

# Moving Cast Shadows Induce Apparent Motion in Depth

---

Daniel Kersten, Pascal Mamassian & David C. Knill

---

## Abstract

We describe phenomenally strong visual illusions in which the motion of an object's cast shadow determines the perceived 3D trajectory of the object. Simply adjusting the motion of a shadow is sufficient to induce dramatically different apparent trajectories of the object casting the shadow. Using 3D graphics, we report psychophysical results which show that: 1) the information provided by the motion of an object's shadow overrides other strong sources of information and perceptual biases, such as the assumption of constant object size and a general viewpoint; 2) the natural constraint of shadow darkness plays a role in the interpretation of a moving image patch as a shadow, but under some conditions even unnatural light shadows can induce apparent motion in depth of an object; 3) when shadow motion is caused by a moving light source, the visual system incorrectly interprets the shadow motion as consistent with a moving object, rather than a moving light source. The results support the hypothesis that the human visual system incorporates a *stationary light source constraint* in the perceptual processing of spatial layout of scenes.

---

Correspondence should be sent to: Daniel Kersten, N218 Elliott Hall, Psychology Department, 75 East River Road, Minneapolis, MN 55455, U.S.A..

Email: kersten@umn.edu

**Perception 26, 171-192 (1997)**

---

## 1.0 Introduction

---

A cast shadow boundary occurs when one object partially blocks the light falling on another surface<sup>1</sup>. It has been appreciated at least since the time of Leonardo da Vinci that cast shadows can provide useful information about the relative depths between two such surfaces. Da Vinci elucidated the principle relating shadow displacement and the perception of depth in his advice to artists: "...when representing objects above the eye and on one side--if you wish them to looked detached from the wall--show, between the shadow on the object and the shadow it casts, a middle light, so that the body will appear to stand away from the wall." (da Vinci, 1970). Figure 1 illustrates how the relative displacement between an object and its cast shadow in an image provides a useful source of visual information about the spatial layout of objects. Although the importance of cast shadows for realism in the arts is widely appreciated,

---

Figure 1 about here

---

there have been only a few systematic studies of the informativeness of cast shadows for depth perception. Yonas, Goldsmith and Hallstrom (1978) showed that the location of a cast shadow can influence the judged depth and height of an object above a ground plane in observers as young as three years old. Using computer graphics simulations of relatively complex environments, Wanger, Ferwerda & Greenberg (1992) found that shadows provided the dominant pictorial cue for spatial and scaling tasks as compared with other manipulations which included texture, projection type, frames of reference, and motion. We know rather little, however, about the mechanisms by which human vision processes shadows, the strength of cast shadow

---

1. A cast shadow boundary can result when a surface region is blocked from a light source either by parts of itself (an *intrinsic shadow*), or by another object (an *extrinsic shadow*). An *attached shadow* boundary occurs where the angle between the surface normal and light source direction changes from acute to obtuse on a single surface. Knill, Mamassian & Kersten (1993) give an account of some of the geometrical properties of shadows.

information relative to other depth cues, or in particular the visual function of cast shadows which are moving.

Cast shadow motion is ubiquitous in everyday scenes where objects are often moving relative to each other. The visual system must cope with shadow motion either by treating shadow regions as noise irrelevant to the task of identifying and localizing surfaces, or by using this information. Moving shadows are used routinely in cartoon animations, video games and computer graphics suggesting an important role, at least for enhancing realism. But identifying and using cast shadows is a hard theoretical problem, and is exacerbated by the fact that the cause of a cast shadow is remote from the surface receiving it. Thus it is perhaps not surprising that, as observers, we occasionally misinterpret both moving and static shadows<sup>2</sup>. One goal of the research reported here was to seek out conditions under which shadows constitute a robust source of information about depth that is resistant to conflicting cues and high-level knowledge. A strong illusion of apparent changes in movement can provide a simple phenomenal test of an hypothesis, that can be convincing based on observation alone. In this context, perhaps the simplest question to ask is: Can shadow motion induce apparent motion in depth even when zero motion flow of the object relative to the background indicates no motion at all? An affirmative answer to this question does not, by necessity, follow from observations based on static shadows (e.g. as in Figure 1). Except for an accidental viewpoint (when an object moves along a fixed line of sight), a change in depth is accompanied by change in the object's image position and size. Thus, shadow movement is almost always accompanied by

---

2. In Rembrandt's well-known painting of the "The Night Watch", the left hand of the militia captain casts a shadow onto the coat of his lieutenant. Because this realistic shadow seems remote from its cause, it can be confused with a "stain" on the coat (William H. Warren, personal communication; Arnheim, 1954). Moving cast shadows have less ambiguity with stain or more generally with reflectance.

motion, in the image, of the casting object. Relative motion is a strong cue to relative depth, and it is an experimental issue as to whether shadow movement could override the motion cues (zero position and size change) to induce apparent movement in depth. In a recent study using a version of what we call the “square-over-checkerboard” scene, we had observers view a square surface whose image was stationary against a checkerboard background. When a shadow was moved in a way consistent with the square’s motion along the line of sight, observers reported an illusion of motion in depth, despite the lack of either image size or position change (Kersten, Knill, Mamassian & Bühlhoff, 1996). The effect was strongest for a shadow below, rather than above the object, and for fuzzy rather than sharp shadows. This simple, yet dramatic, observation indicates that a moving shadow can induce apparent motion in depth even in the presence of other evidence to the contrary. In the next 4 sections, we explore the generality of this conclusion, and the constraints that the visual system respects in making an inference of 3D motion from a shadow trajectory.

In the next section (Section 2), we investigate the generality of the basic effect of cast shadow motion on apparent depth. In our first experiment<sup>3</sup> (Experiment 1, Section 2.1), we use the square-over-checkerboard scene to test whether the shadow has to be dark, rather than (unnaturally) light. The perception of depth from motion parallax is contingent on a general, rather than accidental view of the object against a background. In order to generate animations from a general viewpoint, and to provide a convenient platform for quantitative measurements, we developed the “ball-in-a-box” scene which provides the basic stimulus arrangement for all the subsequent experiments and demonstrations. The first demonstration (Demonstration 1, Sec-

---

3. Empirical observations are reported as either “Demonstrations”, in which the results describe qualitative observations which all observers see, or “Experiments” which report quantitative results which address more specific issues.

tion 2.2) shows the effect of eliminating an accidental view of the object relative to its background. Demonstration 2 (Section 2.3) tests whether the form of an object's spatial trajectory can be influenced by varying the form of the shadow's trajectory.

Section 3 studies the relationship between object size and shadow motion. Demonstration 3 (Section 3.1) tests whether an object's deformation or apparent size can be influenced simply by adjusting the motion of the object's cast shadow. Experiment 2 (Section 3.2) explores how a change in image size interacts with depth from shadows. When an object moves closer, its image size increases. Which source of depth information dominates, a change in image size or shadow displacement?

