Stereo improves 3D shape discrimination even when rich monocular shape cues are available

Young Lim Lee Department of Psychology, University of Hong Kong, Hong Kong

Jeffrey A. Saunders Department of Psychology, University of Hong Kong, Hong Kong

We measured the ability to discriminate 3D shapes across changes in viewpoint and illumination based on rich monocular 3D information and tested whether the addition of stereo information improves shape constancy. Stimuli were images of smoothly curved, random 3D objects. Objects were presented in three viewing conditions that provided different 3D information: shading-only, stereo-only, and combined shading and stereo. Observers performed shape discrimination judgments for sequentially presented objects that differed in orientation by rotation of 0°–60° in depth. We found that rotation in depth markedly impaired discrimination performance in all viewing conditions, as evidenced by reduced sensitivity (d') and increased bias toward judging same shapes as different. We also observed a consistent benefit from stereo, both in conditions with and without change in viewpoint. Results were similar for objects with purely Lambertian reflectance and shiny objects with a large specular component. Our results demonstrate that shape perception for random 3D objects is highly viewpoint-dependent and that stereo improves shape discrimination even when rich monocular shape cues are available.

Keywords: shape discrimination, viewpoint invariance, cue combination, stereo, shading

Citation: Lee, Y. L., & Saunders, J. A. (2011). Stereo improves 3D shape discrimination even when rich monocular shape cues are available. *Journal of Vision, 11*(9):6, 1–12, http://www.journalofvision.org/content/11/9/6, doi:10.1167/11.9.6.

Introduction

There are many sources of information about 3D structure, including both monocular cues and binocular cues. Figure 1 shows an example of a smoothly curved object specified by multiple 3D cues. The 3D shape of the object can be readily perceived from a monocular image due to the information provided by shading, texture, and occlusion contours. When the left and right eyes' views are viewed stereoscopically, binocular disparities provides further 3D information. The availability of multiple 3D cues raises a general question of how such cues are combined for a given perceptual task. Some previous studies have investigated perception of depth from shading and stereo and found that stereo contributes to reliable depth judgments by cooperating with shading cue (Bülthoff & Mallot, 1988; Vuong, Domini, & Caudek, 2006).

The contribution of stereo information is less clear for other perceptual tasks related to 3D shape, such as shape discrimination or shape recognition. Norman, Todd, and Orban (2004) investigated the precision of 3D shape discriminations for images with various combinations of visual shape cues (e.g., shading, specular highlights, texture, motion, stereo). Stereo information improved shape discrimination when combined with some monocular shape cues (e.g., texture, Lambertian shading), but when monocular cues were strong (shading with specularities), shape discrimination thresholds were similar with and without stereo. These results suggest that stereo may not contribute to shape discrimination if sufficient monocular information is available.

However, stereo might be more important for perceiving invariant shape across changes in viewpoint. When a 3D object is viewed from different angles, the projected images can be quite different, which poses a challenge for perceiving shape constancy. Figure 2 shows the same 3D object at different viewing orientations, which differ by rotation in depth by 0°–60°. When the rotation angle is small (15°), most of the same features remain visible, and the 2D image is similar to the original view (e.g., similar occlusion contours). However, with a larger (60°) rotation in depth, the projected image becomes more distinct. There is no simple correspondence between features in the 0° and 60° images and no apparent similarity between the occlusion contours. Subjectively, it is harder to perceive these images to be the same 3D shape. In this more difficult situation, additional 3D information from stereo might facilitate shape recognition.

A number of object recognition studies have tested whether stereo information can facilitate recognition across different viewpoints. Some have found that stereo viewing reduces viewpoint costs. This has been observed for recognition of bent paperclips (Burke, 2005; Edelman & Bülthoff, 1992) and shaded tube-like objects (Bennett & Vuong, 2006). Although observers were generally poor

doi: 10.1167/11.9.6 Received February 8, 2011; published August 16, 2011 ISSN 1534-7362 © ARVO

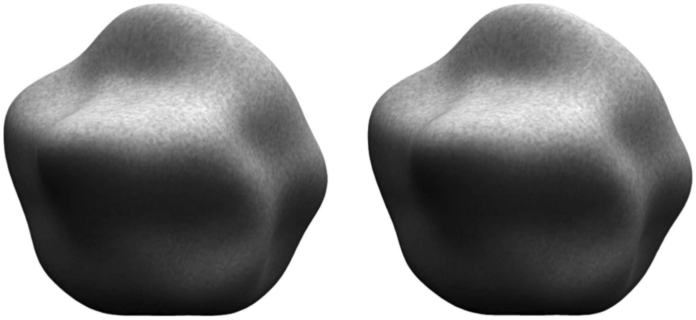

Figure 1. Example of a 3D shape defined by both monocular and binocular cues. The two images are right eye and left eye perspective views of a smoothly curved solid object. The monocular images are sufficient to produce a strong percept of 3D shape due to information from shading, texture, and occlusion contour. When the images are cross fused, binocular disparities provide an additional stereoscopic cue to 3D shape.

at generalizing to novel views, viewpoint costs were reduced with stereo information. However, for these types of stimuli, a monocular image is not very informative for recognition. The observed stereo advantage could be due to limited monocular information.

In the studies of face recognition across different viewpoints, Burke, Taubert, and Higman (2007) found a stereo advantage, while Liu, Ward, and Young (2006) found no benefit of stereo. A possible factor is the different amounts of viewpoint rotation tested in these studies. In Burke et al.'s study, which observed a stereo benefit, viewpoint was changed by 45° or 90°. In contrast, Liu et al. tested viewpoint changes of ±7° and 35° and observed no stereo benefit. Taken together, these results suggest that stereo is advantageous for face recognition only across large viewpoint changes.

Pasqualotto and Hayward (2009) found a stereo disadvantage, rather than advantage, for recognizing familiar objects across changes in viewpoint. They suggest that 2D projected outlines are an important cue for recognition and that 2D outlines are harder to recover and compare with stereo viewing. To the extent that recognition is based on 2D images, stereo might impair performance by interfering with the encoding of 2D information. Results of Liu et al. (2006) support this hypothesis. Liu et al. independently varied whether stereo information was available at encoding and test stages and found that face recognition was less accurate when encoding and test conditions were incongruent (mono→stereo, stereo→mono) than when both faces were viewed in the same way (mono→mono, stereo→stereo).

Some researchers have argued that monocular information is sufficient for shape constancy across changes in viewpoint when objects have structural constraints like symmetry or planarity faces (Chan, Stevenson, Li, & Pizlo, 2006; Liu & Kersten, 2003; Pizlo, Li, & Steinman, 2008; Pizlo & Stevenson, 1999). Pizlo and Stevenson (1999) observed good shape constancy across large viewpoint changes (90°) for polyhedral objects with regularity constraints, even though the stimuli were monocular line drawings. Chan et al. (2006) compared shape discrimination of similar objects under monocular and binocular viewing conditions. In both conditions, structural constraints were a strong determinant in performance, though they also observed a small but consistent benefit from binocular viewing.

