Perception, 2004, volume 33, pages 1305 – 1318

DOI:10.1068/p5288

# Perception of scene layout from optical contact, shadows, and motion

Rui Ni, Myron L Braunstein
Department of Cognitive Sciences, University of California at Irvine, 3151 Social Science Plaza, Irvine, CA 92697-5100, USA; e-mail: mlbrauns@uci.edu

George J Andersen
Department of Psychology, University of California at Riverside, Riverside, CA 92521, USA
Received 23 December 2003, in revised form 24 March 2004; published online on 8 December 2004

**Abstract.** Kersten et al (1997 *Perception* **26** 171 – 192) found that the perceived motion of an object in a 3-D scene was determined by the motion of a shadow. In the present study, we compared the effect of a shadow to that of a second object on the ground in determining the perceived position in depth of a floating object in both dynamic and stationary scenes. Changing the second (lower) object from textured to dark increased the influence of the second object on the judged position of the first object. Giving the second object zero thickness had this effect only if it was also dark. Variations in the height of the floating object were important with a second object but not with a shadow, in motion scenes. With alternative shadows present, the position of the floating object was determined primarily by matching speeds, with matching sizes as a secondary factor. These results show some similarities but important differences between the effect of a second object and that of a shadow.

## 1 Introduction

Gibson (1950) described an important source of information available in the projection of a 3-D scene for the perception of the location of objects in that scene, which he referred to as optical contact. Optical contact describes the contact in a 2-D projection of a 3-D scene between the image of an object and the image of a background surface. Gibson demonstrated that optical contact can provide false information about the relative positions of objects if the objects are not in physical contact with the background surface, for example if one object is floating above the ground.

Observers are able to judge the position of an object in a scene that is not resting on the ground if there are intermediate objects between the one being judged and the ground surface. Meng and Sedgwick (2001, 2002) examined the case of an object resting on a platform which was in turn resting on the ground or floating above the ground. They found that observers used the optical-contact position of the object relative to the platform together with the optical-contact position of the platform relative to the ground to judge the position of the object in the scene. They referred to this as a "nested contact relation". Ni et al (in press) found that the judged position of a higher object was influenced by the position of a lower object (see figure 1) even in the absence of direct optical or physical contact. This effect was weak for stationary objects but increased for objects that were rigidly moving with the scene. Ni et al attributed this to the grouping of the objects by common motion, so that the two objects appeared to be stacked one on top of the other, rather than each lying on the ground surface.

Both Meng and Sedgwick (2001, experiments 1 – 3) and Ni et al (in press) deliberately excluded shadows from their scenes. This is because they were interested in studying the effects of other variables, and it had already been established (eg Yonas et al 1978; Kersten et al 1996, 1997; Mamassian et al 1998) that shadows play an important role in determining the positions of objects in a scene. Now that the effects of nested contact and common motion have been studied in isolation, it is important to determine how

**Figure 1.** A scene with two cylinders in the foreground.

information from shadows is integrated with these other sources of information in determining the perceived positions of objects in a scene. This is the purpose of the present study.

Shadow not only affects the perception of 3-D shape (eg Ramachandran 1988), but also affects object recognition (Braje et al 1998, 2000; Castiello 2001). Castiello (2001) asked observers to recognize familiar objects, such as an apple, a banana, a bottle, etc, while changing the presence, location, and shape of cast and attached shadows. He found that shadows increased response time if the cast shadows and attached shadows were made incongruent by deriving them from different light sources. These results indicated that correct shadow information favored the perception of an object in 3-D space, but this perception could be impaired by incorrect or conflicting shadow information. In studies of the relation of shape to shadows, Knill et al (1997) applied a systematic analysis of local geometric structure of shadows on continuous surfaces. Their results showed that intrinsic shadows could be informative about the properties of 3-D objects, such as illumination direction and surface structure.

In a developmental study of the perception of shape and distance from shadows, Yonas et al (1978) found that young children were able to judge the shapes of objects from cast shadows. Children aged 3 and 4 years were shown picture cards with an ellipse on the wall and an ellipse on the ground plane, each of which represented a circle. When shadow information was provided, shape judgments were improved and perceived distance was affected by the location of shadow. Comparing the effectiveness of shadows to other cues, Wanger et al (1992) found that shadows have a dominant role in visual perception of computer-generated images relative to perspective, texture, reference frames, and motion. Shadows play an important role in determining perceived distance not only in monocular vision but also in binocular vision. Puerta (1989) found that stereo views of shadows in isolation could reveal the depth order in a scene.

The clearest demonstrations of the importance of shadow in the perception of dynamic 3-D displays are found in research by Kersten and his colleagues. Kersten et al (1996) introduced a dramatic phenomenon called "illusory motion from shadow", in which the displacement of the shadow relative to a stationary square induced a strong perception of the motion of the square in depth. Kersten et al (1997) employed a "ball in a box" display, in which observers judged the height of a ball moving diagonally across the bottom of a box. Variations in the motion trajectory of the object's cast shadow determined the perceived motion of the object, even when changes in the projected size of the object provided conflicting information.

Madison et al (2001) had observers judge whether an object was in contact with the ground plane as the presence of a shadow, an inter-reflection, and an extra light source was systematically manipulated. Performance in judging object contact was at chance without shadows and inter-reflection. The most accurate judgments were found with a combination of a correct shadow and inter-reflection.

In our previous research (Ni et al, in press), we found that the judged position of a floating object in a scene was affected both by its optical-contact position and the optical-contact position of a lower object. For stationary scenes, the influence of the lower object was relatively small, but increased for scenes that were translating horizontally. Even in this case, the perceived position of the top object, with two objects present, did not reach the optical-contact position of the lower object. Kersten's research (1996, 1997) suggests that if the lower object was replaced by a shadow, its optical-contact position should fully determine the perceived position of the upper object, especially in a dynamic scene.

