), each plotted as a function of the five different amounts of rotation (11.25°, 22.5°, 30°, 37.5°, 45°). Standard error bars represent between-subjects variability. (b) Mean slopes (filled circles) and mean r^2 (open squares), each plotted as a function of the three rotation speeds (10.5°/s, 14.8°/s, 19°/s). Standard error bars represent between-subjects variability.

each factor to investigate how results varied for each level within each factor. In each case, we first used a Friedman nonparametric ANOVA. In the analysis of slopes, we found that there were significant differences between the different amounts of rotation (Friedman ANOVA χ^2 [$N = 9$, $df = 4$] = 10.0, $p < .05$). We performed pairwise nonparametric Wilcoxon's matched-pairs signed-ranks tests to test which pair of conditions was significantly different. As shown in Table 1, the 45° rotation condition was AQ:7,T1
different from the 11.25°, 22.5°, and 30° rotation conditions. The 37.5° rotation condition was not significantly different from either the smaller rotation conditions or the 45° condition. These results showed that the accuracy of metric shape perception is not a gradual or progressive function of the magnitude of perspective changes. Instead, accurate metric shape perception requires approximately 45° of perspective change and improves relatively suddenly as the amount of perspective change reaches ≈45°.

Table 1
The Results of Pairwise Nonparametric Wilcoxon Tests of Rotation Amounts

	11.25°	22.5°	30°	37.5°	45°
11.25°	x	p = .95	p = .59	p = .68	***p* = .03**
22.5°		x	p = .86	p = .26	***p* = .02**
30°			x	p = .14	***p* = .03**
37.5°				x	p = .11
45°					x

Note. Bolded values show significant differences.

We performed the same Friedman ANOVA nonparametric test to evaluate the slopes for each speed. We found that there were significant differences between the different speeds (Friedman ANOVA χ^2 [$N = 9$, $df = 2$] = 8.4, $p < .05$). But when we performed a pairwise nonparametric Wilcoxon's matched-pairs signed-ranks test to determine which pair of speed levels was significantly different, we found, T2 as shown in Table 2, that only the 10.5°/s and 14.8°/s speeds yielded a significant difference. On average, the observers' performance was poor for the second level of speed (14.8°/s) compared with the first (10.5°/s) and the third (19°/s) level, as shown in Figure 3b. It was unclear why performance should have suffered for the middle speed, but because there were no systematic variations in judged shape as a monotonic function of rotation speed, we judged that our other results could not be an artifact produced by variations in differential speeds of optic flow.

Finally, one might expect possible effects of speed when the information is otherwise ambiguous, that is, when the amount of perspective change was <45° (see Lee, Lind, & Bingham, 2013). However, once good information was available, that is, with rotations ≥45°, then there should be no effect. We tested the three levels of speed with rotation amount restricted to 45° and found no significant difference either for slopes (Friedman ANOVA χ^2 [$N = 9$, $df = 2$] = 1.56, $p > .4$) or for r^2 (Friedman ANOVA χ^2 [$N = 9$, $df = 2$] = 1.56, $p > .4$). Speed of rotation had no effect once the amount of rotation reached 45°.

Experiment 3

The results of Experiment 2 showed that judgments of metric 3D shape did not gradually improve as the amount of rotation or perspective change progressively increased. Instead, judgments improved suddenly as the amount of rotation approached 45°. In linear regressions of judged aspect ratios on actual aspect ratios, the slopes remained ≈0.85 until, with 45° of rotation, they increased to ≈1.0. The next question was whether the slopes would remain ≈1.0 as the amount of rotation is increased significantly beyond 45°. An alternative possibility is that the perceived depth simply continues to increase with the result that the slopes in these regressions also continue to increase.

In Experiments 1 and 2 (and in the other previous experiments of Bingham & Lind, 2008), the only shapes tested were elliptical cylinders. Elliptical cylinders were tested because they are simple in structure in a way that actually makes these tasks more difficult. Nevertheless, it is important to test generalization to other types of objects. Therefore, in Experiment 3, we tested polyhedrons. As discussed at length by Pizlo (2008; and in his earlier work cited therein), such objects contain structure that should make this task easier. Thus, if Pizlo is correct, we should expect performance with amounts of perspective change less than 45° to be better than previously found for the elliptical cylinders. Nevertheless, metric shape should only be perceived accurately once the amount of perspective change reaches or exceeds 45°.

A property of the elliptical cylinders that might have made the task easier, once the amount of rotation reached 45°, was their symmetries. Elliptical cylinders exhibit reflective symmetry left to right and front to back (when the line of sight is along a principal axis of the ellipse). To test the relevance of this symmetry to the previous results, we used an asymmetric polyhedron in Experiment 3. This object exhibited no reflective symmetry whatsoever.