A dynamic image is determined by a complex set of interactions between object, illuminant, and viewer locations. Prior knowledge or estimates of the illumination position could, in principle, simplify the computation of depth from moving shadows. Section 4 investigates how the visual system determines and represents light source location. Experiment 3 (Section 4.1) tests whether observer's position settings are consistent with a fixed light source position, and whether these judgments vary with changes in the shadow properties of contrast and opacity. The data are consistent with an implicit assumed light source position that differs between individuals. Experiment 4 (Section 4.2) tests whether an observer's implicit light source position is determined by prior assumptions or by image data.

In order to arrive at an explanatory principle for our findings, Section 5 introduces the *stationary light source constraint*. We test whether this constraint is used by creating animations of the ball-in-a-box in which the light source itself is moved to simulate shadow motion and then test if observers interpret the scene in terms of changes in ball or light source trajectory (Demonstrations 4-7). We conclude, in Section 6, with a general discussion of the problems of com-

puting shadow identity locally, and the global constraints which could possibly be used to resolve depth ambiguity using moving shadow information.

## **2.0 Shadows affect an object's 3D trajectory**

---

The principle of depth-from-shadows is simple: the closer an object is to its cast shadow in an image, the closer it should appear in depth to the background surface. Using two types of scene simulation, this section investigates conditions under which shadow motion can influence object 3D trajectory, even when other sources of depth information are in conflict.

### **2.1 Experiment 1 - Square-over-checkerboard: Dark shadows are more effective than light shadows.**

In general, depth perception is determined by multiple cues, including stereo, motion parallax, and various pictorial cues (Cutting & Vishton, 1995). Typically when an object's depth changes, its image shifts and under perspective projection its size changes with respect to the background. We noted above that, under conditions when object image motion is zero relative to the background, shadow movement is sufficient to override the strong motion information that the object depth is not changing (Kersten et al., 1996). In this experiment we provide further support for the strength of cast shadows as a depth cue, and specifically test whether the depth percept is sensitive to conditions which might affect the perceptual labelling of the putative shadow as a shadow. In particular, is the percept dependent on the natural condition that the shadow be darker than the surround?

*2.1.1 Methods.* We created a motion analog of the demonstration in Figure 1, in which the shadow cast by a stationary square moves back and forth relative to the square (Figure 2). Unlike Figure 1, however, the shadow in this experiment had a sharp boundary. Observers were asked to look at a fixation mark (+) placed on a checkerboard plane which subtended 6.6

x 10° of visual angle. Viewing distance was 500 mm. At a position 4.1° to the right of the fixation point, a foreground square (2.24° x 2.24°) was superimposed over a shadow of the same size as the square. (Informal observations had indicated that direct fixation was less likely to produce an illusion of motion in depth, presumably because of the greater sensitivity to spatial position changes, or lack thereof, in and near the fovea.) In a 500 msec. animated sequence, the shadow oscillated for one cycle through a 0.34° diagonal displacement from the foreground square. The combined horizontal subtense of the square with the shadow at maximum displacement was 2.48°. The foreground square remained stationary throughout the animation. Observers were asked to indicate whether the foreground square “appeared to move in depth, away from and then back to the checkerboard background”, or not. Six different types of shadow were used for the experiment: three “dark” shadows simulated as film transparencies with transmittances of 12, 16, and 36%; and three physically implausible “light” shadows corresponding to transmittances of 180, 284, and 394% (i.e. light was added within the shadow). The background checkerboard had a mean luminance of 17.4 cd/m<sup>2</sup> with an 82% contrast between dark and light squares. Subjects were split into two groups of ten. The order of presentation of different shadow conditions for one group, in terms of effective transmittance, was: 16, 284, 12, 394, 36, and 180%. The other group saw the stimuli in the order 284, 16, 394, 12, 180 and 36%. Each subject viewed three series of presentations, making a total of 18 trials.

*2.1.2 Results and discussion.* When the shadow was rendered realistically dark, subjects reported seeing the square move toward and away from the background surface 78% of the time. When the shadow was implausibly lighter than its background, subjects only reported seeing the square move in depth 40% of the time. (A Wilcoxon signed rank order test on the difference between light and dark shadows gave  $p < 0.001$ ). Subjects who perceived the motion

reported that the percept was phenomenally strong and “immediate”. The result clearly shows

---

Figure 2 about here

---

an effect of cast shadow motion on observers' perception of 3D motion of an object.

Note that, for the cases in which observers saw the motion in depth, the motion of the shadow overrode conflicting cues which suggested that the square was stationary: the lack of any change in size of the square and the lack of any 2D motion of the square in the image. That these features of the stimulus would suggest object stationarity results from the human visual system's bias to assume, first, that objects do not change size over time (a “constant object size constraint”), and second, that the viewer is viewing the scene from a non-accidental, or general viewpoint (Lowe, 1985; Nakayama, & Shimojo, 1992). The assumption of constant object size would lead the visual system to interpret the non-changing size of the square as information that the square was stationary, since any change in depth of a rigid object would lead to a correlated change in the size of the object's image. The general viewpoint assumption would lead the system to interpret the lack of any 2D motion of the square also as information for stationarity, since for almost all viewpoints (except one accidental view in which the viewer is looking along the direction of motion), motion in depth of an object would cause a correlated 2D motion of the object's image. The cues for stationarity could well have led to the result that on 22% of the trials with dark shadows, subjects did not see the square move in depth. This raises the possibility that elimination of the stationarity cues would lead to greater effects of cast shadow motion on observers' percepts of 3D motion. Unfortunately, one cannot remove the effect of the constant object size constraint from the stimulus, since size change, even if zero, is still a cue. Later in Experiment 2 (Section 3.2), we systematically vary image size and shadow information. In the present case, we set the size cue change to zero, thus pro-



viding conflicting information that the square was not moving in depth. This leaves us with at least two possibilities to increase the strength of the percept of motion in depth for cast shadows: 1) increase the strength of the shadow information to the extent that it overrides the general viewpoint constraint or; 2) remove the effect of the general viewpoint constraint by moving the object. With respect to the first possibility, a shadow with changing penumbral blur (caused by an extended “panel” light source) is sufficient to induce a robust perception of apparent motion in depth 100% of the time on initial viewing of the square-over-checkerboard (Kersten et al., 1996). This is true even with direct fixation of the square patch. When the light source was a point (as in Experiment 1), the percept was seen, but not 100% of the time. The second possibility is to remove the effect of the general viewpoint constraint by simply moving the object, as well as its cast shadow, in the image plane. We do this in the next demonstration using a ball-in-a-box scene.