In summary, stereo has been shown to reduce viewpoint costs in situations where monocular information is weak or insufficient for the task (Bennett & Vuong, 2006; Burke, 2005; Edelman & Bülthoff, 1992), but it is less clear whether stereo facilitates invariant object recognition in conditions with rich monocular information (Chan et al., 2006; Liu et al., 2006).

Present study

We investigated 3D shape discrimination across changes in viewpoint in conditions that provide various combinations of monocular and binocular 3D shape information. Our goals were to test (1) whether stereo facilitates shape discrimination when rich monocular information is available and (2) whether the benefit from stereo depends on change in viewpoint.

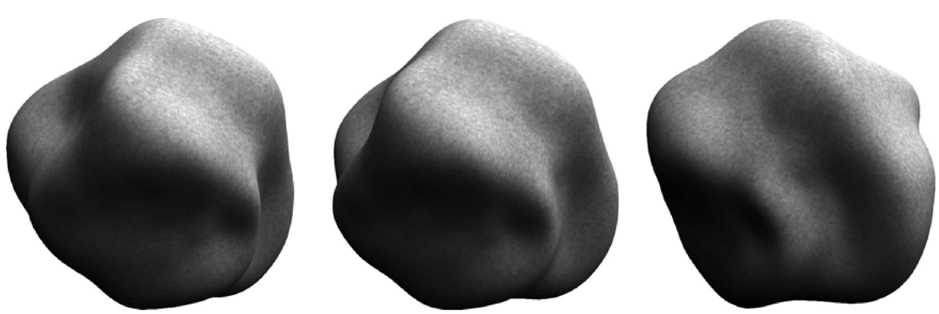

Figure 2. The same 3D object viewed from different orientations. The left image shows the base view of an object. The middle image shows the object after a 15° rotation around the vertical axis, and the right image shows the object after a 60° rotation.

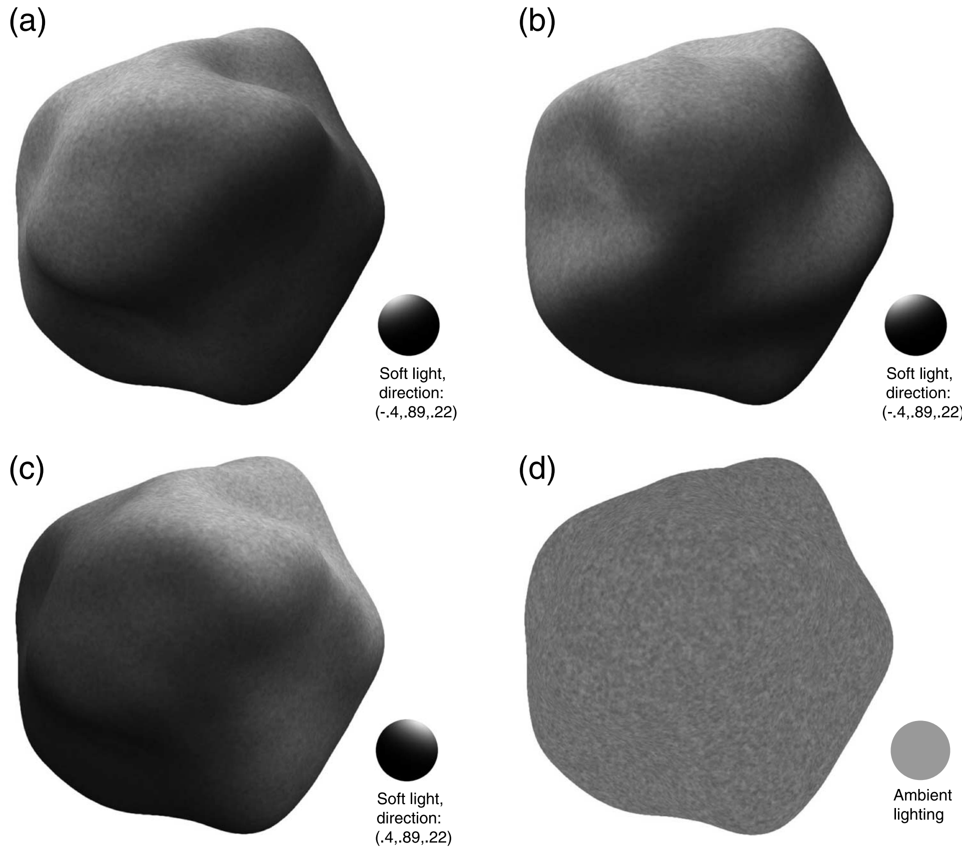

Figure 3. Illustration of shape and lighting conditions. (a, b) Pairs of shapes were constructed to have approximately the same occlusion contours when viewed from a base orientation. The top images show a sample pair. (a, c) For conditions with shading, light source direction was varied across standard and test images. The left images show the same object illuminated by the two light sources used in the experiment. Small spheres show the illumination functions. (d) For the stereo-only condition, simulated illumination was ambient. Texture provided only a weak monocular shape cue.

Stimuli were images of random, smoothly curved 3D objects as shown in Figure 3. There are novel and unfamiliar objects, yet are perceived as having a definite 3D shape due to the information provided by shading and the occlusion contour. Unlike the stimuli used in some previous studies (Bennett & Vuong, 2006; Burke, 2005; Edelman & Bülthoff, 1992), these stimuli elicit a strong subjective sense of 3D structure from monocular information alone.

Because our aim was to study 3D shape discrimination, our conditions were designed to prevent observers from using a purely 2D strategy. Liu, Collin, and Chaudhuri (2000) found that the apparent contribution of stereo to face recognition depended on whether monocular images were similar or reversed polarity, which demonstrates the importance of controlling for image similarity. We constructed pairs of objects to have the same occlusion contour when viewed from a base orientation, so a comparison of 2D outlines would not be informative for the discrimination task. We also varied the light source direction across views to discourage direction comparison of 2D luminance patterns. Figure 3 shows examples of comparison images for "same" and "different" trials, without viewpoint rotation.

We tested shape discrimination across various amounts of viewpoint change, for both monocular and binocular viewing conditions. Norman et al. (2004) found that rich monocular information was sufficient for reliable shape discrimination in the situation where viewpoint remains constant but illumination is varied. When discriminating objects across changes in viewpoint, however, stereo might be more beneficial. Norman, Barthlomew, and Burton (2008) found that structure-from-motion information provided comparatively more benefit for shape discrimination across large viewpoint changes than for small viewpoint changes. There might be an analogous interaction for 3D shape discrimination based on shading and stereo information.

Experiment 1

Experiment 1 tested 3D shape discrimination in three viewing conditions that provided different 3D information: shading-only, stereo-only, and combined shading and stereo. Stereo information might improve shape discrimination, or the rich monocular information in the shading-only condition might be sufficient for reliable judgments.

Observers performed shape discrimination for sequentially presented objects, and viewpoint of the test object was varied by 0°, ±15°, ±30°, or ±60°. Overall, performance would be expected to decrease with viewpoint change. Additional information from stereo could potentially reduce viewpoint costs.