Finally, the tops of the objects are visible in these experiments and thus exhibit a slant. Only a single slant was tested in all of the previous experiments. In Experiment 3, we varied the slant, testing two different values, 8° and 16°. This could not be achieved by simply rotating the object around a frontoparallel axis because this would confound slant with the shape of the horizontal cross-section. (If the object itself was changed in orientation to the vertical, then a horizontal cross section might intersect either the top or bottom surface, for instance.) Instead, we simply manipulated the eye height at which the object sat. A smaller eye height yielded a larger slant. This kept the orientation relative to the vertical constant and thus preserved the shape of the horizontal cross-section.

In summary, in Experiment 3, we tested five different width-to-depth aspect ratios (0.8, 0.9, 1.0, 1.1, 1.2) of asymmetric polyhedrons with variations in slant of the top surface (8°, 16°). We tested amounts of perspective change less than and equal to 45° (25°, 35°, 45°) to determine if the previous results would be replicated with this type of object and with different slants. We tested amounts of perspective change significantly greater than 45° (55°, 65°, 75°) to determine if performance would remain accurate.

Method

Participants. Eighteen adults (12 females and six males) at Indiana University participated in this experiment. All had normal or corrected-to-normal vision and passed a stereo fly test (Stereo Optical Co., Inc.) that was used to check stereoscopic depth perception. All procedures were approved by and conform to the standards of the Indiana University Human Subjects Committee.

Stimuli and Procedure. The stimuli and procedure were similar to Experiments 1 and 2, with the following differences. First, we used five asymmetric polyhedrons with aspect ratios as follows: 0.8, 0.9, 1.0, 1.1, and 1.2 (see Figure 4). Second, we tested F4 six different amounts of rotation as follows: 25°, 35°, 45°, 55°, 65°, and 75°. Third, we varied the eye height of the objects on the

Table 2
The Results of Pairwise Nonparametric Wilcoxon Tests of Rotation Speeds

	10.5°/s	14.8°/s	19.0°/s
10.5°/s	x	***p* = .02**	p = .62
14.8°/s		x	p = .07
19.0°/s			x

Note. Bolded values show significant differences.

APA NLM
tapraid5/zfn-xhp/zfn-xhp/zfn00413/zfn2975d13z | xppws | S=1 | 5/29/13 | 20:43 | Art: 2011-0122

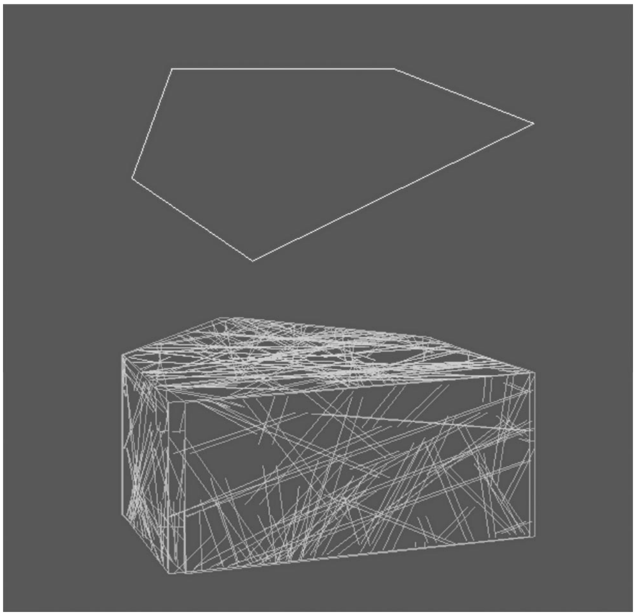

Figure 4. A sample anaglyph display showing a stereoscopic asymmetric polyhedron that would be viewed by a participant wearing red–blue glasses. Also shown above the target object is the 2D response figure that was rescaled by the participant using the arrow keys. A gray-scale version of the display is shown. The actual display was in red and blue.

screen to yield two different slants for the top of the objects: 8° and 16°. Fourth, a single rotation speed (18.75°/s) was used. Fifth, the object remained visible and rotating while participants responded by rescaling the aspect ratio of a 2D outline. Participants hit the space bar when they were finished adjusting the figure, and they were allowed to take as much time as they wished in performing the task. The design allowed participants to take breaks when they preferred. The object for the next trial simply remained on the screen rotating back and forth until they hit the space bar. The three factors—rotation amount (six levels), object aspect ratio (five levels), and slant variation (two levels)—were ordered randomly within each of two repetitions, that is, trials were blocked by repetition, 60 trials per repetition. Thus, a total of 120 judgments was performed by each participant.

Results and Discussion

During debriefing at the end of the experimental session, some of the participants reported that they suspected that they had become confused about the mapping from the 3D object to the 2D response figure. Indeed, this was apparent in their data, because judgments for a given aspect ratio (e.g., 0.8) jumped from one trial to the next from one extreme to the other (e.g., from near 0.8 to near 1.2). This was apparent in the data of six of the participants (all female) whose data was therefore excluded from further analysis, leaving the data of 12 participants.

We performed analysis as in Experiment 2. We performed linear regression of judged aspect ratios on actual aspect ratios separately for each participant and each cell determined by rotation amount and slant, building two new data sets, one of slopes and one of r^2 values. We performed ANOVA on these derived data.

