Exp Brain Res (2013) 224:551–555
DOI 10.1007/s00221-012-3334-y

RESEARCH ARTICLE

Perceived 3D metric (or Euclidean) shape is merely ambiguous, not systematically distorted

Young Lim Lee · Mats Lind · Geoffrey P. Bingham

Received: 29 August 2012 / Accepted: 30 October 2012 / Published online: 23 November 2012

Abstract Many studies have reported that perceived shape is systematically distorted, but Lind et al. (Inf Vis 2:51–57, 2003) and Todd and Norman (Percept Psychophys 65:31–47, 2003) both found that distortions varied with tasks and observers. We now investigated the hypothesis that perception of 3D metric (or Euclidean) shape is ambiguous rather than systematically distorted by testing whether variations in context would systematically alter apparent distortions. The task was to adjust the aspect ratio of an ellipse on a computer screen to match the cross-section of a target elliptical cylinder object viewed in either frontoparallel elliptical cross-section (2D) or elliptical cross-section in depth (3D). Three different groups were tested using two tasks and two different ranges of aspect ratio: Group 1) 2D(Small) → 3-D(Large), Group 2) 2D(Large) → 3D(Small), Group 3a) 2D(Small) → 3D(Small), and Group 3b) 2D(Large) → 3-D(Large). Observers performed the 2D task accurately. This provided the context. The results showed the expected order of slopes when judged aspect ratios were regressed on actual aspect ratios: Group 1 (SL) < Group 3 (SS and LL) < Group 2 (LS). The ambiguity of perceived 3D aspect ratios allowed the range of aspect ratios experienced in the 2D task to affect the 3D judgments systematically. Nevertheless, when the 2D and 3D ranges of aspect ratios were the same (LL and SS) and the 2D were judged accurately, this did not yield accurate 3D judgments. The results supported the hypothesis that perceived 3D metric shape is merely ambiguous rather than systematically distorted.

Keywords 3D shape perception · Structure-from-motion · Stereo · Affine shape

Y. L. Lee
University of Hong Kong, Hong Kong, China

M. Lind
Uppsala University, Uppsala, Sweden

G. P. Bingham (✉)
Department of Psychological and Brain Sciences, Indiana University, 1101 East Tenth Street, Bloomington, IN 47405-7007, USA
e-mail: gbingham@indiana.edu

Introduction

Many studies have reported that perceived 3D metric (or Euclidean) shape is systematically distorted. Correct metric 3D structure was not recovered from binocular stereopsis (Johnston 1991), monocular structure-from-motion (Norman and Lappin 1992; Norman and Todd 1993; Todd and Bressan 1990; Todd and Norman 1991), a combination of motion and binocular stereopsis (Tittle and Braunstein 1983; Tittle et al. 1995), or the integration of multiple sources of information (Norman and Todd 1996; Norman et al. 1995). Johnston (1991) reported that 3D shapes perceived using binocular stereopsis tend to be systematically compressed at larger distances and systematically expanded at shorter distances. Tittle et al. (1995) also reported that observers consistently adjusted the eccentricity of a cylinder viewed in a monocular structure-from-motion display so its shape was compressed in depth. In addition, Tittle et al. found that there was no significant improvement in judgments when 3D objects were specified by combined stereo and motion information. The integration of multiple sources including shading, texture, highlights, stereo, and motion also yielded no significant improvement in the judgments of 3D metric structure (Norman and Todd 1996; Norman et al. 1995).