Methods

Apparatus and stimuli

Twelve pairs of 3D shapes were used for the experiment. Random 3D shapes were generated using a method similar to Norman, Swindle, Jennings, Mullins, and Beers (2009). A sequence of ten sinusoidal distortions was applied to a unit sphere, each in a different randomly chosen direction. The distortion was of the form $(x, y, z) \rightarrow (x, y, z + 0.075\sin(1.6x))$, with the coordinate frame rotated to a different random 3D orientation for each sinusoidal distortion. Shapes were scaled to have an average radius of 10 cm, corresponding to about 9–10° of visual angle at the simulated viewing distance of 60 cm. A large pool of shapes was generated in this manner. We then computed the projected boundary contour for each shape and selected pairs of shapes with similar contours. These pairs of 3D shapes were radially distorted around the line of sight axis to force their projected contours to be closely matched. For "different" trials, the test image was a view of the matched object from a pair. The radial distortion was required to be smooth, and global shapes were approximated by a 15th degree spherical harmonic expansion, so contours were not exactly identical. We manually excluded some pairs of shapes for various reasons: Contour matching was not successful or produced visible artifacts, a base image was near an accidental view, or a shape had an atypical amount of 3D variation (i.e., very bumpy, very smooth). Figure 3 shows an example of a matched pair.

Objects were simulated to have Lambertian surface reflectance and a homogeneous surface texture. The texture modulated the object's reflectance over a range from 0.6 to 1.0. To create the texture for an object, we first generated a random set of 40,000 points on the surface that were uniformly distributed on the surface. These points were used as centers to form a Voronoi tiling of the surface, and each tile was assigned a random reflectance. The resulting surface pattern was then approximated as a cube-map texture, with a resolution of 1024 × 1024 for each side. We repeated this procedure ten times per object and averaged the results to get the final cube-map texture used to render the object.

There were three shape cue conditions: (1) shading-only, (2) stereo-only, and (3) combined shading and stereo cues. For the conditions with shading information, we simulated a diffuse point source light at infinite distance. The illumination map was a Gaussian distribution with width $s = 30°$, centered around the light source direction. Two light source directions were used: (−0.40, 0.89, 0.22) and (0.40, 0.89, 0.22). The light source directions for standard and comparison images on a trial were always different to ensure that the task could not be performed based on 2D image similarity. For the stereo-only condition, illumination was simulated to be ambient, with brightness equal to 60% of the maximum illumination in the shading conditions. This brightness approximately matches the mean luminance contrast of the texture in the conditions with and without shading.

A mirror stereoscope was used to present images. Observers viewed a pair of LCD monitors (Dell SP2208WFP) through two semi-silvered mirrors positioned near the eyes and slanted 45° relative to the line of sight. The monitors had a 47 cm × 29.5 cm visible region with 1680 × 1050 resolution and a frame rate of 60 Hz. The monitors were positioned so that their virtual surfaces (viewed through the mirror) were frontal relative to the viewer and aligned at a distance of 60 cm. We measured interpupillary distance for each observer to compute the accurate stereo projections when rendering. In the shading-only condition, the same apparatus was used, but observers wore an eye patch covering their non-dominant eye.

Procedure

Observers were presented with sequential standard and test images and judged whether the images were the same or different 3D objects. On half of the trials, the images showed the same object, and on the other half of the trials, the test image showed a view of the other object in a pair (see above). The viewing orientation of the test object varied: 0°, ±15°, ±30°, or ±60° relative to the base orientation of an object. A fixation point at the screen distance was shown before presenting each image. The standard image was presented for 2 s, followed by around 600-ms interval with a noise mask, and then the test image. The test image remained visible until observers made a response, and observers had the option to repeat the sequence of images on a trial before making a judgment.

The three shape cue conditions were tested in separate experimental sessions on different days. The order of shape cue conditions was counterbalanced across observers. In each session, the observer first performed a block of 28 practice trials with feedback to become familiar with the stimuli and task. Practice trials used one pair of objects, which was not used in the experimental blocks. No feedback was given during the experimental blocks. An experimental block consisted of 336 trials: one "same" trial and one "different" trial for each combination of object and viewing orientation. The order of conditions within an experimental block was randomized. Trials were self-paced, and the experimental block took approximately 60–70 min to complete.

Participants

Ten adults (three males and seven females) at the University of Hong Kong participated in this experiment. All had normal or corrected-to-normal vision and passed a screening test for stereo acuity. All participants were naive as to the purpose of the study and were paid for participating. All procedures were approved by and conform to the standards of the Human Research Ethics Committee for Non-Clinical Faculties.

Results

We computed sensitivity (d') and criterion parameters from each observer's judgments in a condition. Results from positive and viewpoint rotations (e.g., 15° and −15°) were combined for analysis. Figure 4 shows mean d' (left) and mean criterion (right), averaged across observers, as a function of viewpoint for the three shape cue conditions.

An ANOVA on d' found a main effect of shape cue condition ($F(2,18) = 8.83$, $p = 0.002$), as well as a main effect of viewpoint ($F(3,27) = 70.58$, $p < 0.001$) and a significant interaction ($F(6,54) = 3.46$, $p = 0.006$). Comparisons between the three viewing conditions revealed that overall performance was significantly worse in the shading-only condition than the combined shading and stereo condition ($F(1,9) = 13.6$, $p = 0.005$) and also worse than in the stereo-only condition ($F(1,9) = 8.48$, $p = 0.017$). There was no reliable difference between performance in the stereo-only and combined shading and stereo conditions ($F(1,9) = 4.2$, $p = 0.07$).

We conducted further tests to explore the stereo benefit for different amounts of viewpoint rotation. Sensitivity was significantly higher in the combined condition than the shading-only condition at all viewpoints except the largest (0°: $p = 0.003$, 15°: $p = 0.001$, 30°: $p = 0.021$, 60°: $p = 0.156$), and the trend in the 60° condition was in a consistent direction. Thus, stereo provided an overall benefit for shape discrimination, in conditions both with and without viewpoint change.

In all shape cue conditions, sensitivity decreased with viewpoint rotation. This was confirmed statistically by testing for a linear effect of viewing orientation in each shape cue condition. A significant linear effect was observed in the shading-only condition ($F(1,9) = 79.2$, $p < 0.001$), the stereo-only condition ($F(1,9) = 72.7$, $p < 0.001$), and the combined stereo and shading condition ($F(1,9) = 120$, $p < 0.001$). There were also significant interactions indicating that the linear effect of viewpoint in the shading-only condition was smaller than in the combined cue condition ($F(1,9) = 9.4$, $p = 0.005$) and the stereo-only condition ($F(1,9) = 5.4$, $p = 0.028$). Viewpoint had an equivalent effect in the stereo-only and combined conditions ($F(1,9) = 0.0$, $p = 0.98$).