The results are shown in Figure 5, in which mean slopes and r^2 were plotted as a function of amounts of rotation. First, as expected, performance was better with polyhedrons than with elliptical cylinders when the amount of rotation was below 45°. This was true in respect to slopes, which were ≈ 0.9 for the polyhedrons as compared with ≈ 0.85 for elliptical cylinders. Second, also as expected, performance suddenly improved when the amount of rotation reached 45°. This was evident in respect to both slopes (≈ 1.0) and r^2. Third, amounts of rotation greater than 45° did not yield continued increases in the slope (or r^2). Instead, slopes remained ≈ 1.0. Finally, the means shown in Figure 5 were computed over the two slants (as well as participants and repetitions), because slant was not found to affect judgments significantly. F5

We performed a repeated-measures ANOVA on slopes with rotation amount and slant as factors. This yielded only a main effect of rotation amount, $F(5, 55) = 2.7$, $p < .04$. Neither slant ($p > .3$) nor the interaction ($p > .4$) reached significance. We performed the same ANOVA on r^2. This yielded a main effect of rotation amount, $F(5, 55) = 3.7$, $p < .01$, and a significant

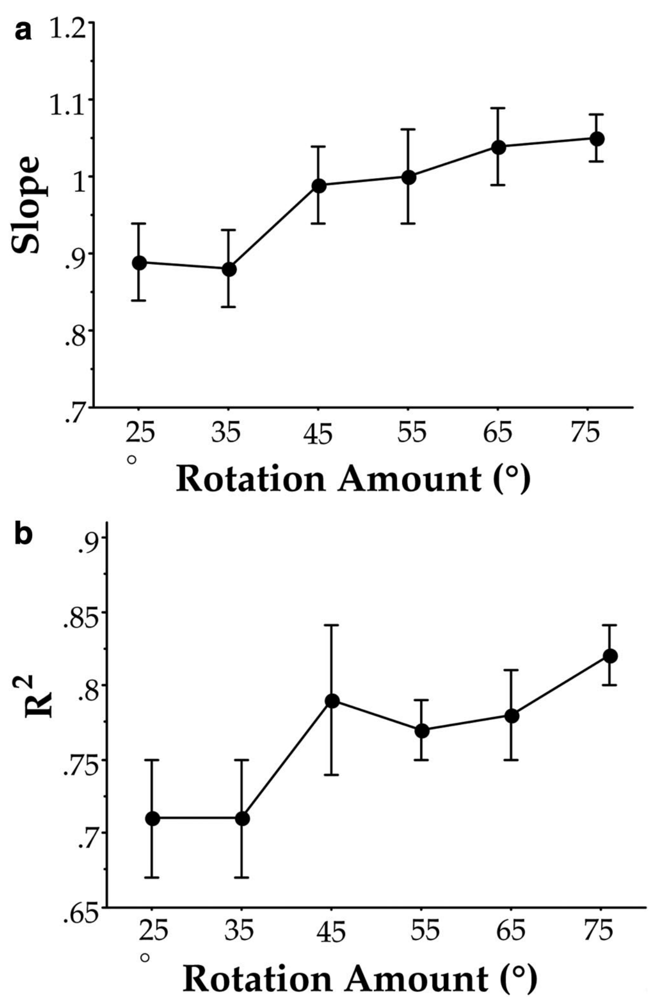

Figure 5. (a) Mean slopes and (b) mean r^2, each plotted as a function of the six different amounts of rotation. Standard error bars represent between-subjects variability.

interaction, $F(5, 55) = 2.6$, $p < .05$. Given the significant interaction, we performed post hoc tests. We found that means were only different as a function of slant with 25° of rotation, otherwise not (Tukey's HSD, $p < .05$). Means for 25° and 35° of rotation were significantly different from that for 75° of rotation (Tukey's HSD, $p < .05$).

Finally, we did analyses to compare gradual and discrete change models. First for slopes, we tested gradual change by regressing the values 0.90, 0.93, 0.96, 0.99, 1.02, and 1.05 on the means shown in Figure 5. The result was an $r^2 = 0.88$ ($p < .01$). We tested discrete change by regressing the values 0.90, 0.90, 1.0, 1.0, 1.0, and 1.0 on the means. The result was an $r^2 = 0.90$ ($p < .01$). Next, for r^2, we tested gradual change by regressing the values 0.70, 0.72, 0.74, 0.76, 0.78, and 0.80 on the means shown in Figure 5. The result was an $r^2 = 0.79$ ($p < .01$). We tested discrete change by regressing the values 0.70, 0.70, 0.80, 0.80, 0.80, and 0.80 on the means. The result was an $r^2 = 0.86$ ($p < .01$). The discrete change model yielded the better fits in respect to the r^2, so the clear appearance of discrete change provided by the graphs was confirmed by the analysis.