It has also been reported that distortions vary with tasks and observers (Lind et al. 2003; Todd and Norman 2003). Todd and Norman (2003) had observers adjust the depth of an object to be same as its width, the planes of dihedral angles to be orthogonal, and the shape of an object to match another at a different viewing distance. The results were variable and inconsistent with changes in viewing distance, orientation, or response task. In a similar vein, Lind et al. (2003) tested perception of 3D metric shape by varying the viewing height and distance of actual wooden cylindrical objects. Observers adjusted the shape of an elliptical outline on a computer screen to be same as the perceived shape of the cylinders. The results were different in different experiments as a function of individual differences, differences of the range of shapes, and differences in task. Observers only judged shape correctly when looking straight down on the tops of the objects so that the elliptical cross-section was frontoparallel. These results imply that perception of 3D metric shape is ambiguous rather than systematically distorted. We investigated this hypothesis under full information conditions by manipulating a contextual 2D judgment task with the expectation that this context would yield systematic variations in judgments of 3D metric shape allowed by the ambiguity of perceived 3D metric shape. Because Lind et al. (2003) found perception of 2D shape (i.e., looking straight down on elliptical cylinders) to be accurate, we used 2D shape information to provide a context that we expected to alter ambiguous 3D metric shape judgments. If 3D aspect ratios are ambiguous, then range of aspect ratios experienced in a preceding 2D task might affect subsequent judgments of 3D shapes. We hypothesized that 2D shape information would act as context to change 3D shape judgments in predictable ways. That is, a small range of aspect ratios in 2D proceeding a large range in 3D (S → L) would yield a decrease in the range of judged aspect ratios in the 3D task, while a large range in 2D proceeding a small range in 3D (L → S) would yield an increase, both in comparison with 2D and 3D tasks with equal ranges of aspect ratio (SS and LL). Using linear regression, we regressed judged aspect ratios on actual aspect ratios to obtain slopes for each group. If we compared these three groups (SL, LS, and SS + LL), we expected the slopes to order as SL < SS + LL < LS.

Methods

Observers

Twenty-six adults at Indiana University participated as observers in this experiment. Eight participated in each of two experimental conditions (five females and three males participated in each condition). Five participated in each of two control conditions (three females and two males participated in each condition). All had normal or corrected to normal vision. All of the participants were naïve as to the purpose of the study and were paid at $7 per hour.

Apparatus and stimuli

Two sets of five elliptical cylinders were used, a small range set and a large range set. The cylinders in the small range set yielded five different depth-to-width aspect ratios: [1] 0.67, [2] 0.83, [3] 1.0, [4] 1.24, and [5] 1.53. The cylinders in the large range set yielded five different aspect ratios: [1] 0.5, [2] 0.67, [3] 1.0, [4] 1.53 and [5] 1.9 (See Fig. 1). The cylinders were hardwood, painted matt black with green phosphorescent dots. The cylinders all were 6.6 cm in width and 4.5 cm in height. Participants sat in front of a table. The stimulus cylinders were shown in either of two tasks, as illustrated in Fig. 2. In a 2D task, the cylinders were placed on a frontoparallel plane, attached to a vertical surface using velcro such that the middle of the cylinder's top surface was at eye level and 50 cm from the participant. Thus, the top of each cylinder was in a frontoparallel plane. In a 3D task, the cylinders were placed on a small box adjusted such that the cylinder's top surface was at 10 cm below and 50 cm viewing distance from the participant's eyes. Thus, the top of each cylinder was just visible from above.

A G3 IMac sat on the table to the left of the stimuli and facing the participant. A white ellipse on a black background appeared on the screen. The ellipse was randomly selected to be either small (1.32 cm) or large (15.8 cm) along the vertically oriented axis when it first appeared. This axis of the ellipse could be altered by pressing two of the arrow keys on the keyboard, one gradually increased the height and the other decreased it. Participants could see both the ellipse presented on the computer monitor and stimulus cylinders at the same time.

Procedure

Before the start of the experiment no cylinders were visible. Each observer's eye height and distance from the stimulus location were set by adjusting chair height and position. Observers always performed the 2D task first followed by the 3D task. In the SL experimental condition, the small range set was used for the 2D task and the large range set was used for the 3D task. In the LS experimental condition, the large range set was used for the 2D task and the small range set was used for the 3D task. In the control conditions, either the small or the large range set was used for both the 2D and 3D tasks (SS or LL conditions). For each trial, a randomly selected cylinder was placed on the