We also analyzed criterion parameters, as shown in the right panel of Figure 4. An ANOVA found a significant main effect of viewpoint ($F(3,27) = 217$, $p < 0.001$) as well as an interaction ($F(6,54) = 13.6$, $p < 0.001$), but no main effect of shape cue condition ($F(2,18) = 3.2$, $p = 0.064$). The effect of viewpoint was due to a bias toward judging objects as "same" when the viewpoint was the same or changed by a small amount and a bias toward judging objects as "different" when the viewpoint changed by a

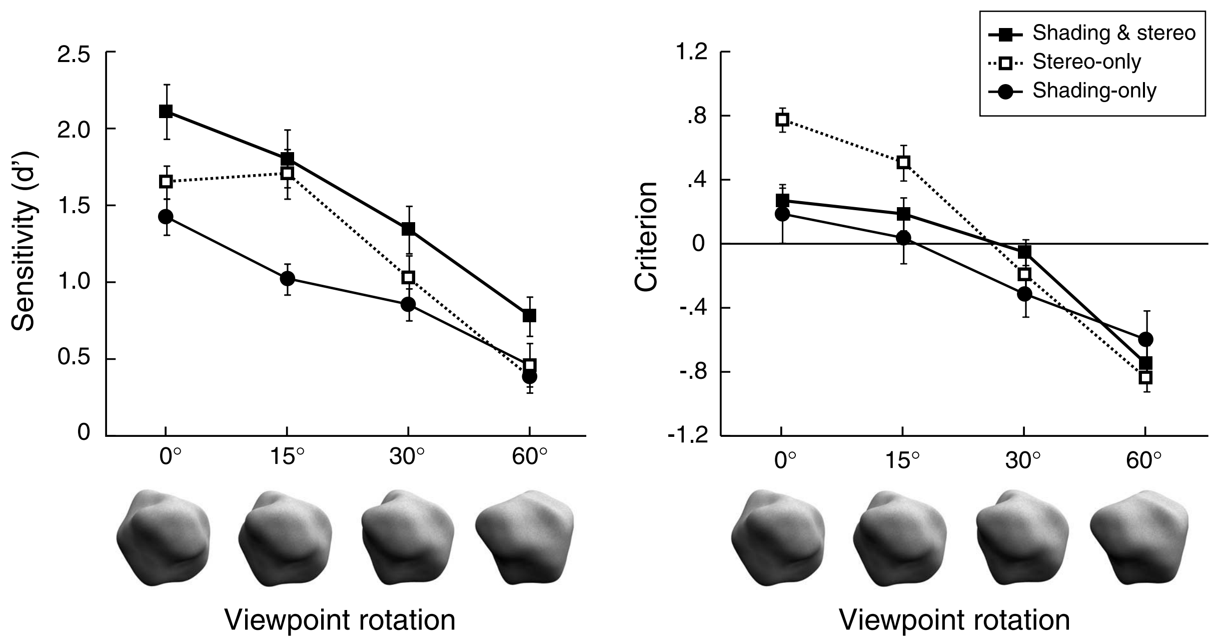

Figure 4. Results of Experiment 1. The graphs show (left) mean sensitivity and (right) criterion, averaged across observers, as a function of change in viewing orientation. The three lines on each graph correspond to the shape cue conditions: shading-only (circles), stereo-only (open squares), and combined shading and stereo (filled squares). Error bars depict ±1 standard error.

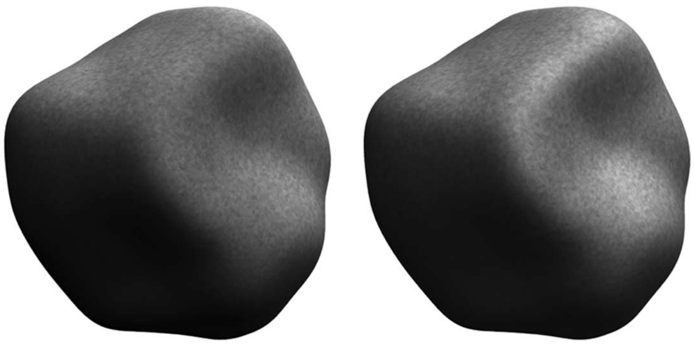

Figure 5. The same 3D object rendered with either purely Lambertian reflectance, as in Experiment 1 (left), or a combination of Lambertian and specular components, as in Experiment 2 (right).

larger amount. We modeled this bias as a linear effect of viewing orientation. The linear effect was significant for all three conditions: shading-only ($F(1,9) = 99.7$, $p < 0.001$), stereo-only ($F(1,9) = 445$, $p < 0.001$), and combined ($F(1,9) = 175$, $p < 0.001$). To explore the interaction, we compared across the shape cue conditions. The response bias was larger in the stereo-only condition than in the shading-only condition ($F(1,9) = 53.1$, $p < 0.001$) and the combined condition ($F(1,9) = 44.7$, $p < 0.001$), while the shading-only and combined conditions produced equivalent response biases ($F(1,9 = 2.64$, $p = 0.12$).

Discussion

We found that rotation in depth markedly impaired discrimination performance in all viewing conditions, as evidenced by reduced sensitivity (d') and increased bias toward judging same shapes as different. In the shading-only condition, viewpoint rotation of 15° was sufficient to reduce d' to one. Although the shading-only images produce a strong subjective percept of 3D shape, observers were not able to reliably discriminate shapes across changes in viewpoint in this condition. Even in the full cue condition with both shading and stereo information, viewpoint rotation of 60° resulted in a d' of less than one and significant response bias. The addition of stereo information improved shape discrimination, but performance remained highly viewpoint-dependent.

The benefit from stereo was not limited to conditions with large change in viewpoint and did not systematically increase with viewpoint. We hypothesized that stereo might provide more benefit with large change in viewpoint, as Norman et al. (2008) observed for the case of structure from motion combined with static monocular cues. However, we observed a small interaction in the opposite direction. This discrepancy may be due to the fact that we controlled for the information from boundary contours, which would be most beneficial across small changes in viewpoint.

Our conditions with no change in viewpoint can be compared to those of Norman et al. (2004). Norman et al. tested shape discrimination across changes in light source direction, but no change in viewpoint, for stimuli that provided various sources of 3D information. For shaded objects with Lambertian reflectance, as in our Experiment 1, Norman et al. similarly observed a stereo benefit. However, they found that objects rendered with specularities yielded better shape discrimination than objects rendered with Lambertian shading and that stereo provided no detectable benefit when specularities were present. It is possible that the lack of specularities in our stimuli reduced the ability to discriminate 3D shape and that stereo would provide less benefit with richer monocular information. We tested this possibility in Experiment 2.

Experiment 2

Previous evidence suggests that Lambertian shading, as used in Experiment 1, is less effective at conveying 3D shape than shading that includes specularities (Norman et al., 2004; Todd, Norman, Koenderink, & Kappers, 1997). Experiment 2 replicated the previous experiment but using a surface reflectance model that includes both Lambertian and specular components. With richer monocular shape information, performance might be less dependent on viewpoint, and stereo might provide less benefit.

Methods

Apparatus and procedure were identical to those in the previous experiment. Stimuli were also the same as those used in the previous experiment except the surface reflectance model had a specular component. We used a Phong model with a 5% ambient component, 70% Lambertian component, and 25% specular component with exponent 100. An example is shown in Figure 5.