In the present study we compared the effect of a second object on the ground to that of a shadow, in determining the perceived location of a floating object in a 3-D scene. A second object can influence the judged position of a floating object if it is grouped with that object, with the greatest influence occurring when the objects are perceived as stacked vertically at one location rather than each lying on the ground at separate locations. This grouping should depend on such factors as proximity and common motion. A floating object and a shadow would also have to be associated with one another, but the constraints are somewhat different. A stationary-light-source constraint (Kersten et al 1997) would require common motion between the object and the shadow. A common-motion constraint for multiple objects is based on an assumption of rigidity (Ullman 1979) or quasi-rigidity. A violation of the common-motion constraint for two objects would mean that the objects would be perceived as moving independently. A violation of this constraint for an object and a cast shadow would mean that the light source was moving. If observers accept a perception of nonrigid motion among objects more readily than a perception of a moving light source, then we would expect the influence of a shadow on the judged position of a floating object to be greater than that of a second object.

The present experiments addressed three specific issues: (i) In a 3-D scene with a floating object, what features determine whether a second object is treated perceptually as another object or as a shadow? (ii) Does the separation of two objects affect responses differently depending on whether the second object has the appearance of an object or of a shadow? (iii) If there is more than one potential shadow, what features determine which shadow is used in judging the location of a floating object? In order to determine the specific differences between a second object and a shadow that influence object location judgments, we varied characteristics of the object along two dimensions in the first experiment: (a) The vertical dimension (thickness) of the second object was varied between zero and the thickness of the floating object. (b) The second object was either a textured object or a dark, shaded object. We expected the second object to have the greatest influence on the judged position of the floating object when its appearance was fully compatible with that of a shadow, that is when it had zero thickness and was dark rather than textured.

Proximity should be less important in associating an object with a shadow than in associating an object with a second object in a scene, as long as there is common motion between the object and shadow. This hypothesis was considered in the second experiment by varying the height of the floating object in both dynamic and stationary scenes. Increasing the height of the floating object was expected to reduce the influence of a second object on the perceived position of the floating object, but not to reduce the influence of a shadow, at least when motion was present.

In the third experiment, we examined the relative importance of a common-motion constraint, based on an assumption of a stationary light source, and a constraint of matched projected size, based on an assumption of a distant overhead light source. Three alternative shadows were presented in each display, with one of the three matching the floating object either in motion, size, or both. On the basis of previous research (eg Kersten et al 1997) it was expected that the judged distance of the floating object would be influenced primarily by the location of the shadow that displayed the same motion. Size was expected to have a secondary influence.

Although most research on the perception of 3-D objects in 3-D scenes has used either computer-generated scenes or directly observed real-world scenes, both of these methods have limitations (Sauer et al 2001). Simulated scenes provide good control over stimulus characteristics but may lack the abundance of information sources available in natural environments. Real-world scenes provide an abundance of information but it is difficult to control all attributes of objects in such scenes. In the present research, we balanced these two methods by adding simulated objects (textured cylinders, shadows, and a track with a marker) to a movie of a real scene.

## 2 Experiment 1: Thickness and texture of the bottom object

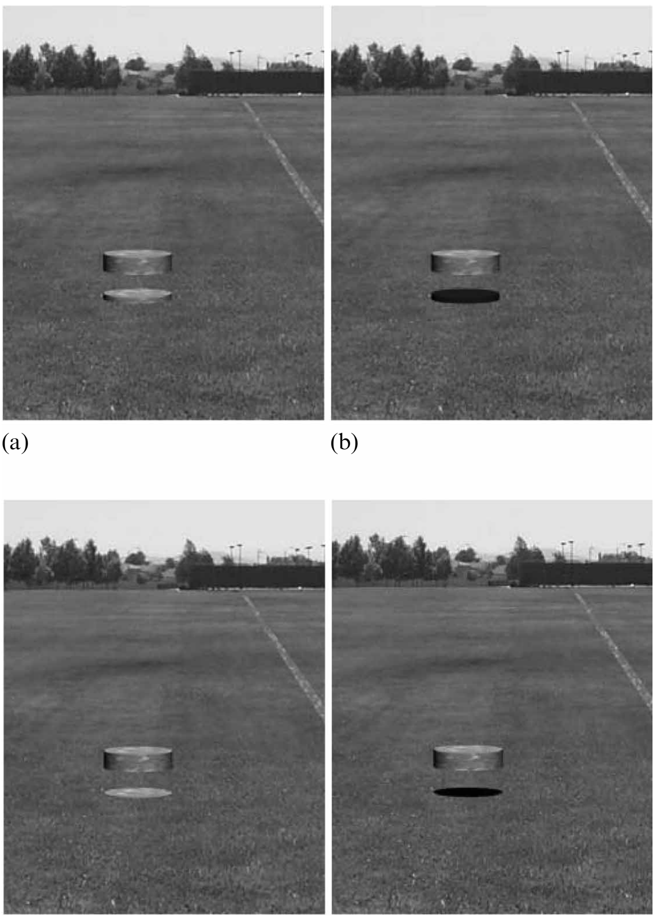
### 2.1 *Method*

2.1.1 *Stimuli.* Each display consisted of two computer-generated cylinders and a computer-generated track with a red marker, superimposed on a background movie of an actual 3-D scene. The background movie consisted of 50 photographs taken with a Kodak DC260 digital camera with a 35 mm equivalent lens. The photographs were taken at fixed positions along a horizontal track that were spaced according to a sinusoidal function, with the larger separations in the center of the track and the smaller separations at the ends. This was intended to approximate the motion of an observer moving his or her head back and forth horizontally. The photographs were cropped from the original size of $1536 \times 1024$ pixels to $1024 \times 768$ pixels, with the original pixel aspect ratio kept unchanged. An example of a cropped photograph with 2 cylinders, a track, and a marker superimposed is shown in figure 1.