## **2.2 Demonstration 1 - Ball-in-a-box: Phenomenally strong illusion of motion in depth with accidental view removed**

We generated a 3D graphics simulation which we call the ball-in-a-box animation (Figure 3), in which we simulated a ball moving inside a box in such a way that it followed a diagonal trajectory in the image plane. In addition to removing the accidental view, the ball-in-a-box provides us with a straightforward way of measuring apparent 3D location of an object in Experiments 2, 3 and 4. As in Experiment 1, the size of the object's image, in this case that of a ball, remained fixed throughout the animation.

---

Figure 3 about here

---

*2.2.1 Methods.* A ball was placed in a small open box, rendered in perspective projection and viewed from an elevation of  $22^\circ$  relative to the floor of the box. Only three sides of the box were visible. The bottom had a checkerboard texture (square tiles of alternate gray). (A checkerboard pattern allows the experimental manipulation of both shadow opacity and contrast-- a property which is exploited later in Experiment 3.) The viewpoint was offset slightly to the right, as shown in Figure 3. Each animation was created in two stages: first, we rendered a scene with a moving ball without cast shadows. Second, we independently added the ball's cast shadow to the images in an animation, so that we could manipulate the motion of the shadow independently of the ball's motion. The shading on the ball was generated by simulating a light source at infinity with a slant of  $63^\circ$  degrees relative to the floor of the box. The ball moved in a linear trajectory in the image at an angle tilted by  $22^\circ$  from the horizontal. Its speed varied sinusoidally (period = 4 sec), so that the ball repeated its motion back and forth between its left- and right-most positions in the image. The shadow moved so that it remained vertically below the ball in the image. Only the distance between the shadow and the ball varied as the shadow and ball moved. The demonstration animations were recorded on videotape, and observers were shown the taped animations. All of the Demonstrations (1-7) were developed with Alias|Wavefront's<sup>TM</sup> The Advanced Visualizer<sup>TM</sup> on a Silicon Graphics Workstation.

There were two different animation sequences: In the first, the ball's cast shadow followed a diagonal trajectory in the image, parallel and nearly identical to that of the ball's image (ending up at the position shown in Figure 3b); in the second, it followed a horizontal trajectory (ending up at the position shown in Figure 3c).

*2.2.2 Results and discussion.* Despite the fact that the ball's image remained the same size and had an identical trajectory in the image plane in both animations, all observers experienced the striking percept of seeing the ball rise above the checkerboard floor when the shadow trajectory was horizontal, and recede smoothly in depth along the floor when the slope of the shadow trajectory matched that of the ball. Because the size of the ball's image remained fixed, it is clear that the apparent depth from the moving cast shadow was sufficient to override the constant size constraint.

### **2.3 Demonstration 2: Non-linear shadow trajectory induces the illusion of a non-linear object trajectory**

The second demonstration explores the versatility of a shadow movement to influence the spatial trajectory of the casting object by testing whether a non-linear shadow trajectory can induce an apparently non-linear spatial trajectory of the ball.

*2.3.1 Methods.* We modified the ball-in-a-box animations in the following way: the shadow was given a non-linear motion trajectory in which it initially touched the ball's image, moved towards the front of the box, at mid-trajectory returned to touch the ball's image, and then swung to the front again (see Figure 4a). The ball's image moved in the same straight, diagonal trajectory as before.

---

Figure 4 about here

---

*2.3.2 Results and discussion.* All observers reported seeing the ball as moving in a non-linear 3D trajectory in which the ball appeared to come forward, retreat in depth, and then come forward again, as it moved from left to right in the box. Moreover, the observers reported seeing a singularity, or bounce, in the path of the ball when the shadow touched the ball's image and

changed direction. Observers saw the bounce despite the fact that the ball's velocity in the image was smooth at that point.

The above ball-in-a-box demonstrations (1 & 2) show that cast shadows can be remarkably effective and versatile in their effects on spatial layout. They can directly influence the perceived form of the spatial trajectory of an object. Related experiments have shown that cast shadows have further versatility in that they can affect apparent rigidity and shape of an object (Bülthoff, Kersten, & Bülthoff, 1994).

### **3.0 Shadows interact with object size**

Simple geometry shows that the image size of an object is proportional to its object size, and inversely proportional to distance. This section investigates how depth-from-shadows influence apparent object size (Demonstration 3, Section 3.1), and how a change in image size influences apparent depth (Experiment 2, Section 3.2).

#### **3.1 Demonstration 3: Apparent depth produced by cast shadows induces apparent size change.**

If observers have an implicit perceptual assumption that objects do not change physical size, one would predict that when the slope of the shadow trajectory matched the ball, the ball would appear to grow in size as it recedes in depth. Indeed, several of our observers reported this perception in Demonstration 1. Yonas et al. (1978) found that static size judgments in adults were influenced by apparent depth induced by cast shadow position. With a moving object, however, the object's identity clearly remains unchanged and thus poses a stiffer constraint on object size. In the next animation, we increased perspective convergence by lengthening the box to see whether the ball's shadow could induce sufficient apparent depth change of the ball to affect its perceived size.

*3.1.1 Methods.* The stimulus specifications for Demonstration 3 were similar to those of Demonstration 1, except that the length of the box in world coordinates was tripled (Figure 5).

---

Figure 5 about here

---

For constant ball size, the image should decrease in size by about 50% if it were indeed receding to the back of the box. However, as before, the image of the ball was kept constant. The ball made a full excursion (in the image) from the lower left corner of the box to the upper right corner.

*3.1.2 Results and discussion.* When the trajectory of the shadow matched the ball, all observers reported seeing the ball apparently inflating as it moved to the back of the box, and then shrinking as it moved forward. When the shadow trajectory was horizontal, the ball appeared to remain fixed in size, although some observers reported seeing a small apparent size change, perhaps due to the fact that increased height in the picture is associated with increased depth.

## **3.2 Experiment 2: Depth-from-shadows vs. image size change**

In this experiment, we measured the apparent position of the ball-in-a-box as a function of both cast shadow slope and size change of the image of the ball. It might be the case that the shadow-induced location change we observed in Demonstration 1 is the result of the lack of salience of a zero size change as a cue to depth. If this is the case, we might expect that if we change the image size of the ball as it moves, this information would dominate the shadow cue. To investigate this issue, we manipulated the information provided by each cue in both consistent and conflicting ways.

*3.2.1 Methods.* A total of twelve observers participated in this study. Eight of these observers were undergraduate students from the University of Minnesota, paid for the time spent in the

experiment. Two other observers were recruited within the vision laboratory, and the last two observers were the first two authors. Except for the authors, all subjects were naive regarding the purposes of the study. All observers had normal or corrected-to-normal vision.

A realistic three-dimensional scene was simulated as described in Demonstration 1 (Section 2.2.1 Methods). The objects of the scene had their positions updated 30 times per second, and were displayed on a 19-inch high-resolution (1280x1024 pixels) monitor. The observer looked at the monitor monocularly behind a reduction screen. Head-movements were restricted by the use of a chin-rest, and the lights of the room were turned off.