Eleven adults (five males and six females) aged from 19 to 27 at the University of Hong Kong participated in this experiment. We excluded the data of one observer from analysis because of unusually poor performance even with no change in viewpoint, which suggested that the observer did not understand the task. All participants had normal or corrected-to-normal vision and passed a screening test for stereo acuity. All were naive as to the purpose of the study and were paid for participating. The procedures were approved by and conform to the standards of the Human Research Ethics Committee for Non-Clinical Faculties.

Results and discussion

Figure 6 shows mean sensitivity (left) and criterion (right), averaged across observers, as a function of change in viewpoint for each of the three viewing conditions.

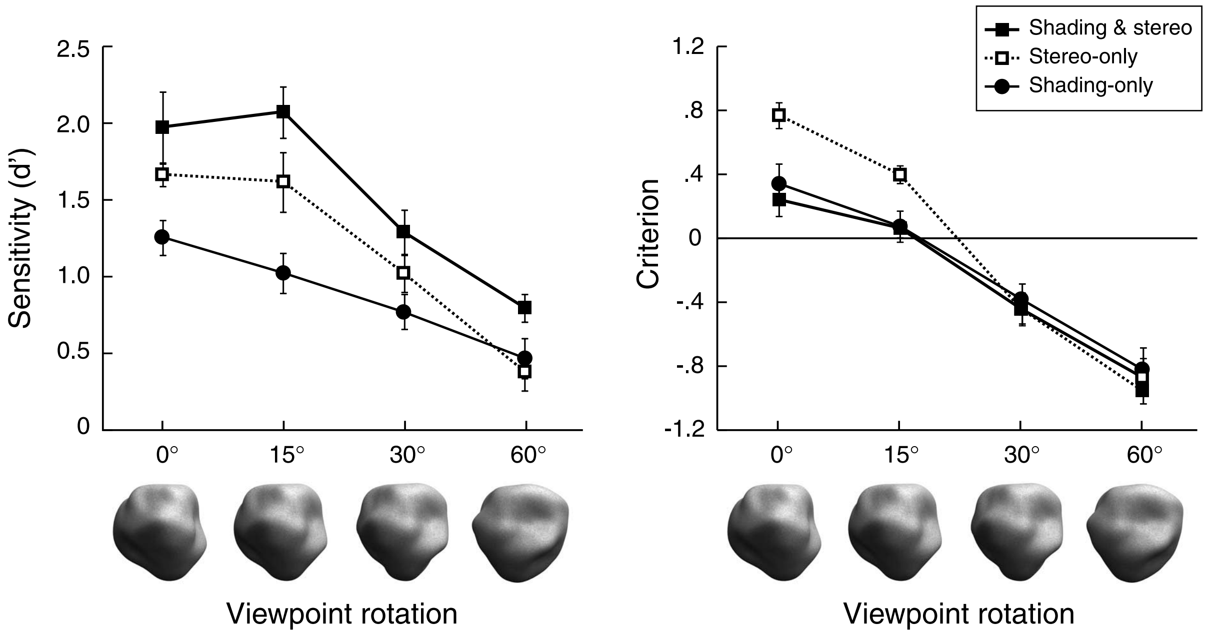

Figure 6. Results of Experiment 2. The graphs show (left) mean sensitivity and (right) criterion, averaged across observers, as a function of change in viewing orientation. The three lines on each graph correspond to the shape cue conditions: shading-only (circles), stereo-only (open squares), and combined shading and stereo (filled squares). Error bars depict ±1 standard error.

Positive and negative viewpoint rotations were combined for analysis, as before. Overall, performance was similar to the previous experiment. In all conditions, performance decreased with change in viewpoint, and stereo improved performance.

An ANOVA on *d′* found a significant main effect of viewing condition, $F(2,18) = 15.86$, $p < 0.001$, as well as a significant main effect of viewpoint change, $F(3,27) = 53.15$, $p < 0.001$, and an interaction, $F(6,54) = 2.83$, $p = 0.015$. Comparisons between the three viewing conditions revealed significant differences between all three pairs. Performance was significantly worse in the shading-only condition than the combined shading and stereo condition ($F(1,9) = 24.9$, $p < 0.001$), as well as worse than the stereo-only condition ($F(1,9) = 6.20$, $p = 0.003$). The combined shading and stereo condition yielded better performance than the stereo-only condition ($F(1,9) = 13.6$, $p = 0.004$).

As in the previous experiment, sensitivity was higher in the combined condition than the shading-only condition at most viewpoints (0°: $p = 0.027$, 15°: $p < 0.001$, 30°: $p = 0.008$), except for the 60° viewpoint rotation condition that showed a non-significant trend (60°: $p = 0.052$). Stereo appeared to provide an overall benefit, not just in cases with viewpoint rotation.

In Experiment 2, we also observed an overall difference in sensitivity between the stereo-only and combined conditions. Further comparisons found significant differences at 15° viewpoint ($p = 0.018$) and 60° viewpoint ($p = 0.047$). At the other viewpoints, the trend was similar but not significant (0°: $p = 0.18$, 30°: $p = 0.078$). Given that the effect of viewpoint appears continuous, we suspect that these null results are due to lack of power and that shading also provides an overall benefit to shape discrimination.

We modeled the effect of viewpoint as a linear effect and found that sensitivity significantly decreased with viewpoint change in all three shape cue conditions: shading-only ($F(1,9) = 38.4$, $p < 0.001$), stereo-only ($F(1,9) = 97.5$, $p < 0.001$), and combined ($F(1,9) = 55.9$, $p < 0.001$). There were significant interactions indicating that the viewpoint effect was smaller in the shading-only condition than either the stereo-only condition ($F(1,9) = 12.0$, $p = 0.002$) or the combined condition ($F(1,9) = 6.16$, $p = 0.02$). The linear effect of viewpoint on sensitivity was equivalent for the stereo-only and combined conditions ($F(1,9) = 0.03$, $p = 0.86$). These results are all consistent with Experiment 1.

The criterion results were again consistent with the previous experiment. An ANOVA on criterion found a significant main effect of viewpoint ($F(3,27) = 160.7$, $p < 0.001$) as well as an interaction ($F(6,54) = 8.44$, $p < 0.001$) but no main effect of shape cue condition ($F(2,18) = 2.87$, $p = 0.083$). In all conditions, observers tended to judge "same" more often when viewpoint was similar and "different" more often with larger change in viewpoint. This bias was observed in all three shape cue conditions, as evidenced by significant linear effects ($p < 0.001$). The response bias was larger in the stereo-only condition than either the shading-only condition ($F(1,9) = 32.7$, $p < 0.001$) or the combined cue condition ($F(1,9) = 24.5$, $p < 0.001$), which is also consistent with Experiment 1.

Although specularities potentially provide stronger shading information, we found that stereo still provided a benefit. By both measures of performance, d' and criterion, performance was better with stereo information than from shading alone, consistent with Experiment 1. We directly compared observed d' and criterion values from Experiments 1 and 2 in the shading-only condition and found no significant difference at any viewpoint rotation angle ($p > 0.05$). Thus, rendering objects with specularities did not measurably improve performance and did not eliminate a stereo benefit.