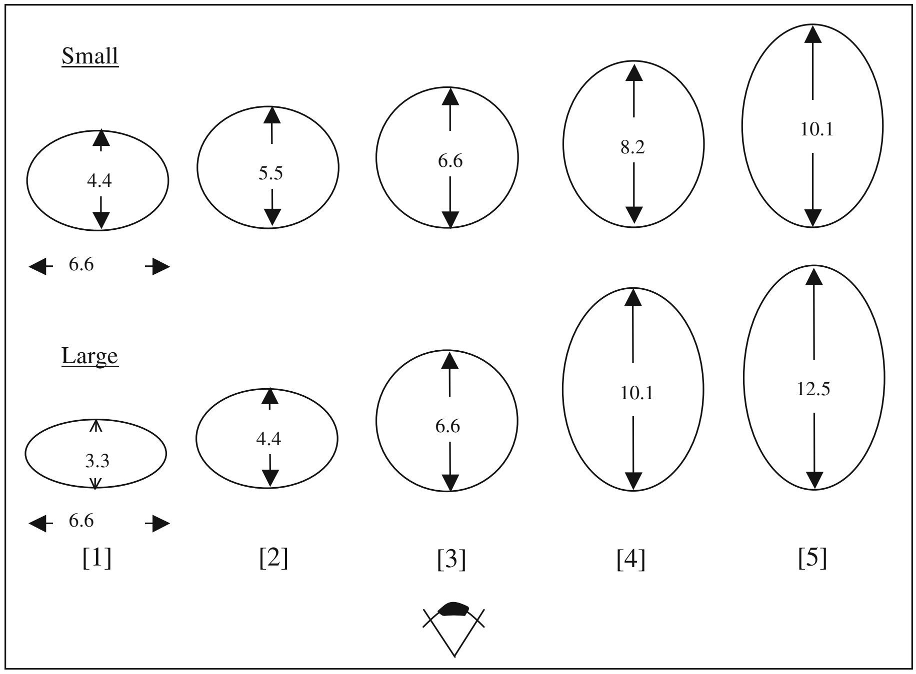

Fig. 1 A schematic representation of the objects used in this experiment. The *cylinders* in the *small range* set were of five different depth-to-width aspect ratios: [1] 0.67, [2] 0.83, [3] 1.0, [4] 1.24, and [5] 1.53. The *cylinders* in the *large range* set were also of five different aspect ratios: [1] 0.5, [2] 0.67, [3] 1.0, [4] 1.53, and [5] 1.9

Fig. 2 The *left panel* illustrates the stimulus in the 2D task and the *right panel* illustrates the stimulus in the 3D task

viewing apparatus while the observer's vision was occluded, and then, the observer was asked to view the target and adjust the eccentricity of the ellipse on the computer screen so as to match it with the eccentricity of the top of the stimulus cylinder. Time to make this judgment was not limited, but observers generally took about 20 s. When the observer was satisfied with the ellipse on the computer screen, he or she hit the space bar to finish the judgment. Then, he or she closed his or her eyes until the experimenter had placed the stimulus for the next trial. Each observer saw five cylindrical shapes in random order three times in each task. A total of 30 judgments were performed by each observer ($5 \times 3 \times 2$).

Results

In this experiment, we investigated whether 2D shape judgments would act as context to systematically affect 3D metric shape judgments. First, as expected, we found that observers performed the 2D task accurately. We regressed judged on actual aspect ratios separately for each condition in the 2D task. All regressions were significant, $p < 0.001$. For LL, $r^2 = 0.97$, $y = 0.92x + 0.02$. For SS, $r^2 = 0.97$, $y = 1.04x - 0.03$. For LS, $r^2 = 0.98$, $y = 1.01x - 0.05$. For SL, $r^2 = 0.95$, $y = 1.07x - 0.08$. Next, we used both simple and multiple regression analyses to analyze judgments in the 3D task. We plotted mean judged aspect ratios as a function of actual object aspect ratios for each condition (SL, LS and control) in Fig. 3 together with lines fit by linear regression. The results showed the expected slope orders, SL < control (SS and LL) <LS. To compare the conditions pairwise (i.e., SL vs. LS, SL vs. control, LS vs. control), we performed multiple regressions on the 3D judgments, using actual aspect ratios, a categorical variable (coding the conditions as ± 1), and an interaction vector as independent variables. (Because there was no significant difference between the results for the two control conditions (SS and LL), we combined them as 'control'.) In each of the three comparisons (all $r^2 = 0.90$ or better and $p < 0.001$), in addition to a significant aspect ratio factor, we found a significant main effect of condition (SL vs. LS: $t(236) = 8.37$, $p < 0.001$; SL vs. SS + LL: $t(266) = 3.87$, $p < 0.001$; and LS vs. SS + LL: $t(266) = 5.47$, $p < 0.001$). Also, in each case, the interaction was significant (SL vs. LS: $t(236) = 9.22$, $p < 0.001$; SL vs. SS + LL: $t(266) = 3.64$, $p < 0.001$; and LS vs. SS + LL: $t(266) = 6.52$, $p < 0.001$). The latter result showed that there was a significant difference in slope between each of the conditions. Thus, each condition was significantly different from the other conditions, in terms of both slope and intercept.