The box was deeper than wide (9 x 6 tiles), and was simulated to be 150 mm deep, 100 mm wide, and 55 mm high; the center of its bottom was 330 mm in front and 130 mm below the viewpoint, and the plane of the bottom was slanted by 22 deg towards the viewer. The shaded ball had a diameter of 18 mm. The cast shadow was rendered as a transparent elliptical patch such that the intensity of the tiles covered by the shadow was reduced by 20 per cent. The shading on the ball was computed from a fixed light source positioned at infinity, the direction of the light rays being perpendicular to the bottom of the box (therefore slanted from the vertical by an angle of 22 deg towards the viewer).

The ball was given an oscillating motion along a linear trajectory in the image plane (the tilt of the trajectory relative to the horizontal was 33 deg). The shadow was given a similar oscillating motion, but the tilt of its trajectory was manipulated between trials. At the left-most position of the ball, the shadow was always touching the ball, but at the right-most position, the distance between the ball and its shadow varied according to the first independent variable named shadow slope. The shadow slope varied between zero and one: a value of zero referred to a horizontal shadow trajectory, while a value of one referred to a slope identical to the slope

of the ball trajectory (thereby consistent with a ball moving on the bottom of the box). The slopes were 0.1, 0.42, 0.68, and 0.9. Figure 4b shows a schematic illustrating how shadow slopes varied (slopes are shown are not exact). The second independent variable was the degree by which the ball and its shadow were shrinking. This variable was named size change and also varied between zero and one: a value of zero corresponded to no change of size and a value of one was consistent with a ball moving on the bottom of the box (a value of one corresponded to a shrinking of the ball image by 29 per cent from one end to the other of its trajectory). Four levels were selected for each of these two variables, chosen such that for two consecutive values, the right-most position of the ball would be equi-distant in 3-D space. The two variables were manipulated independently, their values being either equal or different, providing respectively either consistent or conflicting information about the spatial location of the ball in the box.

Observers were requested to follow the ball with their eyes for a full cycle of the ball trajectory (from left to right, and back to left). Once the ball reached again its left-most position, a red line (horizontal relative to the floor) was displayed on the right side of the box. The task of the subject was then to adjust the height of this line (shown in Figures 3b and c, and schematically in Figure 6 ) to the perceived height of the center of the ball relative to the bottom of the

---

Figure 6 about here

---

box, for the right-most position of the ball trajectory. Subjects adjusted the height of the line by moving the computer's mouse and indicated a match by pressing the mouse button. Before running the experiment, the subject ran four practice trials for which the shadow and size cues were always consistent. Thereafter, the subject ran 8 repeated trials for each of the 16 conditions, all randomized within two consecutive blocks.

3.2.2 *Results and discussion.* Figure 7 shows data from four observers in which perceived height is plotted as a function of size change. If size change vetoed shadow slope information, we would expect all four shadow conditions to lie on the same line. This was not the case for any of the observers. For all but two observers (who ignored size change as a cue to spatial location), both cues contributed to the perceived height. However, the individual differences between subjects were large, similar to what has previously been reported in studies involving two conflicting sources of information (cf. Doshier, Sperling and Wurst, 1986). To analyze the contribution of each cue to the perceived location of the ball, we performed an analysis of variance for each subject. A main effect of either the shadow slope or the size change variable would indicate a significant contribution of the corresponding cue for the evaluation of the ball position. Moreover, an interaction between these two variables would indicate that the information from the two cues does not merely accumulate, but rather is being combined in a non-linear way.

We divided the twelve subjects into four groups according to the significant effects of their ANOVAs, under an alpha level of 0.05. Data from four observers, representative of these four groups, are illustrated in Figure 7 (panels a-d), and were characterized by: (a) a main effect of both independent variables and no interaction; (b) a main effect of both independent variables and a significant interaction; (c) a main effect of shadow slope and an interaction of this variable with size change, but no main effect of size change alone; and (d) only a significant main effect of shadow slope. The distribution of the observers within the four groups was as fol-

---

Figure 7 about here

---

lows: six subjects (including the first author) in group (a), two subjects (including the second author) in group (b), two subjects in group (c), and two subjects in group (d). For two subjects



in group (a), the interaction between shadow slope and size change was almost significant ( $F(9, 112) = 1.96, p = 0.051$ , and  $F(9, 112) = 1.93, p = 0.056$ ), so that under a laxer criterion, the population of group (b) would be doubled.

A careful analysis of the data from observers in group (c) showed that the main effect of size change was actually annihilated by the interaction effect (see Figure 7c). The effect of size change on the perceived height of the ball was different when the shadow slope was 0.9 than for the other levels of shadow slope. In a post-experiment interview, these observers claimed that at some occasions the ball seemed “to grow as it moved to the back of the box”, consistent with the observation of Demonstration 3. This perceived inflation of the ball might well occur when the cast shadow trajectory was consistent with a ball almost on the bottom of the box (shadow slope = 0.9) while the size of the ball stayed roughly the same (size change = 0.1), and account for the differential performance between observers from group (c) and those from group (b)<sup>4</sup>.

The main conclusion is that for all observers, cast shadow information was a strong and salient cue for spatial location, even when a size change of the ball’s image indicated a conflicting spatial location.

#### **4.0 Implicit light source position**

---

Like many other monocular cues, the relative displacement of an object’s image and its cast shadow provides theoretically ambiguous information for spatial layout. In order to interpret

---

4. A reasonable interpretation can also be advanced for the lack of main effect of size change for the observers in group (d). Several subjects complained that the stimulus “didn’t look right”, even though they were usually unable to report what was wrong. As a consequence, some of these subjects might have chosen (voluntarily or not) to attend only to the most salient cue, which turned out to be the moving cast shadow in our experiment.

the cues, the visual system must use other information about the scene and make prior assumptions about the world. Since cast shadow displacement is a function of both object position and light source position (Figure 8), the visual system must make implicit assumptions, or inferences from image data, about the position of the light source creating the shadows in order to infer the spatial positions of the casting objects. In this section, we present experimental data which reveal the nature of the information and prior assumptions about light source position which the visual system brings to bear on the interpretation of cast shadow motion.