Bootstrapping From Affine to Metric Structure With 45° of Perspective Change: A Model

To discuss the theoretical aspects of this issue, it is advantageous to distinguish between three different entities: the distal 3D object, the instantaneous projection of the object onto a projection surface, and the information about the 3D structure of the distal object extracted by the visual system at each instant of time (see F6 Figure 6).

As described previously in this article, there is an increasing body of results supporting the idea that the extracted instantaneous 3D information can be characterized as the result of a mapping from the distal 3D object that only preserves properties invariant over affine transformations. Most metric properties are not extracted. All relative distances between points on the distal object lying on a surface orthogonal to the line of sight are well preserved, relative distances between points on the distal object parallel to the line of sight are likewise well preserved, but the scaling between these two sets of distances is unknown (Todd & Norman, 2003; Todd, Oomes, Koenderink & Kappers, 2001). The effect is that the extracted object has an unknown amount of stretching or compression along the line of sight compared with the distal object. This is true whether the observer has access to monocular information from relative motion between the object and herself, static binocular information, or both. The mathematical basis for this type of extraction in the case of "structure from motion" is well analyzed (e.g., Koenderink & van Doorn, 1991; Shapiro, Zisserman, & Brady, 1995). Interestingly enough, these analyses also show that whether the distal object is instantaneously rigid or not is easy to extract as well, with one important exception. Just as distances along the line of sight contain an unknown (but, for all points, constant) stretch or compression, instantaneous nonrigidity strictly in directions along the line of sight cannot be perceived (Norman & Todd, 1993).

One way of analyzing how larger amounts of object rotation can be utilized by a visual system is to look at how the instantaneously extracted 3D information about an object can be further analyzed over time. The approach does not try to extend the structure from motion analysis of the 2D projections, but instead presumes that process to work only instantaneously without any type of "memory" and leaves the analysis over time to deal with the extracted 3D information only. This is the strategy adopted here. In doing this, we presume that the distal 3D object is rigid over the time period it is being viewed.[3] In the following, we Fn3 presume that there are identifiable texture elements on the viewed object.

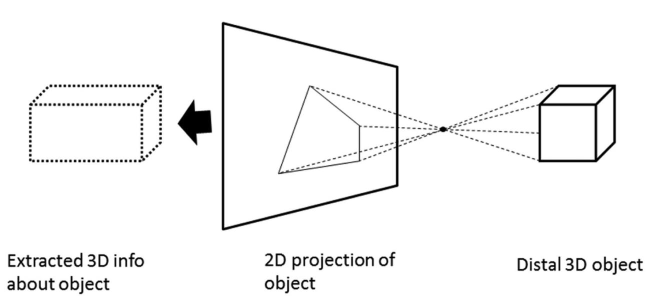

Figure 6. Illustration of the mappings entailed in the perception of metric shape.

The First Instantaneous View: Establishing Orthogonal Axes

An arbitrary instantaneous view of the object is chosen as the first view. The extracted 3D information is a mapping of the distal 3D object, as described earlier, containing an unknown scaling factor for distances along the line of sight. This extracted information has some interesting properties. First, it is easy to identify all texture elements on the extracted object that are equally far away from the observer. An unknown stretching or compression along the line of sight does not affect this type of equidistance judgment. Therefore, two texture elements that have the same distance from the observer, given that two such elements exist, can be chosen. Call these texture elements A and B. They define a line on or through both the extracted and the distal object. This line is by definition *orthogonal* to the line of sight (see Figure 7). Let the F7 line of sight define the direction of the z-axis in a 3D Cartesian coordinate system (x, y, and z). The line between A and B is then parallel to the x–y plane of that coordinate system.

Next, we look for a texture element C lying on a straight line through A and having the same x-coordinate as texture element A. The x-coordinates are not affected by the unknown scaling factor and are thus the same in both the extracted object and the distal

[3] This might seem like a severe restriction, but it is not. No assumptions about the rigidity of all viewed objects is required or intended. The nonrigidity of objects in the environment is easy to perceive (see Bingham & Lind, 2008, for a demonstration and evidence). Instantaneously, a nonrigid component of motion directed strictly along the line of sight will present problems, as described, but for a moving observer, the viewpoint constantly changes. In turn, this means that what was a nonrigid motion strictly along the line of sight a moment ago no longer is so. The only exception would be a nonrigid object having only one direction of nonrigidity that constantly turns in synchrony with the observer's motion to keep the nonrigid motion at all times directed along the observer's line of sight. This is a circumstance that we deem safe to ignore.

object. The elements A, B, and C form a triangle, with a 90° angle between the lines AB and AC.

Subsequent Instantaneous View: Extracting Metric Structure

Fn4 To recover the metric (similarity[4]) structure of the distal object, we next keep track of these texture elements in the extracted object as the distal object moves relative to the observer. As soon as the distal object has moved in such a way that the direction of the line of sight through the distal object is changed compared with the first view, the two texture elements defining the baseline will no longer be equidistant from the observer, and the metric structure of the distal object can be recovered. This is achieved by finding the scaling value along the new line of sight, relative to the current newly extracted object, that brings the angle between the tracked texture elements *in the newly extracted object* to 90°. This can be
F8 achieved in the following manner (please refer to Figure 8).