The top object was always a textured cylinder and the bottom object was either a textured cylinder or a dark, shaded cylinder. The simulated diameters of both cylinders were 36 cm. (In this experiment, and in experiment 2, the projected sizes of the cylinders indicated that they were at the same distance. Effects of variations in projected size were examined in experiment 3.) The thickness (vertical extent) of the top cylinder was 9 cm. The bottom cylinder varied in thickness between zero and the thickness of the top cylinder. Lambertian shading was applied to the top cylinder and to the bottom cylinder when its thickness was greater than zero. We will refer to the two types of cylinders as textured and shaded. The zero-thickness cylinder is, of course, a circle (an ellipse in the image). The dark cylinder with zero thickness had the appearance of a shadow. (For convenience in describing the stimuli, we will sometimes refer to the circular patches—the cylinders with zero thickness—as objects, whether they are textured or shaded. When a distinction is necessary, we will use the terms "second object" and "shadow".) In the 3-D simulation, the front edge of the bottom cylinder was 21.2 m from the observer and the top cylinder was aligned vertically with the bottom cylinder.

With the height of the top cylinder constant, the variations in the thickness of the bottom cylinder resulted in variations in the separation between the two cylinders. For this reason, a control experiment was conducted in which the separation between the cylinders was kept constant but the height of the top cylinder was allowed to vary.

Figures 2a and 2b show frames from displays with the bottom cylinder textured and the bottom cylinder shaded. The bottom cylinder in both cases has a thickness of 1/3 that of the top cylinder. Figures 2c and 2d show similar conditions with the bottom cylinder having zero thickness. In all conditions, the top cylinder had the same motion as the bottom cylinder.



(a) (b)

(c) (d)

**Figure 2.** Pairs of objects with the thickness of the bottom object 1/3 that of the top [(a) and (b)] or zero [(c) and (d)]. Note that these and subsequent scene illustrations are cropped from the full scene, shown in figure 1, that was actually displayed.

2.1.2 *Observers.* The observers were twenty-four students from the University of California at Irvine. All observers had normal or corrected-to-normal visual acuity and all were naïve about the purpose of the experiment. They received extra credit for participating.

2.1.3 *Design.* The independent variables were whether the scene was in motion or stationary, whether the bottom cylinder was textured or shaded, and the thickness of the bottom cylinder (0/3, 1/3, 2/3, or 3/3 that of the top cylinder). The twenty-four observers were divided into four groups and the first two independent variables were run as between-subjects variables, to avoid any possible influence of viewing displays in one combination of motion and shading conditions on judgments in another combination of conditions. The thickness variable was run within subjects. Each observer also responded to a control condition in which the gap between the two cylinders was held constant, but the height of the top cylinder was allowed to vary. The constant-height and constant-gap conditions were run in separate blocks for each group. The order of the blocks was counterbalanced across observers. Each block contained six replications of each of the four thickness levels. The 24 trials in each block were randomly arranged for each observer. A practice block, with the same conditions as the first block but in a different random order, preceded the first block.

2.1.4 *Apparatus.* A Pentium 4 2-GHz computer displayed the stimuli on a 21-inch (53-cm) flat-screen CRT monitor at a resolution of 1024 (horizontal) $\times$ 768 (vertical) and a refresh rate of 85 Hz. The experiment was carried out in a darkened room. Observers viewed the displays binocularly from a distance of 85 cm through a 19-cm-diameter collimating lens with a focal length of 75 cm, with their heads stabilized by a chin-rest and head-rest. A black viewing hood was placed between the collimating lens and the monitor, limiting the field of view to the display area, and black cloth separated the observer from the apparatus to ensure that the observer would not see the location of the monitor. Responses were made with a Microsoft SideWinder joystick.

2.1.5 *Procedure.* The observers' task was to adjust the red marker on the track on the right side of the scene, as shown in figure 1, until it matched the distance of the front surface of the cylinder. Observers pressed the trigger button on the joystick when they were satisfied with their response.

### 2.2 *Results and discussion*

Judged cylinder position as a function of bottom-cylinder thickness is shown in figure 3a. The observed data are the mean judged distances of the front edge of the top object for the six observers, for the conditions with two textured cylinders and with one textured and one dark, shaded cylinder. The graphs also show the optical-contact positions in the image for the top and bottom objects. These are the distances from the observer at which the bottom front edge of the top object and the bottom front edge of the bottom object contact the ground in the image. Because the top object is always positioned directly above the bottom object in the simulation, the optical-contact position of the bottom object is always the same as the simulated position of both objects. In the motion conditions, the means of the observed judgments fell between the optical-contact positions of the top and bottom objects. The judgments for the conditions with one textured object and one dark object were closer to the optical-contact position of the bottom object than were the judgments for the conditions with two textured cylinders, suggesting a greater influence of the bottom object in the former case. The mean for the zero-thickness dark object, intended to simulate a shadow, is closest to the optical-contact position of the bottom object. The results are similar, but displaced upward, for the stationary conditions.

