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

Mats Lind
Uppsala University

Young Lim Lee
Dankook University

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

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 Forssell (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 variability. 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...
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.
We addressed two questions in this first experiment. First, we compared judgments of aspect ratios with either 45° of rotation or 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 $\pm 1$), an interaction vector (computed as the product of the first two vectors), 45° versus 22.5° rotation amount (coded as $\pm 1$), 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 that 45° of continuous rotation yielded good perception of metric shape (slope = 1.00 and intercept $\approx -0.01 \approx 0$), whereas less than this (that is, 22.5°) yielded low slope = .84 and high intercept = .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 1](image1.png)

*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.

![Figure 2](image2.png)

*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.
ement 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 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 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 $\chi^2 [N = 9, df = 4] = 10.0$, $p < .05$). 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 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 $\approx$45°.
The Results of Pairwise Nonparametric Wilcoxon Tests of Rotation Amounts

<table>
<thead>
<tr>
<th></th>
<th>11.25°</th>
<th>22.5°</th>
<th>30°</th>
<th>37.5°</th>
<th>45°</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.25°</td>
<td>x</td>
<td>p = .95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22.5°</td>
<td>x</td>
<td>p = .59</td>
<td>p = .68</td>
<td>p = .03</td>
<td></td>
</tr>
<tr>
<td>30°</td>
<td>x</td>
<td>p = .86</td>
<td>p = .26</td>
<td>p = .02</td>
<td></td>
</tr>
<tr>
<td>37.5°</td>
<td>x</td>
<td>p = .14</td>
<td>p = .03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45°</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 $\chi^2 [N = 9, df = 2] = 8.4, p < .05$). 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, 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.

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 perspective change progressively increased. Instead, judgments of 3D shape did not gradually improve as the amount of rotation 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 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 horizon.

Experiment 3

The Results of Pairwise Nonparametric Wilcoxon Tests of Rotation Speeds

<table>
<thead>
<tr>
<th></th>
<th>10.5°/s</th>
<th>14.8°/s</th>
<th>19.0°/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.5°/s</td>
<td>x</td>
<td>p = .02</td>
<td>p = .62</td>
</tr>
<tr>
<td>14.8°/s</td>
<td>x</td>
<td></td>
<td>p = .07</td>
</tr>
<tr>
<td>19.0°/s</td>
<td>x</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Bolded values show significant differences.
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 $\approx 0.9$ for the polyhedrons as compared with $\approx 0.85$ 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 ($\approx 1.0$) and $r^2$. Third, amounts of rotation greater than 45° did not yield continued increases in the slope (or $r^2$). Instead, slopes remained $\approx 1.0$. 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.

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 interaction.
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 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 presume that there are identifiable texture elements on the viewed object.

**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 line of sight define the direction of the z-axis in a 3D Cartesian coordinate system $(x, y, z)$. 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

---

3This 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

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 achieved in the following manner (please refer to Figure 8).

Call the coordinates of the three points A, B, and C; in this view, \(A = [x_a, y_a, z_a]\), \(B = [x_b, y_b, z_b]\), and \(C = [x_c, y_c, z_c]\). 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 in the distal 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):

\[
AB = [(xb - xa), (yb - ya), (q*zb - q*za)]
\]
\[
AC = [(xc - xa), (yc - ya), (q*zc - q*za)]
\]
\[
0 = (xb - xa)^2 + (yc - ya)^2 + (zc - za)^2 + q^2 * (zb - za)^2
\]

Solving for \(q\) yields

\[
q = (+) \sqrt{\frac{-(xb - xa)^2 + (yc - ya)^2}{(zc - za)^2}}
\]

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).

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.

---

\(^4\) Metric shape requires similarity geometry. This falls just below Euclidean geometry in the Klein hierarchy of geometries, and above affine geometry.
ments viewed target objects under conditions that were represen-
jects with various elliptical shapes. The observers in those exper-
that such poor perception of 3D metric shape was evident in
Lee, Crabtree, Norman, and Bingham (2008). This study showed
ers seem unable to perceive metric shape accurately was one by
angle.
leftmost triangle,
sion along the line of sight (in this case, no stretching is applied to the
the found 90° angle is small, even large amounts of stretching or compres-
When the amount of rotation away from the original position of
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 compres-

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 observ-
ers 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 ob-
jects with various elliptical shapes. The observers in those exper-
iments viewed target objects under conditions that were represen-
tative 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 feed-
forward 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 repre-
sentative. 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 pro-
ceeded 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 9](image_url)

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 compres-

![Figure 10](image_url)

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°. 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 polyhedrons 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) also had 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.

---

References


---

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.


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