Experiment 3

The results of Experiment 2 for conditions with no viewpoint rotation appear to conflict with the findings of Norman et al. (2004), who observed no stereo benefit when specular highlights were present. However, the simulated surface reflectance used in Experiment 2 was not equivalent to the specular conditions of Norman et al. The specular objects used by Norman et al. had larger specular components, and the matte component had a chromatic hue that enhanced the subjective appearance of shininess.

Experiment 3 replicated the previous experiments using chromatic objects with larger specular components. An example is shown in Figure 7. Subjectively, these objects appear much more shiny than the objects in Experiment 2, and the larger specular component potentially provides richer monocular shape information.

Methods

Apparatus and procedure were identical to those in the previous experiments. Stimuli were also the same except for the simulated surface reflectance. We used a Phong model with no ambient component, 40% Lambertian component, and 60% specular component with exponent 100. The Lambertian component had an overall blue hue, and the texture had chromatic variation as well as lightness variation. The reflectance of the surface texture, expressed as RGB components, varied between (0.0, 0.4, 0.8) and (1.0, 1.0, 1.0). The specular component was achromatic.

Figure 7. Example of the shiny objects used in Experiment 3. The surface reflectance had a larger specular component, and the Lambertian component had an overall blue hue. The surface texture also had chromatic variation in addition to lightness variation.

Seventeen adults (six males and eleven females) aged from 19 to 34 at the University of Hong Kong participated in this experiment. All participants had normal or corrected-to-normal vision and passed a screening test for stereo acuity. Two subjects were excluded from analysis on the basis of poor performance in the stereo-only condition with no viewpoint change. These two subjects were essentially at chance in these conditions ($d' < 0.15$), indicating that they could not effectively use stereo information for 3D judgments. All participants were naive as to the purpose of the study and were paid for participating. The procedures were approved by and conform to the standards of the Human Research Ethics Committee for Non-Clinical Faculties.

Results and discussion

Figure 8 shows mean sensitivity (left) and criterion (right), averaged across fifteen observers, as a function of change in viewpoint for each of the three viewing conditions. Positive and negative viewpoint rotations were combined for analysis, as before. Performance in the binocular viewing conditions was worse overall than in the previous experiments, but otherwise the results were similar: Performance decreased with change in viewpoint, and stereo improved performance.

An ANOVA on d' found a significant main effect of viewing condition, $F(2,28) = 9.4$, $p = 0.001$, and a main effect of viewpoint change, $F(3,42) = 50.68$, $p < 0.001$, but no interaction ($F(6,84) = 1.29$, $p = 0.27$). Comparisons between the three viewing conditions revealed significant differences between the combined shading and stereo condition and two single-cue conditions. Performance was significantly better in the combined shading and stereo condition than the shading-only condition ($F(1,14) = 28.96$, $p < 0.001$) and the stereo-only condition ($F(1,14) = 11.18$, $p = 0.005$). There was no reliable difference between performance in the shading-only and stereo-only conditions ($F(1,14) = 0.006$, $p = 0.94$).

In Experiment 3, sensitivity was higher in the combined condition than the shading-only condition at most viewpoints (0°: $p = 0.035$, 15°: $p = 0.01$, 30°: $p = 0.002$), except for the 60° viewpoint rotation condition ($p = 0.39$). Stereo appeared to provide an overall benefit, not just in cases with viewpoint rotation. Sensitivity was also higher in the combined condition than the stereo-only condition at 0° viewpoint ($p = 0.029$) and 30° viewpoint ($p < 0.001$) but not at the 15° and 60° viewpoint rotation conditions. At the 15° viewpoint, however, the trend was similar but

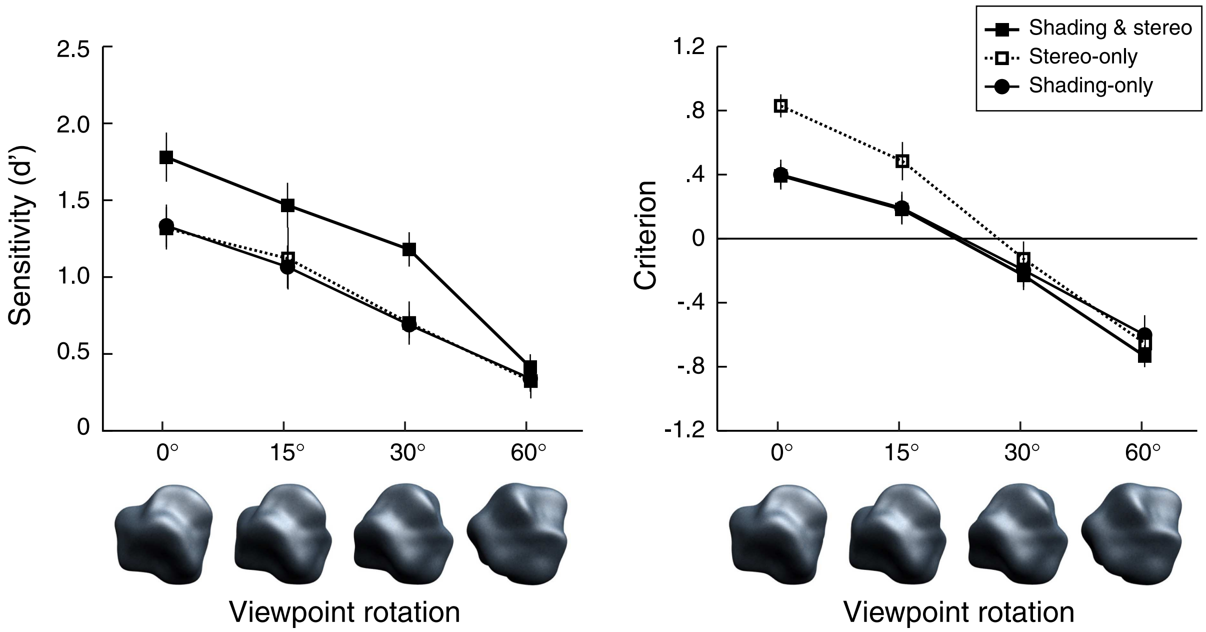

Figure 8. Results of Experiment 3. The graphs show (left) mean sensitivity and (right) criterion, averaged across observers, as a function of change in viewing orientation. The three lines on each graph correspond to the shape cue conditions: shading-only (circles), stereo-only (open squares), and combined shading and stereo (filled squares). Error bars depict ±1 standard error.

not significant ($p = 0.072$). Given the effect of viewpoint appears continuous except 60° viewpoint, which is consistent with previous experiments, we suspect that shading also provides an overall benefit to shape discrimination.

We modeled the effect of viewpoint as a linear effect and found that sensitivity was significantly reduced in all three shape cue conditions: shading-only ($F(1,14) = 47.82$, $p < 0.001$), stereo-only ($F(1,14) = 67.14$, $p < 0.001$), and combined ($F(1,14) = 83.59$, $p < 0.001$). There was, however, no significant interaction between any pairs of shape cue conditions (shading vs. combined: $F(1,15) = 2.56$, $p = 0.11$; shading vs. stereo, $F(1,15) = 0.005$, $p = 0.95$; stereo vs. combined, $F(1,15) = 2.97$, $p = 0.09$). In the previous experiments, the viewpoint effect was larger for the combined condition than the shading-only condition. In Experiment 3, this interaction was not significant, though the trend was in the same direction.