A three-way ANOVA with two between-subjects variables (motion versus stationary and textured versus shaded) and one within-subjects variable (thickness) showed significant main effects for motion versus stationary ($F_{1,20} = 13.26$, $p < 0.01$) and textured versus shaded ($F_{1,20} = 4.63$, $p < 0.05$). The main effect of thickness was not significant ($F_{3,60} = 2.12$, $p > 0.05$), but there was a significant interaction between textured versus

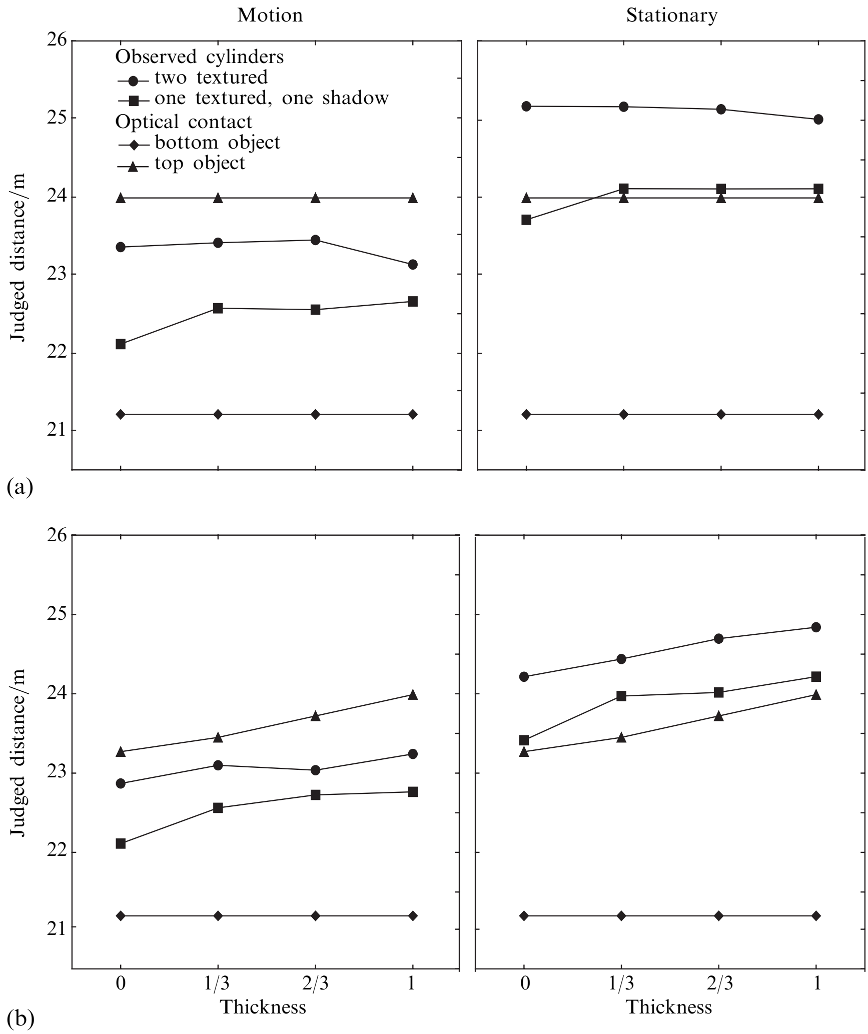



**Figure 3.** Judged distance as a function of the thickness of the lower object in experiment 1. Panel (a) shows the results with the top object at a constant height. Panel (b) shows the results of a control experiment with the gap between the objects constant.

shaded and thickness ($F_{3,60} = 3.56$, $p < 0.05$). As shown in figure 3a, there was a greater change in judged distance between the zero-thickness condition and the other thickness conditions for the shaded objects than for the textured objects. There were no other significant interactions.

2.2.1 *Control condition.* The results for the constant-gap control condition are shown in figure 3b. The optical-contact position of the top object changed with the thickness of the lower object in this experiment, to maintain a constant vertical separation between the objects. As would be expected, judged distance also increased with the thickness of the lower object. In all other respects, these results are quite similar to those in the main experiment. Observed judgments for the conditions with two textured cylinders were closer to the optical-contact location of the top cylinder than were the observed judgments for the conditions with one textured cylinder and one dark shaded cylinder. Again, judgments for the shaded cylinder with zero thickness were closest to the optical-contact position of the lower object.

A three-way ANOVA showed a significant main effect for motion versus stationary ($F_{1,20} = 12.02$, $p < 0.01$). The main effect for textured versus shaded was not significant ($F_{1,20} = 3.07$, $p > 0.05$). The main effect of thickness was significant ($F_{3,60} = 15.94$, $p < 0.01$). This would be expected because the height of the top cylinder, and therefore its optical-contact position, varied with thickness in the constant-gap control displays. There were no significant interactions. Although the main effect of textured versus shaded was not significant in the ANOVA, figure 3b shows that judged distance was greater for the textured stimuli than for the shaded stimuli for all eight combinations of motion versus stationary and thickness.

Overall, judged distances were affected by variations in whether the lower object was a textured or a dark, shaded object, regardless of its thickness. Variations in thickness of the lower object, on the other hand, had little effect on the judged distance of the floating object, with one exception: for the dark lower object there was a drop in the perceived distance of the floating object when the thickness of the lower object was reduced to zero, that is when both shading and thickness indicated that the lower object was a shadow.

## 3 Experiment 2: Variation in gap size

In experiment 1, small variations in the gap between the cylinders were introduced in the main experiment as a result of the variations in the thickness of the bottom cylinder, with the top cylinder kept at a constant height. These variations had little effect on the judgments. The purpose of experiment 2 was to determine whether larger variations in gap size would reduce the tendency to group two textured cylinders in a scene, but would not affect the grouping of one cylinder with a shadow.