---

Figure 8 about here

---

#### **4.1 Experiment 3: Consistency of apparent position judgments with varying shadow opacity and contrast.**

Figure 1 suggests that, at least when no information about multiple light sources is provided in an image, the visual system relies on the assumption of a single light source (a constraint similar to the light source from above constraint used to explain certain effects in the perception of shape from shading (Gibson, 1950; Ramachandran, 1988)). In this experiment, we wanted to know whether the spatial percepts are consistent with an implicit light source position for different shadow trajectories over trials, and to what extent the judgments relate to the actual light position. We also wanted to know how apparent ball height in the ball-in-a-box depends on the photometric properties of the shadow. Experiment 1 showed that light shadows were less effective than dark shadows in producing apparent motion in depth of a stationary square. Does a similar result hold for a moving ball in a box? We had observers make settings of apparent ball height for the simulation described in Demonstration 1, with the slope of the shadow trajectory (Figure 4b), the shadow opacity and contrast as independent variables.

*4.1.1 Methods.* Three subjects viewed the ball-in-a-box animations as described in Demonstration 1 and Experiment 2. The box was now 132 x 132 mm (6 x 6 tiles) and was viewed from a point 355 mm from the center of the floor of the box at an angle of 22° relative to the horizontal. The viewpoint was also 132 mm vertically above and 330 mm horizontally from the center of the box floor. The screen width of the front of the box was 190 mm or 30° of visual angle. The ball's image size was 2.8° (16 mm in screen coordinates) and never changed during the animation. The vertical and horizontal sweep amplitudes of the ball were 32 mm and 79 mm in screen coordinates.

The main independent variable was the slope of the shadow trajectory which was either 0.0, 0.38, 0.71 or 1.0, where 0 corresponds to a horizontal trajectory, and 1 to a trajectory that matches that of the ball (Figure 4b). In terms of angle relative to the screen horizontal, these slopes corresponded to 0, 8.5, 16, or 22 degrees. The four shadow slopes in the experiment correspond to 20 mm steps in world coordinates in depth along the box floor. In addition to the shadow slope, we manipulated two photometric variables defining the shadow. Because the shadow was projected on a checkerboard, and not on a surface of uniform reflectance, the shadow was defined by two luminances which were specified in terms of opacity,  $P$  and contrast  $C$ . Let  $B_1$  and  $B_2$  represent the luminances of the checkerboard without the shadow patch, and  $S_1$  and  $S_2$  the two luminances of the shadow patch.  $S_1$  and  $S_2$  are determined by the following equations:

For  $-1 < P < 0$ ,

$$\begin{aligned} S_1 - S_2 &= C \times (1 + P) \times (B_1 - B_2) \\ S_1 + S_2 &= (1 + P) \times (B_1 + B_2) - 2P \end{aligned}$$

For  $0 < P < 1$

$$\begin{aligned} S_1 - S_2 &= C \times (1 - P) \times (B_1 - B_2) \\ S_1 + S_2 &= (1 - P) \times (B_1 + B_2) \end{aligned}$$

The opacity dimension determined the overall darkness of the shadow. A positive opacity served to decrease the brightness of the shadow. On the other hand, a negative opacity produced a bright patch in the image, inconsistent with a plausible shadow. Opacities were: -1, -0.66, -0.33, 0, 0.33, 0.66, 1. The contrast variable determined the consistency of the luminances inside and outside the shadow along the shadow boundary. Again, a positive value corresponded to a natural shadow, while a negative value produced a patch in the image inconsistent with the checkerboard colors. The unshadowed dark and light squares of the checkerboard were 15 and 36 cd/m<sup>2</sup>, respectively.

As in Experiment 2, subjects were given the task of interactively adjusting a line along the right wall (see again Figures 3b and c, and Figure 6) to match the apparent height of the middle of the ball at the right-most point of its trajectory. The motion of the ball and its shadow continued throughout the course of a trial. The order of the 84 conditions was randomized. The three subjects each made 8 settings for each of the conditions.

*4.1.2 Results and discussion.* Of the three factors, only shadow slope produced a main effect ( $p < 0.001$ ;  $p > 0.1$  for contrast and opacity). Figure 9 shows ball height as a function of opacity for a contrast of -1 (Figure 9a) and +1 (Figure 9b) for one observer. Observers' settings of

---

Figure 9 about here

---

ball height did not depend on the photometric properties of the shadow (the shadow was invisible for a contrast of +1, and opacity of zero, Figure 9b). Further, there were no significant interactions between the three conditions when the zero opacity condition (no visible shadow) was excluded.

How consistent are observers' settings with a fixed light source position? Figure 10 shows results obtained for three observers for the natural shadow condition corresponding to a transparent dark shadow (contrast = 1 and opacity = 0.66). The height estimates of all three subjects varied systematically with the slope of the shadow trajectory: smaller slopes, corresponding to larger divergences between the shadow and the ball, resulted in larger height estimates. This reflects differences in the perceived 3D motion of the ball between that of receding along the floor (for large slopes) to that of rising above the floor (for small slopes). If the observers based their setting on the actual light source position (which was at infinity), the settings would have fallen on the solid lines shown in the plots. While this was a good fit for only one observer (subject WB), we were able to obtain a better fit to each subject's data by finding what would amount to a perceptually implicit fixed light source position for the subject. In terms of distance (mm) from the middle of the checkerboard floor and slant (deg) with the floor, the light positions which fit the data best were: (419 mm, 60.8°); (105 mm, 50.4°); and, (67 mm, 46.8°) for observers WB, GDA, and PB, respectively. These fits are shown with dashed lines. Observers behaved as if they had fabricated a fixed illumination arrangement with which to interpret the scene. Any such fabrication, however, would have to have been unconscious, for when queried after the experiment as to where the light source was, observers claimed to have not thought about it.

---

Figure 10 about here

---

#### **4.2 Experiment 4: Is effective light source direction determined by prior assumptions or image data?**

The data from the previous experiment (Experiment 3) suggest that the visual system uses a strategy in which it effectively accounts for light source position when interpreting cast shadow motion over different shadow trajectories. The question remains as to how the human

visual system incorporates knowledge of light source position in generating percepts of 3D object motion from cast shadow motion. In this experiment, we tested whether subjects' implicit light source direction is determined by the shading information on the ball or a prior bias.

*4.2.1 Methods.* We ran the same ball-in-a-box experiment used for Experiment 3 but with three different shading conditions for the ball, corresponding to three different, fixed light source positions. 40 subjects were split into 3 groups (13, 13 and 14). Each group was shown a different shading pattern, corresponding to being illuminated by a light source from one of three angles above the checkerboard:  $60^\circ$ ,  $90^\circ$  and  $120^\circ$  (recall that the viewing direction was  $22^\circ$  above the checkerboard). All light sources were at infinity. Each observer made 16 settings at each of four shadow trajectory slopes: 0.1, 0.42, 0.68, and 0.9.