Call the coordinates of the three points A, B, and C; in this view, A = [xa,ya,za], B = [xb,yb,zb], and C = [xc,yc,zc]. Because we know the distal object is rigid (see previous discussion), we also know that the angle between vectors AB and AC in the distal object still is 90°. Due to the fact that the x- and y-coordinates in the extracted view are identical to the ones in the distal view, and the z-coordinates in the extracted view are scaled by an unknown rotational velocity (e.g., Koenderink & van Doorn, 1991), we also know that by applying the inverse of this unknown rotational velocity to the coordinates za, zb, zc, we can make the angle between the vectors AB and AC also in the extracted object equal to 90°. If we call the inverse of the unknown scaling factor q, we can thus set up the following definitions and equations (because the dot product of two vectors forming a 90° angle is 0):

$$\mathrm{AB} = [(xb - xa), (yb - ya), (q * zb - q * za)]$$

$$\mathrm{AC} = [(xc - xa), (yc - ya), (q * zc - q * za)]$$

$$0 = (xb - xa) * (xc - xa) + (yb - ya) * (yc - ya) + q * q * (zb - za) * (zc - za)$$

Solving for q yields

$$q = (+)\sqrt{\frac{-[(xb - xa) * (xc - xa) + (yb - ya) * (yc - ya)]}{[(zb - za) * (zc - za)]}}$$

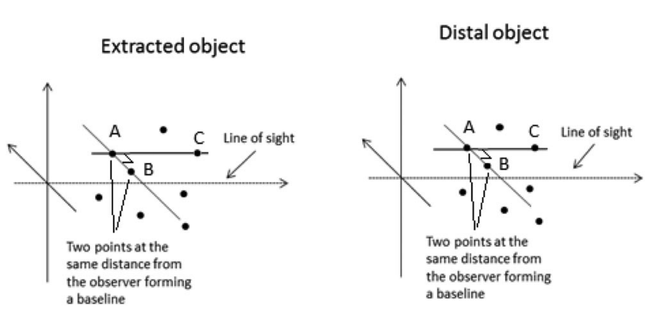

Figure 7. An illustration of the first view in which the two orthogonal axes can be established.

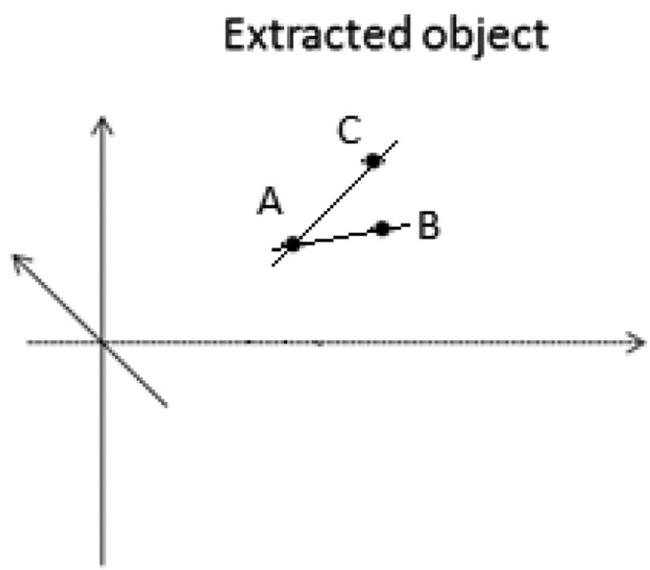

Figure 8. The extracted axes after some rotation of the object, yielding change in perspective.

If the new position of the distal object is such that the direction of the sides of the triangle forming the 90° angle are close to being parallel either to the line of sight or to a line perpendicular to it, the scaling factor is difficult to determine exactly, especially in the presence of noise in the "retinal" measurements. This is because the analyzed angle in these cases will change very little from 90°, even with relatively large changes in the unknown scaling factor (see Figure 9). F9

The Role of Symmetry (Yielded by 45° of Perspective Change) in Bootstrapping Metric Shape

The most informative new position, as a result of rotation, is when the direction of a line, defined through bisection of the angle between the lines AB and AC, is parallel either to the line of sight or to a line perpendicular to it, and the angle between the plane defined by the points A, B, and C and the projection surface is close to 90°. (If it is equal to 90°, then of course, the angle is no longer visible). Bisection of any angle is invariant over affine transformations and, therefore, whether the found angle is bisected is easily determined in a single instantaneous view. (Bisection is a symmetry operation.) In our experiments this bisection criterion corresponds to a rotation of 45°, and this is also the point in our data at which our observers "get" the metric structure. Once found, the no longer unknown scaling factor can be applied to the whole object in this instantaneous view. This reveals the entire object's similarity structure and no further processing is needed. Rotation beyond an amount that brings the angle to be bisected by the x- or y-axis (that is, 45°) is not providing any new information. This is illustrated in Figure 10. F10

[4] Metric shape requires similarity geometry. This falls just below Euclidean geometry in the Klein hierarchy of geometries, and above affine geometry.