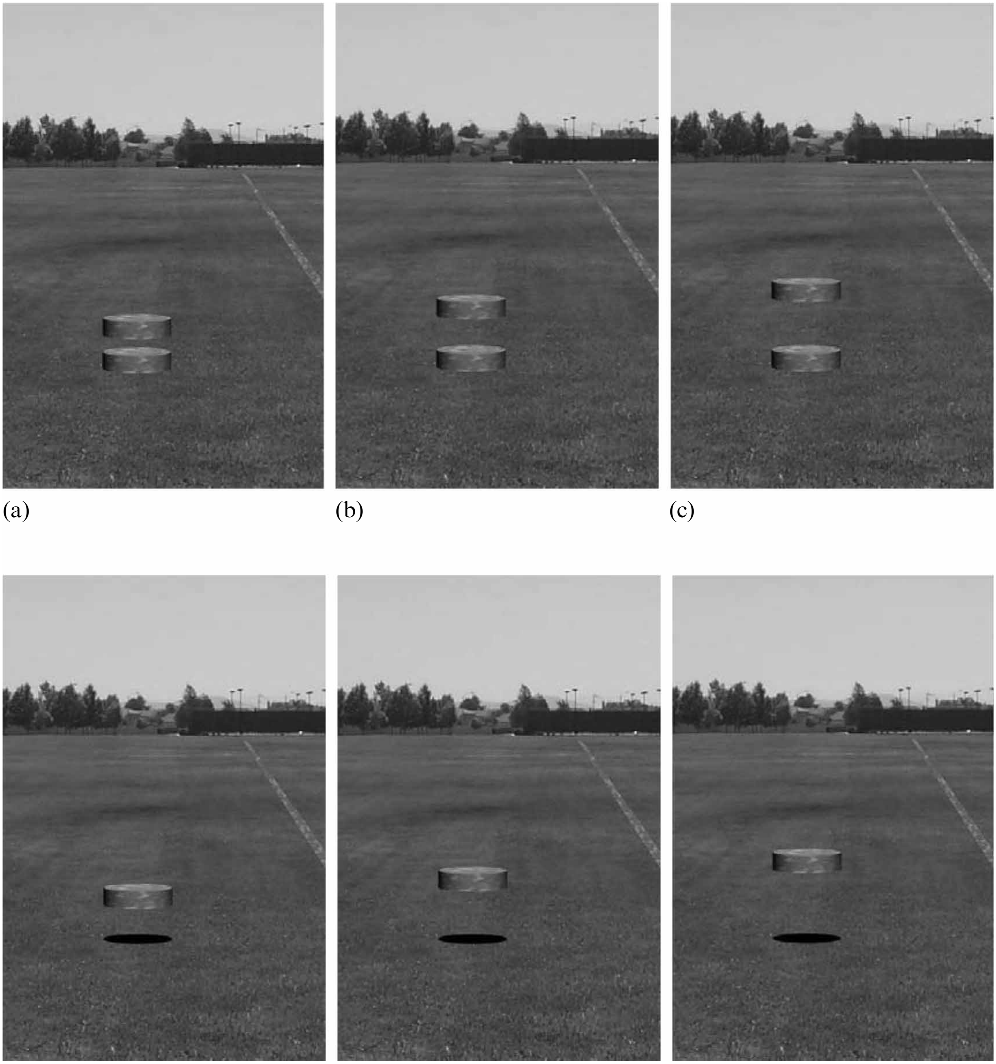
### 3.1 *Method*

3.1.1 *Observers.* The observers were twenty-four students from the University of California at Irvine. All observers had normal or corrected-to-normal visual acuity and were naïve about the purpose of the experiment. They received extra credit in a psychology course for participating.

3.1.2 *Stimuli.* The background scene was the same as in experiment 1. The top cylinder was the same projected height and diameter as the top cylinder in experiment 1. The bottom object was either an identical cylinder or a shadow. The top cylinder was either 18 cm, 27 cm, or 36 cm above the ground, or, equivalently, above the shadow. The gap between the top cylinder and the bottom cylinder was 9 cm, 18 cm, or 27 cm. This set of values allows us to compare either conditions in which the distance from the top cylinder to the ground is equal or conditions in which the gap size is equal. Figures 4a – 4c show the three cylinder-height conditions for two cylinders and figures 4d – 4f show the three cylinder-height conditions for a cylinder and shadow. In the motion conditions, the top cylinder moved rigidly with the bottom cylinder or the shadow.

3.1.3 *Design.* The independent variables were whether the scene was in motion or stationary, whether the bottom object was a cylinder or a shadow, and the height of the top cylinder above the ground (18 cm, 27 cm, or 36 cm). The first two variables were run between subjects (to avoid any possible influence of viewing displays of one type on responses to displays of another type) with six observers in each condition; the third variable was run within subjects. Each observer responded to 6 replications of each gap condition. The 18 trials were presented in one block, in a different random order for each observer, preceded by a practice block consisting of the same trials in a different random order.

3.1.4 *Apparatus and procedure.* The apparatus and procedure were the same as in experiment 1.



(a) (b) (c) (d) (e) (f)

**Figure 4.** Examples of displays with two cylinders [(a) – (c)] and one cylinder and a shadow [(d) – (f)], with variations in the height of the top object.