*4.2.2 Results and discussion.* If observers used the ball's shading to determine a light source direction for the estimation of 3D object motion from shadow motion, subjects' estimates of the ball's height at the end of its trajectory should have varied accordingly. The dotted lines in Figure 11 show theoretical predictions of ball height as a function of shadow slope for the three illumination angles based on exact knowledge of the light source direction. Figure 11 also shows the perceived height, averaged over all observers for each of the three conditions. The data show a small effect of illumination direction consistent with observers' usage of shading information to indicate light source direction. It can be seen from the figure, however,

---

Figure 11 about here

---

that the size of the effect was far from what would be predicted theoretically, suggesting that in this experiment, a strong prior bias for a default light source position determined perfor-

mance. Our results, of course, do not rule out the possibility that stronger image information for light source position, than that provided by the ball's shading, may have a greater influence on the subjects' interpretation of cast shadow motion.

## **5.0 Stationary light source constraint**

In order to explain the perception of motion in depth from moving cast shadows, we conjecture that the visual system assumes that the light source casting a shadow is fixed on the time scale of the motion. We call this the *stationary light source constraint*. The idea is that if image motion in a scene could be accounted for in terms of either object or light source motion, the visual system assumes the object is moving, rather than the light source. Consider the apparent non-linear motion in Demonstration 2. If we ignore for the moment the lack of change in shading on the ball and elsewhere in the box, as well as the lack of changes in shadow shape, the non-linear motion of the shadow could be interpreted as due to an appropriately moving light source. Our assumptions, however, leave us with the question of whether cues elsewhere in the box provide additional information that the light source is not moving.

### **5.1 Demonstrations 4-7: Can the visual system account for a moving light source?**

A test of the stationary light source constraint would be to find out whether the visual system can account for a moving light source in its interpretation of cast shadow motion when appropriate information about the motion of the light is provided in a sequence of images. We made four animations using a moving light source to generate the cast shadows. The animations were designed so that observers should see qualitatively different object motions if they assume a fixed light source constraint than if they accounted for the light source motion.

*5.1.1 Methods.* All the animations were based on a realistic 3D simulation of a ball oscillating in the front plane of the box. The motion of the ball was chosen to give the same image trajectory as was used in the previous ball-in-a-box demonstrations and experiments--moving diagonally in the image plane, with no change in image size. Constant image size was achieved through small continuous adjustments of the 3D size. The simulated trajectory of the ball was not only fixed in the image, but was also fixed in 3D for all animations. The previous demonstrations and experiments manipulated the shadow trajectory without varying the lighting on the rest of the scene. In Demonstrations 4-7, we generated shadows by rendering the scene with ray-tracing from a moving light source whose motion gave rise to different trajectories for the cast shadows. In these animations, the continuously changing shading on the ball and in the room provided information for the motion of the light source. A system which could effectively discount this motion should see the *same* 3D motion of the ball in all the animations. That is, the system should infer the correct 3D trajectory as specified by the geometrical model defining the animations.

For the first of the demonstrations in this section (Demonstration 4), we made two animations in which the simulated light source motions gave rise to cast shadow trajectories mimicking those used in Demonstration 1 (one following the ball, the other moving horizontally in the image). Demonstration 5 was identical to Demonstration 4, except that the spherical ball was replaced by an ellipsoid. In Demonstration 6, we added further information about the moving light source by including other stationary objects (vertically elongated parallelepipeds) placed on the floor of the box (as also used in Demonstration 7, and shown in Figure 12). Finally, we generated an animation (Demonstration 7; Figure 12) in which the motion of the light source caused a non-linear shadow motion which mimicked that of Demonstration 2, but with the objects of Demonstration 6.



---

Figure 12 about here

---

*5.1.2 Results and discussion.* The four demonstrations support the hypothesis that the visual system relies on a fixed light source constraint when interpreting shadow motion. For Demonstration 4, all observers reported seeing the ball as moving along different 3D trajectories in the two animations as in Demonstration 1. When asked to compare the perceived object motions in these animations with those in the animations used for Demonstration 1, all observers reported that they appeared the same. This suggests that the observers were not able to incorporate the information for a moving light source into their estimation of object motion. The result, however, may have arisen either because observers interpreted the changing shading of the ball as being due to something other than a moving light source or because the changing shading on the ball and in the room did not provide sufficient information to induce the percept of a moving light source. In support of the former hypothesis, several observers reported that the ball appeared to rotate and that the shading on the ball then appeared to be from markings on the ball's surface. In order to control for this effect, we repeated Demonstration 4 using an ellipsoidal instead of a spherical ball (Demonstration 5). This led to a correct interpretation of the shading pattern (the ellipsoid did not appear to rotate); however, the phenomenon remained unchanged--observers still reported seeing different motions for the ellipsoid in the two animations. The animation for Demonstration 6 included several visible moving cast shadows for the stationary objects, providing even more information for the motion of the light source, yet we found no effect on the apparent trajectory of the ball. When the animation for Demonstration 7 was shown after the animation used in Demonstration 2, observers reported that their percepts of non-linear 3D motion were similar for both animations. We cannot rule out the possibility that adding even more information about the presence

of a moving light source would outweigh the stationary light source constraint; however, taken together, Demonstrations 4-7 provide strong evidence that the human visual system incorporates an assumption of a fixed light source in its interpretation of 3D object motion from cast shadow motion, and that it ignores even clear evidence to the contrary.

## 6.0 General Discussion

---

The above results clearly show that cast shadows can be a powerful source of information for the perception of spatial layout. The importance of cast shadows for human vision, however, stands in contrast to some standard assumptions about how vision functions. If vision's primary jobs are to determine the identity and spatial layout of surfaces and objects, one could argue that variation of intensity in the image due to illumination should be discounted early in the visual system. With respect to variations in mean light level, this is certainly consistent with the function of retinal adaptation. The idea that the visual system also discounts slow spatial variations in illumination in order to determine surface color has been discussed since Helmholtz. It can be accomplished through lateral inhibition, and is the basis of a number of lightness algorithms. However, we can already see a complication in this line of reasoning in that slow variations in shading can arise due to either illumination or a smooth change in surface shape (Knill & Kersten, 1991a). Proceeding nonetheless, a logical progression is to assume that cast shadows also be discounted early in visual processing. Consistent with this strategy, work in computer vision has typically treated shadows as noise, to be filtered out during the first steps of edge detection (for exceptions, see Waltz, 1972; Shafer, 1985; Kender, J. R., & Smith, E. M., 1987). This approach, however, has not worked because of the well-known problems involved in extracting significant edges--it is too easy to confuse shadow boundaries with depth and material changes. The general difficulty of discounting shadows

early in visual processing raises the possibilities that human visual performance may either suffer because of it, or perception may in fact use shadow information not only in determining spatial layout, but also in other high-level tasks such as object recognition. It is not too hard to find examples of the problems shadows pose for human image understanding (e.g. the “stain” in Rembrandt’s “The Night Watch”, or the Mooney pictures of unfamiliar objects, Moore and Cavanagh, 1996). However, a couple of recent studies are also consistent with the latter possibility.