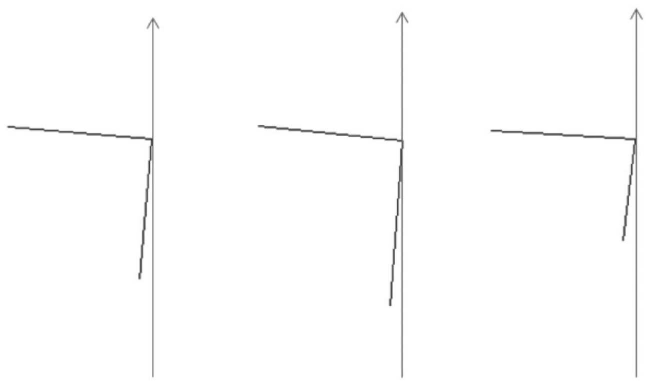

Figure 9. When the amount of rotation away from the original position of the found 90° angle is small, even large amounts of stretching or compression along the line of sight (in this case, no stretching is applied to the leftmost triangle, +25% to the middle triangle, and −35% to the rightmost one) produces comparatively small deviations from 90° in the analyzed angle.

Examining Figure 10 might lead to the conclusion that reliable information exists already after 30° of rotation or so. Although true, this is impossible for an observer to use, because the amount of rotation of the object cannot be judged correctly from a single instantaneous affine 3D view. The unknown scaling factor in the depth dimension leads to the amount of rotation being unknown to the observer. Therefore, an observer cannot be certain when "enough" rotation has taken place, except by waiting for the moment when the 90° angle is bisected by either an axis parallel to the line of sight or an axis orthogonal to it. In our experiments, that corresponds to a 45° rotation. This moment is easy to perceive because of the symmetry available in affine structure, and it also defines the moment when the most information about metric structure is available.

Bingham and Lind (2008) showed that allowing observers to view an object only at the beginning and end of a 90° rotation (that is, occluding vision of the actual continuous rotation) does not allow observers to recover the metric structure of the object. This might seem a bit counter intuitive because a 90° rotation around the y-axis (as in our experiments) reveals metric structure to a system with perfect memory of scenes over time. The unknown stretching of the object along the line of sight, at 0°, would be perceivable after the rotation by 90°, because the distances along that direction would be mapped onto distances in the image plane after the rotation. However, the result is perfectly consistent with the proposed model, given the observations made in reference to Figure 9. The model requires continuous rotation by at least 45° for recovery of metric structure.

General Discussion

Among the numerous studies demonstrating that human observers seem unable to perceive metric shape accurately was one by Lee, Crabtree, Norman, and Bingham (2008). This study showed that such poor perception of 3D metric shape was evident in inaccurate control of feedforward reaches-to-grasp cylindrical objects with various elliptical shapes. The observers in those experiments viewed target objects under conditions that were representative in respect to lighting, use of stereo vision, and free head movements by seated participants. However, Bingham and Lind (2008) then discovered that sufficiently large perspective changes (≥45°) enabled human observers to perceive metric shape well. Their study had used feedforward reaching performance measures. So, Lee and Bingham (2010) replicated the reach-to-grasp task investigated in Lee et al. (2008), but with the addition of large perspective changes that were made available before each reach-to-grasp. The result was that participants performed accurate feedforward reaches-to-grasp, grasps that reflected accurate perception of metric shape.

In that article, Lee and Bingham (2010) wondered about the relevance of such large perspective changes because the conditions of the previous study (Lee et al., 2008) had seemed to be representative. They noted, however, something missing from that design, that is, a way in which perhaps it was not so representative. People typically locomote into the work spaces (one's kitchen or office) in which they then stand or sit and perform manual actions. Locomotion through the surrounds does generate large perspective changes. With this in mind, Lee and Bingham next tested whether the large perspective changes would continue to be effective in allowing accurate feedforward reaches-to-grasp under conditions representative of locomotion into a workspace in which one proceeded to interact with a number of different objects. Participants successively reached-to-grasp a number of objects after viewing them with large perspective changes, and then a delay during which the objects were viewed without the large perspective changes. Performance remained good. Thus, it appears that large perspective changes that occur in optic flow when we locomote through the environment inform or calibrate subsequent visual structure about the location and shapes of objects in the surround-