### 3.2 *Results and discussion*

Figure 5 shows judged distance as a function of the height of the top cylinder, for the conditions with two cylinders and with one cylinder and a shadow, in the motion and stationary displays. In the motion displays, the judged distances fell between the optical-contact position of the top cylinder and the optical-contact position of the bottom cylinder or shadow. They were closer to the top-cylinder position when the bottom object was a cylinder and closer to the bottom-object position when it was a shadow. The effect of the height of the top cylinder was thus larger in the two-cylinder case than in the cylinder-and-shadow case. The results are similar for the stationary displays, except that there was a greater effect of the height of the top cylinder, when the bottom object was a shadow, for the stationary displays than for the motion displays. A three-way ANOVA showed significant main effects for motion versus stationary ($F_{1,20} = 10.13$, $p < 0.01$), two cylinders versus cylinder plus shadow ($F_{1,20} = 10.20$, $p < 0.01$), and cylinder height ($F_{2,40} = 54.39$, $p < 0.01$). The interaction of motion versus stationary

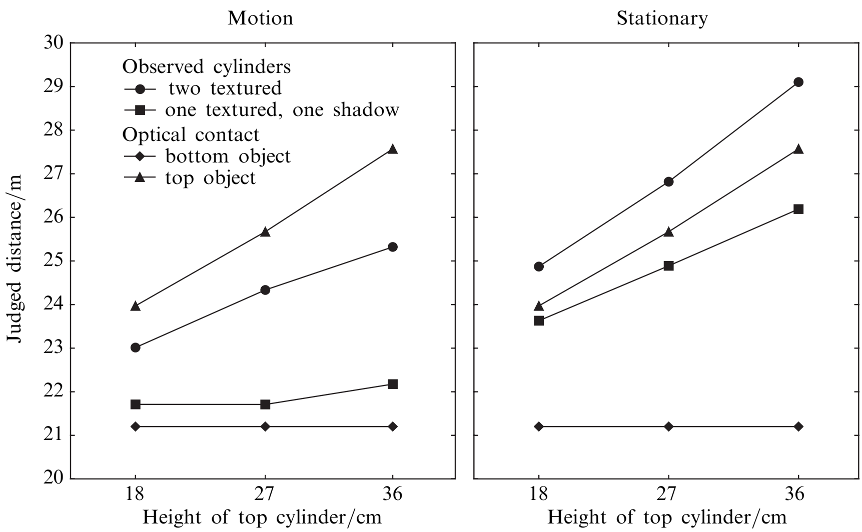



**Figure 5.** Judged distance as a function of height of the top cylinder in experiment 2.

and cylinder height was significant ($F_{2,40} = 8.02$, $p < 0.01$). The slopes of the two curves in figure 4 indicate that the motion displays were less affected by the height of the top cylinder. This suggests that common motion can substitute for proximity in grouping the two cylinders, or the cylinder and the shadow. The interaction of two cylinders versus one cylinder plus shadow with gap was also significant ($F_{2,40} = 11.14$, $p < 0.01$). This interaction reflects a larger effect of cylinder height for two cylinders than for one cylinder plus a shadow, suggesting that grouping by proximity is less important when matching a shadow with an object than when grouping two objects.

## 4 Experiment 3

In experiment 2, when there was one cylinder and a shadow which had the appropriate size and motion in the scene, judged distance of the cylinder was very close to the distance of the shadow. The same result might be expected if the size or motion of the shadow did not match that of the cylinder (Kersten et al 1997). If there is more than one potential shadow present, however, the perception of the distance of the cylinder would depend on which, if any, of the shadows was attributed to the cylinder. In this experiment we presented three alternative shadows with each cylinder, with the shadows varying in size and speed.

### 4.1 *Method*

4.1.1 *Observers.* The observers were nine students from the University of California at Irvine. All observers had normal or corrected-to-normal visual acuity and were naïve about the purpose of the experiment. They received extra credit for participating in the experiment.

4.1.2 *Stimuli.* The background scene was the same as in experiments 1 and 2. Three dark circular patches (elliptical in the image) were placed at simulated distances of 17.6 m, 21.2 m, and 24.8 m from the observer (see figure 6). Each patch had a simulated diameter of 36 cm. The projected sizes of the dark patches at 17.6 m, 21.2 m, and 24.8 m distances were 4.3, 3.6, and 3.1 cm. The simulated diameter of the cylinder was the same as in experiment 1. The projected size of the cylinder could take on one of three values, matching the projected size of one of the shadows. Similarly, the projected speed of the cylinder could take on one of three values, corresponding to the projected speeds of the three shadows. The combination of these three possible

**Figure 6.** Cylinders varying in projected size with three alternative shadows.

**Table 1.** Projected diameters and translations speeds of the cylinders and shadows. Only one cylinder was present on a trial; all three shadows were present. Sizes are in cm; speeds (in parentheses) are in cm $s^{-1}$.

| Cylinders | | | Shadows |
|---|---|---|---|
| 4.3 (4.6) | 3.6 (4.6) | 3.1 (4.6) | 4.3 (4.6) |
| 4.3 (3.6) | 3.6 (3.6) | 3.1 (3.6) | 3.6 (3.6) |
| 4.3 (3.0) | 3.6 (3.0) | 3.1 (3.0) | 3.1 (3.0) |

projected sizes and three possible projected speeds resulted in 9 conditions, as shown in table 1. In only three of these conditions, both the projected size and speed of the cylinder were consistent with one of the three shadows (with a light source assumed directly overhead at a great distance). In the other conditions, either the size or speed was consistent with one of the shadows. The optical-contact position of the cylinder was kept constant at 27.5 m. The scene and the shadows translated horizontally at the same speed as in experiments 1 and 2.

4.1.3 *Design.* There were two independent variables: the size of the cylinder, which was matched to one of the three shadows; and the speed of the cylinder, which was matched to the speed of one of the three shadows. The two variables were run within observers. Each observer responded to 6 repetitions of each of the 9 conditions. The 54 trials were divided into two blocks, with 3 repetitions of each condition randomly ordered in each block. The two experimental blocks were preceded by a practice block consisting of 3 replications of each condition.