The criterion results were consistent with the previous experiments. An ANOVA on criterion found a significant main effect of shape cue condition ($F(2,28) = 5.91$, $p = 0.007$), as well as a significant main effect of viewpoint changes ($F(3,42) = 191.58$, $p < 0.001$) and an interaction ($F(6,84) = 6.53$, $p < 0.001$). In all conditions, observers tended to judge "same" more often when viewpoint was similar and "different" more often with larger change in viewpoint. This bias was observed in all three shape cue conditions, as evidenced by significant linear effects ($p < 0.001$). The response bias was larger in the stereo-only condition than either the shading-only condition ($F(1,14) = 31.14$, $p < 0.001$) or the combined cue condition ($F(1,14) = 28.28$, $p < 0.001$) but no reliable difference between the shading-only and combined cue conditions ($F(1,14) = 1.92$, $p = 0.17$), which is also consistent with previous experiments.

Overall performance for the stereo-only and combined conditions in Experiment 3 appeared lower than in the previous experiments. To test this, we computed mean sensitivity averaged across viewpoints as a measure of overall performance and compared performance in Experiment 3 to the aggregate data from Experiments 1 and 2. Mean sensitivity was significantly lower in Experiment 3 than in the previous experiments for both the stereo-only condition ($p = 0.021$) and the combined stereo and shading condition ($p = 0.021$) but not the shading-only condition ($p = 0.56$). We suspect that the poorer overall performance in the binocular conditions of Experiment 3 was due to individual differences in ability to use stereo information. The stereo-only condition was identical across the three experiments, and viewing conditions were blocked, yet overall sensitivity was lower in Experiment 3. Given that stimuli were identical, the difference in this condition is most likely due to the particular sample of observers. The difference in performance in the combined condition was the same as the difference in the stereo-only condition, so it could also be explained by differences in ability to use stereo information.

With regard to the contribution of stereo information, Experiment 3 replicated the main findings of the previous experiments: Binocular viewing improved shape discrimination, both for conditions with and without change in viewpoint. Although the shinier objects, in principle, could have provided stronger monocular information about 3D shape, we still observed a consistent benefit from stereo.

The only qualitative discrepancy compared to the previous experiments was that we did not observe a significant stereo benefit for viewpoint rotation of 60°. Given that performance was lower overall in Experiment 3, this discrepancy may have been due to lack of sensitivity.

General discussion

Benefit from stereo information

One goal was to test whether stereo information improves 3D shape discrimination across changes in viewpoint. We used stimuli that provided rich monocular shape cues: shading, specularities, texture, and occlusion contour. These monocular cues were found by Norman et al. (2004) to be sufficient for discrimination of the types of random solid objects tested here. However, we observed a consistent benefit from stereo viewing, across all changes in viewpoint.

Other studies have similarly found a significant cost of changing in viewpoint and a benefit from stereo (Bennett & Vuong, 2006; Burke, 2005; Edelman & Bülthoff, 1992). Our results are consistent with these studies and demonstrate that a stereo benefit can also be observed in conditions with richer cues. Our results are also generally consistent with those of Chan et al. (2006). Chan et al. observed a benefit from stereo for shape discrimination across large (90°) changes in viewpoint. However, the stereo advantage observed by Chan et al. was comparatively small. The difference could be due to the type of objects. The stimuli used in Chan et al. were polyhedra, which are more structured than the smooth random shapes tested here.

We hypothesized that stereo information might be especially beneficial with larger change in viewpoint. Some previous results suggest that stereo provides a benefit for face recognition only in the case of large viewpoint changes (Burke et al., 2007; Liu et al., 2006; see Introduction section). Based on these results, one might have expected an interaction in our experiment, with viewpoint costs being smaller in the condition with combined shading and stereo information than in the monocular shading-only condition. Our results are contrary to this hypothesis. We observed that the improvement in binocular conditions was constant or decreased with change in viewpoint. This may be due to a difference in tasks. Our conditions required observers to discriminate smoothly curved 3D surfaces, while face recognition involves comparison between configurations of salient features. Another possible reason would be because we controlled information provided by boundary contours. Boundary contours might provide more beneficial information across small changes than large changes in viewpoint. Since there was additional information from boundary contours in our study, stereo benefit for small viewpoint changes was not different from that for large viewpoint changes.

We found that stereo improved shape discrimination even in the case where there was no change in viewing angle. In this case, the challenge for perceiving shape constancy comes from variation in simulated lighting direction across first and second presentations. Although our task required discrimination between qualitatively different shapes (for example, see Figure 3), change in lighting made the task challenging. Our results indicate that stereo information facilitated discriminating shapes across lighting changes. This is consistent with the results of Liu et al. (2000), who found a stereo benefit in discrimination of faces that were illuminated from different lighting directions.

Our results for the condition without viewpoint change appear to conflict with previous results of Norman et al. (2004). They found that when shapes were defined by matte shading, specularities, and occlusion contours, shape discrimination from a monocular view was as good as when stereo information was present. Our stimuli in Experiment 3 had similar shading and specularities, but we observed a consistent stereo advantage. Norman et al. point out that the lack of stereo benefit in their study could be a ceiling effect. Another potential factor is the presence of surface texture. In Norman et al., the conditions with smooth shading and specularities did not include surface texture, which could have reduced the effectiveness of stereo information. Specularities pose a challenge for binocular correspondence, because the location of specular highlights on an object differs for the right and left eyes' views. In our experiment, objects with shading also had surface texture, which would facilitate binocular correspondence. Todd et al. (1997) found that observers judged local orientation slightly more accurately and reliably for the textured surface than for the smoothly shaded shiny surfaces when viewed binocularly.

A possible explanation for the overall stereo benefit is that stereo resolves ambiguities in monocular shape information. Although shading can produce a vivid subjective percept of shape, shading information by itself is formally ambiguous (Belhumeur, Kriegman, & Yuille, 1999). Studies of perceived shape from shading have found that 3D surfaces tend to be perceived as distorted relative to veridical (Battu, Kappers, & Koenderink, 2007; Di Luca, Domini, & Caudek, 2010; Koenderink, van Doorn, Kappers, & Todd, 2001; Nefs, Koenderink, & Kappers, 2005, 2006). In our experiment, uniform compression or expansion of depth would not necessarily interfere with shape discrimination, because comparison objects had qualitative differences that would remain distinct. However, perceptual distortions that change relative surface relief could be more problematic. Nefs et al. (2005, 2006) found that changing the direction of illumination produced systematic distortions of perceived surface relief, both for Lambertian and specular surfaces.

Such illumination-dependent distortions might have impaired shape discrimination in our experiment, particularly in the monocular conditions. Additional information from stereo information could compensate for such distortions, thereby improving performance. This would be consistent with other evidence suggesting that stereo helps resolve the ambiguity of shading information (Di Luca et al., 2010; Norman, Todd, & Phillips, 1995; O'shea, Agrawala, & Banks, 2010; Vuong et al., 2006).