Tarr, Kersten & Bühlhoff (submitted) have obtained experimental results suggesting that cast shadows are not simply discounted in a visual recognition task (a same/different judgment of object identity) and can improve recognition speed. This speed gain, however, comes at a cost—a change of illumination direction slows down accurate same judgments when intrinsic cast shadows are present, but not in the absence of cast shadows. Rensink & Cavanagh (1993) reported evidence suggesting that cast shadows are rapidly identified in order to establish object structures, but then effectively removed so that they are difficult to access in a visual search task. These results and ours indicate that at some level the visual system must distinguish between shadows and other types of boundaries (e.g. shape or reflectance). Distinguishing whether an image edge is due to a cast shadow or some other cause, however, is not a simple computational problem. Solving this problem is even more complex when one realizes that identification of a static cast shadow boundary carries different information than a moving shadow. Static cast shadow boundaries (unlike specularities) can be treated as surface markings (e.g. potentially useful for stereopsis and, in cases when only the observer is moving, for structure-from-motion). In contrast, moving shadows are clearly not surface markings, and while they pose some of their own problems, we will see below how motion may provide for

their solutions (Section 6.2). What information can be used to distinguish cast shadows from other intensity gradients for either static or moving shadows?

### **6.1 Computing shadow identity & local constraints**

There are local, highly ambiguous constraints for computing shadow identity. These include the X-junction constraints of Metelli (1975), fuzziness of the penumbra, weak constancy of luminance variance and chromaticity values across shadow boundaries (see Funke-Lea, 1995). Unfortunately, none of these are robust enough to avoid confusing a shadow boundary with one caused by variations in shape, reflectance, or focus. Shadow motion may provide more reliable information for shadow identity, through averaging of the above local cues, or by providing more diagnostic information in some instances. An example of the latter is the dynamically changing penumbral blur of an extended light source, which is less likely to be confused with a material change. Nevertheless, local ambiguity remains and in general, the form and evolution of optic flow is influenced by changes in the relative positions of objects and illumination as well as viewpoint. The effects of illumination in the image can be quite distant from the image of the object causing it. A change in shape causes a local change in image intensity, whereas a cast shadow can be caused by an object either unseen or distant from its effect in the image--the effects of illumination are not just local.

### **6.2 Global Constraints**

Given the ambiguity of local cues, how are shadows identified? It would seem that identification is a prerequisite to appropriately linking a shadow with its casting object, but the identification cannot be made uniquely based on local cues alone. We can gain some insight into this problem by contrasting the results with the stationary square-over-checkerboard (as in Experiment 1) with the ball-in-a-box. We have seen that apparent motion of the stationary square is sensitive to the specifics of shadow properties. Observers are more likely to experience appar-

ent depth changes for dark shadows than light, and with fuzzy rather than sharp shadows (Experiment 1; Kersten et al., 1996). In contrast, we have found that the strength of the ball-in-a-box illusion persists under a wide variety of manipulations of the shadow. In the ball-in-a-box simulation of Experiment 3, we found that an object's cast shadow does not have to be physically reasonable--it can have the wrong contrast polarity or lightness--for observers to consistently see different motions in depth which depend on shadow trajectory<sup>5</sup>. These observations also stand in contrast to those obtained for the role of shadows in shape judgments in static images, which show that manipulations of shadow brightness and contrast interfere with shape perception (Cavanagh & Leclerc, 1989). Why do the ball-in-a-box demonstrations produce a strong percept of motion in depth, even when several properties of shadows such as contrast polarity, transparency, and correct shape are wrong?

The robustness of the ball-in-a-box illusions may in fact be a consequence of perception's ability to use global constraints which are needed to cope with the complexity and ambiguity of natural viewing. Specifically, the robustness of the percept may owe in part to the fact that dynamic displays contain an important piece of information not available in static displays--the correlation between the motion of an object and its cast shadow. The detection of a particular kind of correlated motion could indicate an event category corresponding to an object/shadow pair. The informativeness of correlated motion as a cue depends on at least two assumptions: 1) a stationary light source, and 2) a general viewpoint. A stationary light source constrains the relative image positions of an object and its shadow to be along a line connecting the shadow, object and light source and thus defines the correlated motion. If the light

---

5. We have replaced the shadow with a square transparent surface, as well as a thick coin-like object. Informal observations suggest that as long as the shadow substitute appears to be on the floor of the box, the apparent ball trajectories are similar to those seen with a natural shadow.

source is at infinity, the line makes a fixed angle in the image, thus an object and its shadow, while changing in relative distance during motion, are constrained to maintain the same relative angle. If the light source is at a finite position, the line sweeps through an angle anchored to a fixed location in the image plane.

The assumption of a general viewpoint says that the scene is viewed from a generic, rather than special or accidental viewpoint<sup>6</sup>. The idea that the visual system assumes a general viewpoint has been shown to be useful in explaining why humans can make some categorical perceptual inferences as reliably as they do (Nakayama & Shimojo, 1992). For the case of correlated motion, it is clear that the type of correlated motion we have described for objects and their shadows is one which would appear in just about any view of the object (assuming both object and shadow were visible). It is also clear that such motions do not normally arise from independent motions of objects in the world. Even for objects whose motions are causally related (e.g. the ends of a rigid baton moving arbitrarily in 3D space), the occurrence of the specific type of correlated motion we have described would be rare, and found only under a limited range of viewpoints. This particular motion path, then, is a reliable indicator that an

---

6. To summarize the basic idea, suppose an image feature is detected (e.g. correlated motion such that two image regions move constrained by a line of fixed orientation) and one wants to use the information provided by that feature to infer something about the world, say state “A” (e.g. object/shadow hypothesis). Assumption of a general viewpoint allows one to make such an inference reliably when the proportion of views giving rise to the feature given one state of the world (state “A”) is much greater than the proportion of views giving rise to that feature for other states of the world (“not A”) (Jepson, Richards & Knill, 1996). Expressed in probabilistic terms, we require:

$$p(\text{feature}|A) \gg p(\text{feature}|\text{not}A)$$

When this condition holds, we can say that detection of a feature (correlated motion) reliably implies the existence of state A of the world. As it turns out, the informativeness of a feature relies on a weaker version of this rule: the *a priori* probability of the world state must have a significant non-zero probability (Knill and Kersten, 1991b; Richards and Jepson, 1992). In the terms of Richards and Jepson (1992), correlated object and shadow image motion would be a “key” feature for labelling two image regions as being matched object/shadow pairs.

object/shadow pair is present; though, by itself, it does not indicate which image patch corresponds to the object and which to the shadow, nor does it specify where the shadow is. Shadow identity could be resolved by assuming that the light source is above, and thus the lower region is the shadow. But we still require knowledge of the shadow location to compute a unique depth trajectory.