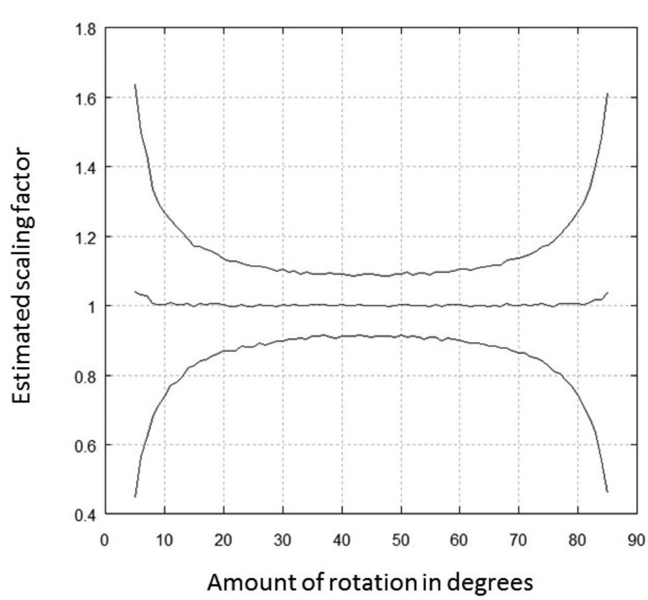

Figure 10. A simulation using our model and adding 10% noise to the extracted variables in each instantaneous view. Each simulation was run 1,000 times for every degree of rotational angle between 5° and 85°. The true scaling factor was 1 and the plot shows the found mean value of all 1,000 observations for each degree as well as ±2 standard deviations.

ings with which we might then manually interact. See Pan, Bingham, & Bingham (2013) for additional evidence to support this idea.

The original Bingham and Lind (2008) study, however, left some questions about the source of improvements in perception of 3D metric shape, questions also not addressed by the subsequent investigations of reaches-to-grasp. One question was whether, in fact, generic 45° (or greater) perspective changes yield good perception of metric shape, or instead whether a change yielding a specific perspective on the objects was responsible, namely, a view at 45° relative to a principal axis of the elliptical cylinders used in the investigations. We tested these possibilities in Experiment 1 and found that the unique perspective was not required, only generic continuous changes in perspective of 45° (or more). A second question arose naturally in the context of this and the previous findings, namely, whether accuracy in perceiving metric shape might gradually improve with increasing amounts of generic perspective change. Was it simply just that progressively more perspective change was progressively better, or instead is a continuous perspective change specifically of 45° or more what is required? We investigated this question in Experiments 2 and 3 and found that improvements in performance did not occur gradually in proportion to increases in the amount of perspective change that was made available. Instead, improvement occurred only when the amount of perspective change was ≈45°.[5] In Experiment 2, we also established that the speed of rotation failed to affect judgments once 45° of perspective change was available. Finally, in Experiment 3, we tested (a) other asymmetric shapes, (b) the potential effects of changes in slant, and, perhaps most important, (c) whether greater internal depth would be seen with increases in perspective change beyond 45°. Given the analyses and results of Pizlo (2008), we expected judgments of the polyhedons to be better with perspective changes less than 45°. This expectation was confirmed. However, we also expected the usual result to appear, namely, sudden improvement in performance once perspective change equaled 45°. This expectation also was confirmed, showing that this result generalizes to shapes other than elliptical cylinders. In fact, Lee, Lind, Bingham, and Bingham (2012) had also investigated perception of metric shape using both symmetric and asymmetric polyhedrons. Their study investigated the use of metric shape for object recognition, and the results were consistent with those in the current study in respect to the various shapes. Large perspective changes yield good perception of metric shape both for the more difficult elliptical cylinders and for the easier polyhedrons, both symmetric and asymmetric. We found that slant failed to affect these results. Finally, judgments not only became accurate once the amount of perspective change reached 45° but also remained so as the amount of change exceeded 45°.

So why would a continuous change in perspective of 45° or more be uniquely effective in allowing perception of 3D metric shape? The answer lies in the twofold change that this brings about. First, a rotation of 45° or more provides the means for an observer, using only the affine instantaneous 3D view, to easily identify the moment when 45° of rotation is reached by observing when the originally identified 90° angle is being bisected by the line of sight or a line orthogonal to it. The symmetry of bisection is readily apprehended in affine structure. Second, knowing that 45° of rotation is present, a simple calculation based on properties present in the affine instantaneous 3D view reveals the hitherto unknown scaling factor.

In addition, with continuous 45° change, this comparison is swept through a uniquely significant portion of an object, enough that, with reflection front to back, the entire structure of the object might be specified. Of course, hidden elements at the back of the object not brought into view by a 45° perspective change could not be perceived via this means and would naturally be judged incorrectly. One can only see that about which one has information available. Shy of this circumstance, however, it appears that we have a solution to the problem of perceiving metric shape. Future investigations should investigate the role of object symmetries in perceiving metric shape via this means.

[5] In Experiment 2, improvement in performance began to appear with 37.5° of rotation, although judgments were not yet fully accurate until 45° of rotation. As described in the model section, 45° of rotation yields recognizably sufficient rotation. However, perception always requires consideration of resolution. What is recognized at 45° may begin to be perceived at 42° or 40°. In the data, we see that the effect begins to appear at 37.5°, but is not yet complete until 45° of rotation.