4.1.4 *Apparatus and procedure.* The apparatus and procedure were the same as in experiments 1 and 2.

### 4.2 *Results and discussion*

Judged distance as a function of the projected speed of the top cylinder and the projected diameter of the top cylinder, averaged across observers, is shown in figure 7. These results indicate consistent effects of both cylinder speed and cylinder size, with cylinder speed having the larger effect. A two-way ANOVA showed significant main effects for speed ($F_{2,16} = 24.0$, $p < 0.01$), and for projected size ($F_{2,16} = 14.6$, $p < 0.01$). The interaction was not significant ($F_{4,32} = 1.9$, $p > 0.05$). Judged distance of the cylinder increased with decreasing speed and was also affected by projected size, indicating that observers used both motion and size in matching the cylinder to one of the alternative shadows.

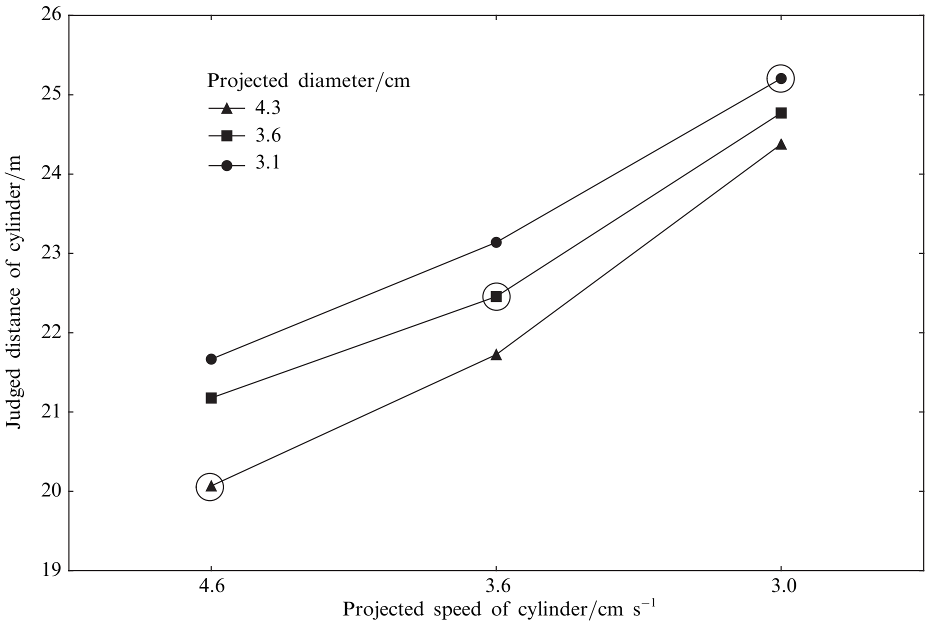



**Figure 7.** Judged distance of the top cylinder as a function of its projected speed and projected diameter. The circles indicate conditions in which one of the shadows matches the top cylinder in both projected speed and projected diameter.

### 4.3 *Control experiment*

A control experiment with stationary scenes was conducted with twelve naïve observers. The independent variable was the projected size of the top cylinder, which had the same three levels as in the main experiment (see figure 6). The mean judged distances for projected diameters of 4.3, 3.6, and 3.1 cm were 24.0, 25.2, and 26.6 m, respectively. An ANOVA showed a significant main effect for projected size ($F_{2,22} = 3.72$, $p < 0.05$).

## 5 Conclusion

Previous studies concerned with the role of optical contact in scene perception (Meng and Sedgwick 2001, 2002; Ni et al, in press) demonstrated that optical contact between an object and the ground, either direct or mediated by intermediate objects, is a major factor in determining the perceived position of an object in a 3-D scene. This research supported Gibson's (1946, 1950) theory that the perception of the distances of objects in a 3-D scene depends on how those objects are related to the ground surface. Research on the role of shadows in scene perception has found that shadows can determine the perceived position of a floating object (Kersten et al 1997) and disambiguate information for surface contact (Madison et al 2001). The present study was intended to bridge these lines of research by comparing effects on perceived layout with multiple objects to effects with objects and shadows, especially for dynamic scenes.

Two features that distinguish an object from a shadow were considered: an object can be textured and thick; a shadow is typically dark and has no thickness. Varying these two factors independently, we found that the judged position of a floating object was closer to the optical-contact position of a lower object when the lower object was a dark, shaded object than when it was a lighter, textured object, regardless of the thickness of the lower object. On the other hand, reducing the thickness of the lower object to zero only affected the judged position of the floating object when the lower object was a dark, shaded object. Note that this finding that the distance of the floating object was judged closer to the optical-contact position of the lower object, with a shaded lower object, is opposite to a prediction based on grouping by similarity. Grouping by similarity would be expected to result in two textured objects, rather than a textured object and a dark, shaded object, being perceived at the same distance. Instead, it appears more likely that a floating object will be associated with any dark object on the ground than with another textured object. This suggests that the tendency to match a floating object with a shadow persists even when the "shadow" is a thick dark object with shading. The change in slope of the function relating distance judgments to thickness that occurs between zero and other thickness levels for the shaded object does not occur when the lower object is a textured object. This suggests that the textured lower object continues to be perceived as an object, rather than as a shadow, even at zero thickness.

The differences in judged distance between two objects and one object and a shadow are similar in the stationary case, but both conditions show greater judged distance for the top object, with the judged distance in some conditions exceeding the optical-contact position of the top object. The similar upward shifts in the curves in the stationary conditions, relative to the motion conditions, in both figures 3a and 3b, suggest that these shifts are not related to differences between two objects and one object and a shadow. Instead, they are likely to be a result of errors in aligning the marker with the perceived position of the top object when the scene is stationary.