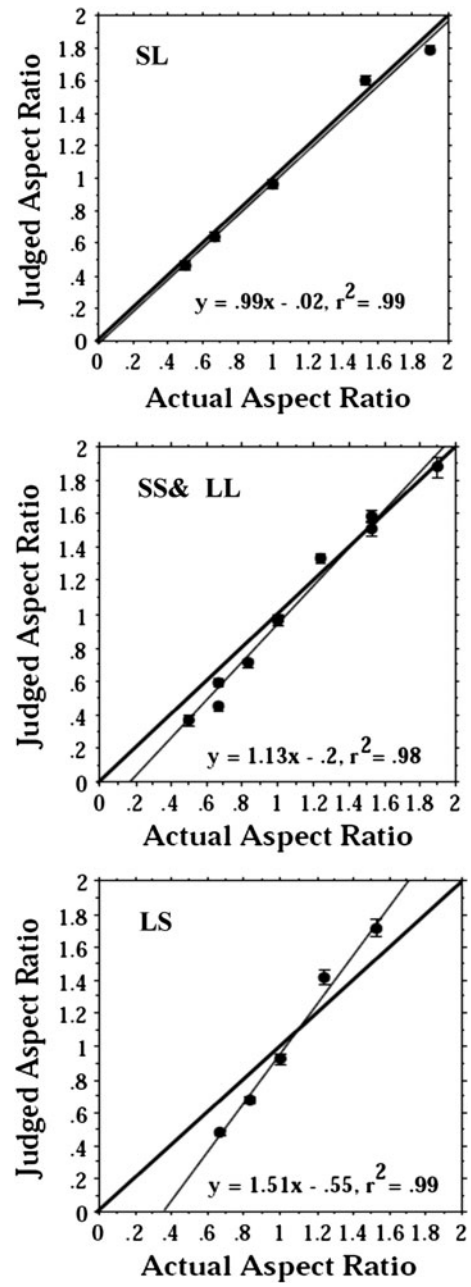

Fig. 3 Mean judged aspect ratios plotted against actual aspect ratios for each of the 3 conditions in the 3D task. A *line* fitted by linear regression is also shown in each case together with a *darker line* representing slope 1 and intercept 0. The *upper panel* shows the SL condition, the *middle panel* shows the control conditions (SS and LL), and the *bottom panel* shows the LS condition. The *error bars* represent the SE of the within-subject variability

Discussion

Proponents of the affine theory of visual space perception argue that visual information is affine, meaning, in part, that the relation between distances or metric extents in different directions is not well specified (e.g., Koenderink 1990; Todd and Bressan 1990). Accordingly, frontoparallel object width could not be reliably compared with object depth using vision. The depth-to-width aspect ratio that measures an aspect of metric 3D shape would be ambiguous. Studies in the literature on perception of 3D shape have well demonstrated that performance is poor, that is, fairly inaccurate and imprecise. Inaccuracy is sometimes reported as systematic distortions (e.g., Johnston 1991). However, other studies find more variability in judgments and less systematicity, at least, when it comes to perception of metric 3D shape. See, for instance, Di Luca et al. (2010) for a review.

The task in the current experiment required observers to judge the 3D metric shape of actual wooden objects viewed sitting in front of them, more or less within reach, on a table in full lighting and information conditions, constrained only by the requirement that the observers stay in their seat and only look at the objects (not, for instance, reach out and touch them). The advantage of this task is that it is representative of everyday object perception. Lind et al. (2003) investigated this same task with different groups of observers who judged different ranges of aspect ratios and objects viewed at different distances and different heights. The resulting judgments yielded slopes, in regressions of judged on actual aspect ratio, that were variable. Sometimes they were significantly greater than 1, other times less than 1 and then again, approximately equal to 1. In each case, the results were systematic enough to yield a significant regression and significant differences in slope, for instance, between cases. Nevertheless, the results varied among cases. These results, like others of the sort reviewed by Di Luca et al. (2010), suggest that the perception of 3D metric shape is indeed ambiguous. Nevertheless, in the separate tests, the results seem to reflect rather systematic distortions in perception. We investigated this situation by manipulating a contextual perceptual task with the expectation that subsequent judgments of 3D metric shape would vary systematically as a result. Ambiguity of perceived 3D shape was expected to allow the contextual task to affect the judgments systematically. If the judgments are readily affected and altered in this way, then we conclude that other contextual factors yield task-specific systematicity in other studies despite the ambiguity of perceived 3D metric shape.