View dependence

Although shading provided a strong subjective percept of 3D shape, shape constancy in the shading-only conditions was poor. A small rotation of viewpoint (15°) was sufficient to significantly impair performance. Sensitivity was modest even with no change in viewpoint ($d' \approx 1.4$), suggesting that variation in light direction also imposed a cost. Results were similar whether the reflectance model for the object was Lambertian (Experiment 1) or had a specular component (Experiments 2 and 3). Additional shape information from stereo improved performance relative to the shading-only condition, but shape constancy remained poor even in the full cue conditions. This is consistent with some previous studies (Bennett & Vuong, 2006; Burke, 2005; Edelman & Bülthoff, 1992).

Norman et al. have tested shape discrimination across changes in viewpoint for random 3D objects like those tested here and also observed a large effect of viewpoint change (Norman et al., 2008, 2009). The stimuli tested by Norman et al. (2009) were comparable to those in our combined shading and stereo condition. They found a significant decrease in sensitivity and increase in negative response bias with change in viewpoint, as observed here, but better performance overall and less viewpoint cost than in our experiment. This quantitative difference could be explained by the fact that we controlled for information provided by the occlusion boundary contours, while Norman et al. did not. In our study, comparison objects had the same boundary contours at their base viewpoint, which made the task challenging even with small changes in viewpoint. In Norman et al., the occlusion boundary was potentially informative.

The poor shape constancy observed here may be a function of the class of objects that we tested. Some researchers have argued that structural constraints like symmetry or planarity are important for view-invariant shape perception from a monocular image. Pizlo and Stevenson (1999) found that shape constancy from novel views was more reliable for structured polyhedrons with symmetry and planar faces than for unstructured polyhedrons. While our stimuli produced a strong subjective percept of 3D shape, the objects were random and comparatively unstructured. If viewpoint-invariant perception of 3D shape depends strongly on structural constraints, then objects with symmetry or other constraints might show greater shape constancy.

Conclusion

We found that stereo information improved 3D shape discrimination even when rich monocular information is available. In contrast to some previous studies, we observed a stereo advantage both with and without change in viewpoint. However, even with stereo, observers showed limited ability to generalize across changes in viewpoint.

Our results suggest that perception of 3D shape is highly view-dependent even when rich 3D information is available. Shape discrimination was significantly impaired by change in viewing orientation, for stimuli and viewing conditions that produce a vivid subjective percept of 3D shape.

Acknowledgments

We would like to thank two anonymous reviewers for their comments. This work was supported by a grant from the Hong Kong Research Grants Council (GRF HKU-752010H). Part of this work was presented at the 2011 Meeting of Vision Sciences Society.

Commercial relationships: none.
Corresponding author: Young Lim Lee.
Email: younglee@hku.hk.
Address: Department of Psychology, University of Hong Kong, Hong Kong.

References

Battu, B., Kappers, A. M. L., & Koenderink, J. J. (2007). Ambiguity in pictorial depth. *Perception, 36,* 1290–1304.

Belhumeur, P. N., Kriegman, D., & Yuille, A. L. (1999). The bas-relief ambiguity. *International Journal of Computer Vision, 35,* 33–44.

Bennett, D. J., & Vuong, Q. C. (2006). A stereo advantage in generalizing over changes in viewpoint on object recognition tasks. *Perception & Psychophysics, 68,* 1082–1093.

Bülthoff, H. H., & Mallot, H. A. (1988). Integration of depth modules: Stereo and shading. *Journal of the Optical Society of America A, 5,* 1749–1758.

Burke, D. (2005). Combining disparate views of objects: Viewpoint costs are reduced by stereopsis. *Visual Cognition, 12,* 705–719.

Burke, D., Taubert, J., & Higman, T. (2007). Are face representations viewpoint dependent? A stereo advantage for generalizing across different views of faces. *Vision Research, 47,* 2164–2169.

Chan, M. W., Stevenson, A. K., Li, Y., & Pizlo, Z. (2006). Binocular shape constancy from novel views: The role of a prior constraints. *Perception & Psychophysics, 68,* 1124–1139.

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

Edelman, S., & Bülthoff, H. H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision Research, 32,* 2385–2400.

Koenderink, J. J., van Doorn, A. J., Kappers, A. M. L., & Todd, J. T. (2001). Ambiguity and the 'mental eye' in pictorial relief. *Perception, 30,* 431–448.

Liu, C. H., Collin, C. A., & Chaudhuri, A. (2000). Does face recognition rely on encoding of 3-D surface? Examining the role of shape-from-shading and shape-from-stereo. *Perception, 29,* 729–743.

Liu, C. H., & Kersten, D. (2003). Three-dimensional symmetric shapes are discriminated more efficiently than asymmetric ones. *Journal of the Optical Society of America A, 20,* 1331–1339.

Liu, C. H., Ward, J., & Young, A. W. (2006). Transfer between two- and three-dimensional representations of faces. *Visual Cognition, 13,* 51–64.

Nefs, H. T., Koenderink, J. J., & Kappers, A. M. L. (2005). The influence of illumination direction on the pictorial reliefs of Lambertian surfaces. *Perception, 34,* 275–287.

Nefs, H. T., Koenderink, J. J., & Kappers, A. M. L. (2006). Shape-from-shading for matte and glossy objects. *Acta Psychologica, 121,* 297–316.

Norman, J. F., Bartholomew, A. N., & Burton, C. L. (2008). Aging preserves the ability to perceive 3D object shape from static but not deforming boundary contours. *Acta Psychologica, 129,* 198–207.

Norman, J. F., Swindle, J. M., Jennings, R., Mullins, E. M., & Beers, A. M. (2009). Stereoscopic shape discrimination is well preserved across changes in object size. *Acta Psychologica, 131,* 129–135.

Norman, J. F., Todd, J. T., & Orban, G. A. (2004). Perception of three-dimensional shape from specular highlights, deformations of shading, and other types of visual information. *Psychological Science, 15,* 565–570.

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

O'Shea, J. P., Agrawala, M., & Banks, M. S. (2010). The influence of shape cues on the perception of lighting direction. *Journal of Vision, 10*(12):21, 1–21, http://www.journalofvision.org/content/10/12/21, doi:10.1167/10.12.21. [PubMed] [Article]

Pasqualotto, A., & Hayward, W. G. (2009). A stereo disadvantage for recognizing rotated familiar objects. *Psychonomic Bulletin & Review, 16,* 832–838.

Pizlo, Z., Li, Y., & Steinman, R. M. (2008). Binocular disparity only comes into play when everything else fails; a finding with broader implications than one might suppose. *Spatial Vision, 21,* 495–508.

Pizlo, Z., & Stevenson, A. K. (1999). Shape constancy from novel view. *Perception & Psychophysics, 61,* 1299–1307.

Todd, J. T., Norman, J. F., Koenderink, J. J., & Kappers, A. M. L. (1997). Effects of texture, illumination and surface reflectance on stereoscopic shape perception. *Perception, 26,* 807–822.

Vuong, Q. C., Domini, F., & Caudek, C. (2006). Disparity and shading cues cooperate for surface interpolation, *Perception, 35,* 145–155.