Variations in the height of a floating object largely determined its judged distance with two objects but had little effect on judged distance for an object and a shadow, in the motion displays. Height of the floating object had an effect for stationary displays, either with two objects or an object and a shadow. This indicates the importance of the common motion of an object and shadow for the dominance of the shadow location over other cues, such as optical contact, in determining the perceived position of objects.

If a shadow is to determine the perceived position of an object in a scene, the observer must associate the object with the shadow. This is straightforward if there is only one potential shadow present, but in natural scenes it is possible that several potential shadows might be available (eg shaded areas that are not shadows or shadows of other objects). Kersten et al's (1997) results indicate that the motion of the shadow is more important than the correspondence between the size of the object and the size of the shadow in determining the perceived position of an object. The results of experiment 3 demonstrate that, when alternative shadows are present, the relative projected sizes of the object and shadow affect perceived object location. However, in agreement with Kersten et al, the effect of motion was greater than that of size in determining which shadow would be matched to the object.

In the present experiments, we examined displays with objects moving together with the background scene, producing images similar to those produced by head movements and displays with the object and scene stationary. Our previous research (Ni et al, in press) suggested that in the motion case a constraint of common motion would group a floating object with an object on the ground, resulting in the floating object's perceived position moving away from its own optical-contact position and closer to the optical-contact position of the object on the ground. The results reported here demonstrate that the judged position of the floating object moves even closer to that of the object on the ground if the lower object can be interpreted as a shadow. These results can be interpreted in terms of a common-motion constraint, related to the rigidity constraint (Ullman 1979). With a moving scene, the common-motion constraint is effective either for multiple objects or for an object paired with a shadow. But this constraint is more effective when an object is paired with a shadow, because an assumption of a stationary light source (Kersten et al 1997) is less easily violated than an assumption of rigid motion among objects.[1] The relation between studies with multiple objects moving together and studies with moving shadows thus appears to be based on the use of a common-motion constraint in both cases, one relying on an assumption of rigidity or quasi-rigidity and the other relying on an assumption of a stationary light source.

**Acknowledgments.** This research was supported by NIH grant EY-12437. We thank Bart Anderson and an anonymous reviewer for helpful comments on an earlier draft, and Zheng Bian, Cary Feria, and Huiying Zhong for useful discussions.

**References**

Braje W L, Kersten D, Tarr M J, Troje N F, 1998 "Illumination effects in face recognition" *Psychobiology* **26** 371 – 380

Braje W L, Legge G E, Kersten D, 2000 "Invariant recognition of natural objects in the presence of shadows" *Perception* **29** 383 – 398

Braunstein M L, Andersen G J, 1984 "Shape and depth perception from parallel projections of three-dimensional motion" *Journal of Experimental Psychology: Human Perception and Performance* **10** 749 – 760

Castiello U, 2001 "Implicit processing of shadows" *Vision Research* **41** 2305 – 2309

Gibson J J, 1946 "Perception of distance and space in the open air", in *Motion Picture Testing and Research* Ed. J J Gibson (AAF program Report #7), reprinted in *Readings in Perception* Eds D C Beardslee, M Wertheimer (1958, Princeton, NJ: Von Nostrand)

Gibson J J, 1950 *The Perception of the Visual World* (Boston, MA: Houghton Mifflin)

Kersten D, Knill D C, Mamassian P, Bülthoff I, 1996 "Illusory motion from shadow" *Nature* **379** 31

Kersten D, Mamassian P, Knill D C, 1997 "Moving cast shadows induce apparent motion in depth" *Perception* **26** 171 – 192

Knill D C, Mamassian P, Kersten D, 1997 "The geometry of shadows" *Journal of the Optical Society of America A* **14** 3216 – 3232

Madison C, Thompson W, Kersten D, Shirley P, Smits B, 2001 "Use of interreflection and shadow for surface contact" *Perception & Psychophysics* **63** 187 – 194

Mamassian P, Knill D C, Kersten D, 1998 "The perception of cast shadows" *Trends in Cognitive Sciences* **2** 288 – 295

Meng J C, Sedgwick H A, 2001 "Distance perception mediated through nested contact relations among surfaces" *Perception & Psychophysics* **63** 1 – 15

Meng J C, Sedgwick H A, 2002 "Distance perception across spatial discontinuities" *Perception & Psychophysics* **64** 1 – 14

Ni R, Braunstein M L, Andersen G J, in press "Distance perception from motion parallax and ground contact" *Visual Cognition*

Puerta A M, 1989 "The power of shadows: Shadow stereopsis" *Journal of the Optical Society of America A* **6** 309 – 311

Ramachandran V S, 1988 "Perception of shape from shading" *Nature* **331** 163 – 166

Sauer C W, Braunstein M L, Andersen G J, Bian Z, 2001 "Judged shape of ground plane regions in realistic 3-D scenes" *Journal of Vision* **1**(3) 41

Ullman S, 1979 *The Interpretation of Visual Motion* (Cambridge, MA: MIT Press)

Wanger L R, Ferwerda J A, Greenberg D P, 1992 "Perceiving spatial relationships in computer-generated images" *IEEE Computer Graphics & Applications* **12** 44 – 58

Yonas A, Goldsmith L T, Halstrom J L, 1978 "Development of sensitivity to information provided by cast shadows in pictures" *Perception* **7** 333 – 341

[1] Although the rigidity constraint can influence judgments in structure-from-motion displays, 3-D structure can still be perceived when this constraint is violated (Braunstein and Andersen 1984).

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION

VOLUME 33 2004

www.perceptionweb.com