References

Bingham, G. P., Bradley, A., Bailey, M., & Vinner, R. (2001). Accommodation, occlusion, and disparity matching are used to guide reaching: A comparison of actual versus virtual environments. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 1314–1334. doi:10.1037/0096-1523.27.6.1314

Bingham, G. P., & Lind, M. (2008). Large continuous perspective transformations are necessary and sufficient for accurate perception of metric shape. *Perception & Psychophysics, 70,* 524–540. doi:10.3758/PP.70.3.524

Brenner, E., & van Damme, W. J. M. (1999). Perceived distance, shape, and size. *Vision Research, 39,* 975–986. doi:10.1016/S0042-6989(98)00162-X

Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. *Vision Research, 31,* 1351–1360. doi:10.1016/0042-6989(91)90056-B

Koenderink, J. J. (1990). *Solid shape*. Cambridge, MA: MIT Press.

Koenderink, J. J., & van Doorn, A. J. (1991). Affine structure from motion. *Journal of the Optical Society of America A, Optics, Image & Science, 8,* 377–385. doi:10.1364/JOSAA.8.000377

Lee, Y. L., & 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. doi:10.1007/s00221-010-2323-2

Lee, Y. L., Crabtree, C. E., Norman, J. F., & Bingham, G. P. (2008). Poor shape perception is the reason that reaches-to-grasp are visually guided online. *Perception & Psychophysics, 70,* 1032–1046. doi:10.3758/PP.70.6.1032

Lee, Y. L., Lind, M., & Bingham, G. P. (2013). Perceived 3D 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. doi:10.1016/j.visres.2012.07.013

Lind, M. (1996). Perceiving motion and rigid structure from optic flow: A combined weak- perspective and polar-perspective approach. *Perception & Psychophysics, 58,* 1085–1102. doi:10.3758/BF03206835

Lind, M., Bingham, G. P., & Forsell, C. (2003). Metric 3D structure in visualizations. *Information Visualization, 2,* 51–57. doi:10.1057/palgrave.ivs.9500038

Norman, J. F., & Lappin, J. S. (1992). The detection of surface curvatures defined by optical motion. *Perception & Psychophysics, 51,* 386–396. doi:10.3758/BF03211632

Norman, J. F., & Todd, J. T. (1993). The perceptual analysis of structure from motion for rotating objects undergoing affine stretching transformations. *Perception & Psychophysics, 53,* 279–291. doi:10.3758/BF03205183

Norman, J. F., & Todd, J. T. (1996). The discriminability of local surface structure. *Perception, 25,* 381–398. doi:10.1068/pp.250381

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,* 1057–1069. doi:10.1016/j.visres.2005.09.034

Norman, J. F., Todd, J. T., & Phillips, F. (1995). The visual perception of surface orientation from multiple sources of optical information. *Perception & Psychophysics, 57,* 629–636. doi:10.3758/BF03213268

Pan, J. S., Bingham, N., & Bingham, G. P. (2013). Embodied memory: Effective and stable perception by combining optic flow and image structure. *Journal of Experimental Psychology: Human Perception and Performance*. Advance online publication.

AQ: 8 Pedhazur, E. J. (1982). *Multiple regression in behavioral research: Explanation and prediction* (2nd ed.). New York, NY: Holt, Rinehart & Winston.

Perotti, V. J., Todd, J. T., Lappin, J. S., & Phillips, F. (1998). The perception of surface curvature from optical motion. *Perception & Psychophysics, 60,* 377–388. doi:10.3758/BF03206861

Pizlo, Z. (2008). *3D shape: Its unique place in visual perception*. Cambridge, MA: MIT Press.

Shapiro, L. S., Zisserman, A., & Brady, J. M. (1995). 3D motion recovery via affine epipolar geometry. *International Journal of Computer Vision, 16,* 147–182. doi:10.1007/BF01539553

Tittle, J. S., & Braunstein, M. L. (1993). Recovery of 3-D shape from binocular disparity and structure from motion. *Perception & Psychophysics, 54,* 157–169. doi:10.3758/BF03211751

Tittle, J. S., Todd, J. T., Perotti, V. J., & Norman, J. F. (1995). Systematic distortion of perceived three-dimensional structure from motion and binocular stereopsis. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 663–678. doi:10.1037/0096-1523.21.3.663

Todd, J. T., & Bressan, P. (1990). The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Perception & Psychophysics, 48,* 419–430. doi:10.3758/BF03211585

Todd, J. T., & Norman, J. F. (1991). The visual perception of smoothly curved surfaces from minimal apparent motion sequences. *Perception & Psychophysics, 50,* 509–523. doi:10.3758/BF03207535

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,* 31–47. doi:10.3758/BF03194781

Todd, J. T., Oomes, A. H. J., Koenderink, J. J., & Kappers, A. M. L. (2001). On the affine structure of perceptual space. *Psychological Science, 12,* 191–196. doi:10.1111/1467-9280.00335

Todd, J. T., Tittle, J. S., & Norman, J. F. (1995). Distortions of three-dimensional space in the perceptual analysis of motion and stereo. *Perception, 24,* 75–86. doi:10.1068/pp.240075 AQ: 9

Received March 3, 2011
Revision received April 22, 2013
Accepted April 25, 2013 ■