Accurate judgments of 2D shape provided the context for judgments of 3D metric shape. A large range of aspect ratios following a small (SL) yielded a decrease in the range of judged aspect ratios while a small range following a large (LS) yielded an increase, when those judgments were compared to judgments in the equal range conditions. The results were as predicted by the hypothesis that perception of 3D metric shape is ambiguous, rather than systematically distorted, and that contextual information yields systematic patterns of apparent distortion.

Notably, even though participants succeeded in judging 2D shapes accurately, in the conditions where the ranges of aspect ratios in the 2D and 3D tasks were the same, the judgments of 3D metric shape were still inaccurate! As shown in Fig. 3, the slope ≈ 1 and intercept ≈ 0 in the SL condition and these were significantly different from the slope ≈ 1.13 and intercept ≈ -0.2 in the control condition (i.e., LL and SS, where the 2D and 3D aspect ratios were the same). So, the range of aspect ratios in the 2D task definitely affected the range of judged aspect ratios in the following 3D task, but the 2D range did not act as information allowing the 3D range to be judged accurately when the ranges were the same. (The question is exactly how information about the range of 2D aspect ratios might be used to scale the following 3D aspect ratios.) Similarly, Lee et al. (2008) found that accurate feedback from grasping object widths and depths did not yield accurate pre-shaping of grasps in subsequent reaches-to-grasp the same set of objects (similar to those used in the current experiment). Such feedback was found to calibrate accurate reaches-to-grasp in respect to the distance and size of target objects, but not the 3D metric shape of those objects.

Finally, note that if two 2D tasks were to be performed in sequence with the first providing the context for the second and each involving a different range of aspect ratios, the 2D aspect ratios should be judged equally well in both cases simply because there is no ambiguity in the 2D case.

We have shown that perception of Euclidean shape is ambiguous. It has been shown in other work that tasks that require, for instance, Affine scaling rather than Euclidean, yield results showing the perception is accurate (e.g., Todd et al. 2001). So, the bottom line is that perception of 3D metric shape is ambiguous, and reliability of judgments in specific tasks (requiring Euclidean scaling) and situations comparable to that investigated in the current study (i.e., seated observer with free viewing and full information) should be attributed to extra-perceptual factors.

References

Di Luca M, Domini F, Caudek C (2010) Inconsistency of perceived 3D shape. Vision Res 50:1519–1531

Johnston EB (1991) Systematic distortions of shape from stereopsis. Vision Res 31:1351–1360

Koenderink JJ (1990) Solid shape. MIT Press, Cambridge

Lee Y, Crabtree CE, Norman JF, Bingham GP (2008) Poor shape perception is the reason reaches-to-grasp are visually guided online. Percept Psychophys 70:1032–1046

Lind M, Bingham GP, Forsell C (2003) Metric 3D structure in visualization. Inf Vis 2:51–57

Norman JF, Lappin JS (1992) The detection of surface curvatures defined by optical motion. Percept Psychophys 51:386–396

Norman JF, Todd JT (1993) The perceptual analysis of structure from motion for rotating objects undergoing affine stretching transformations. Percept Psychophys 53:279–291

Norman JF, Todd JT (1996) The discriminability of local surface structure. Perception 25:381–398

Norman JF, Todd JT, Phillips F (1995) The visual perception of surface orientation from multiple sources of optical information. Percept Psychophys 57:629–636

Tittle JS, Braunstein ML (1983) Recovery of 3-D shape from binocular disparity and structure from motion. Percept Psychophys 54:157–169

Tittle JS, Todd JT, Perotti VJ, Norman JF (1995) Systematic distortion of perceived three-dimensional structure from motion and binocular stereopsis. J Exp Psychol Hum Percept Perform 21:663–678

Todd JT, Bressan P (1990) The perception of 3-dimensional affine structure from minimal apparent motion sequences. Percept Psychophys 48:419–430

Todd JT, Norman JF (1991) The visual perception of smoothly curved surfaces from minimal apparent motion sequences. Percept Psychophys 50:509–523

Todd JT, Norman JF (2003) The visual perception of 3D shape from multiple cues: are observers capable of perceiving metric structure? Percept Psychophys 65:31–47

Todd JT, Oomes AHJ, Koenderink JJ, Kappers AML (2001) On the affine structure of perceptual space. Psychol Sci 12:191–196