Consider the geometric ambiguities present in the ball-in-a-box animations (Figure 13). To determine the ball's location along the line of sight, the visual system requires three pieces of information: 1) the location of the shadow, 2) the direction of the light source, and 3) the viewpoint. The first two pieces of information constrain the ball to be on a line between the shadow and light source. The third specifies a line from the eye through the ball. The intersection of these two lines determines the ball's position along the line of sight. We have seen in Experiment 3 that the visual system assumes a fixed, if incorrect light source position, but what information determines the location of the shadow? Local photometric constraints could con-

---

Figure 13 about here

---

tribute to labelling a region as a shadow, which is by necessity on the receiving surface. But in Experiment 3, we found that opaque white shadows, which are photometrically wrong, had no significant effect on the depth illusion. Another source of information is the non-accidental alignment of the canonical axis of the shadow patch with that of the floor -- a perpendicular through the floor coincides with a perpendicular through the shadow<sup>7</sup>. For example, a slant estimate based on the aspect ratio of the images of a checkerboard square would match that of

---

7. See Richards, Jepson, & Feldman, J. (1996) for a discussion of the theoretical basis for this kind of inference. A second source of information would be the assumption that in the absence of evidence to the contrary, an object (ersatz shadow) should have support.

the elliptical shadow patch. An economical explanation for this coincidence is that the opaque white shadow and floor are coplanar<sup>8</sup>.

### 6.3 Psychophysics of constraints

The previous section underscored the need for non-local computations to integrate cast shadow motion with object motion. An example of global consistency checking in the box world is the classic work on the utilization of static shadow contour information by Waltz (1972). But virtually all biologically motivated computational models of depth perception (e.g. stereo and motion) rely on local computations. The kind of brain computation required to support the perceptual processing we have described here resembles a more global process in which the visual system seeks a logical and probable interpretation of the image based on a knowledge of how images could be formed from objects, their spatial relations, the illumination, and viewpoint together with the prior assumptions about the nature of the world (Gregory, 1970; Rock, 1983; Kersten, 1990). Assuming such a framework for visual system processing suggests a program of psychophysics which we refer to as a “psychophysics of constraints” (Knill, Kersten & Yuille, 1996). The objects of experimental study become the nature of the image features used for perception of scene characteristics, the constraints assumed by the visual system on how such features are generated from real scenes and the prior constraints assumed on the values of scene characteristics.

The discussion of Section 6.2 suggests that one can, with a set of quite general prior assumptions about the world and the assumption of a general viewpoint, build up a network of infer-

---

8. One prediction of this analysis is that ambiguity of the shadow location should reduce the influence of a change in shadow trajectory, if the shadow is replaced by an object whose location is ambiguous. Although we haven't formally tested this prediction, we have observed that replacing the shadow by another ball tends to reduce the apparent trajectory difference between the horizontal and diagonal conditions of Demonstration 1. The information to place this second ball, ersatz shadow, on the floor is weaker than for a patch whose orientation matches the floor.



ences about the world which one can reliably make from image data, along with a list of the image features on which they are based. Jepson, Richards and Knill (1996) have proposed a general framework for modeling this type of inference and have applied it to modeling the qualitative inferences one can make about object motion. They refer to the qualitative inferences as motion events (e.g. stationary, resting on ground, etc.). Including shadow information into their framework would seem straightforward, and including relative object and shadow motion within their system of motion events is a natural direction to take in modeling the visual interpretation of cast shadow motion. This paper has presented the beginning of such a program of research to the perception of 3D spatial layout and motion from cast shadow information.

## Acknowledgments

---

This work was supported by the National Science Foundation (BNS-9109514), the Max Planck Society and by the Air Force Office for Scientific Research (AFOSR 90-2074) and NIH (EY09383-01A1). We thank Albert Yonas, Isabelle Bühlhoff, and Deborah Rossen for their comments and suggestions. We also especially thank Heinrich Bühlhoff for providing an environment conducive to the development and expansion of the ideas in this paper. A number of the results described in this paper were first reported at the annual meeting of the Association for Research in Vision and Ophthalmology (Kersten, Mamassian & Knill, 1991; Mamassian, Kersten & Knill, 1992). A brief description of some of the results of Experiments 1, 3 and 4 have been described by Knill, Kersten and Mamassian (1996).

## References

---

Arnheim, R. (1954). *Art and Visual Perception*. Berkeley: University of California Press.

Bülthoff, I., Kersten, D., & Bülthoff, H. H. (1994). General lighting can overcome accidental viewing. (Abstract) Association for Research in Vision and Ophthalmology. Sarasota, Florida.

Cavanagh, P., & Leclerc, Y. G. (1989). Shape from shadows. *Journal of Experimental Psychology, Human Perception and Performance*, **15**, 3-27.

Cutting, J. E., & Vishton, P. M. (1995). Perceiving Layout and Knowing Distances: The Integration, Relative Potency, and Contextual Use of Different Information about Depth, Perception of Space and Motion, (pp. 69-117). *Perception of Space and Motion*, Academic Press, Inc.

Dosher, B. A., Sperling, G., & Wurst, S. (1986). Trade-offs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure. *Vision Research*, **26**, 973-990.

Funka-Lea, G. D. (1995). *The Visual Recognition of Shadows by an Active Observer* (MS-CIS-95-11, GRASP LAB 389). Computer and Information Science, University of Pennsylvania.

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

Gregory, R. L. (1970). *The Intelligent Eye*. New York: McGraw-Hill Paperbacks.

Gombrich, E. H. (1995) *Shadows: The Depiction of Cast Shadows in Western art*.

Jepson, A., Richards, W. & Knill, D. (1996). Modal structure and reliable inference. In K. D.C., & R. W. (Ed.), *Perception as Bayesian Inference*. Cambridge, UK: Cambridge University Press.

Kender, J. R., & Smith, E. M. (1987). Shape from darkness: deriving surface information from dynamic shadows. *Proceedings of the First International Conference on Computer Vision*. London, UK. 539-546.

Kersten, D. (1990). Statistical limits to image understanding. In C. Blakemore (Ed.), *Vision: Coding and Efficiency*. Cambridge: Cambridge University Press.

Kersten, D. J., Mamassian, P., & Knill, D. (1991). Moving cast shadows generate illusory object trajectories. (Abstract). *Investigative Ophthalmology and Visual Science*, **32**, 1179.

Kersten, D., Knill, D.C., Mamassian, P. and Bühlhoff, I. (1996) Illusory motion from shadows. *Nature*, **379**, 31.

Knill, D. C., & Kersten, D. (1991a). Apparent surface curvature affects lightness perception. *Nature*, **351**, 228-230.

Knill, D. C. & Kersten, D. (1991b). Ideal Perceptual Observers for Computation, Psychophysics, and Neural Networks. In R. J. Watt (Ed.), *Pattern Recognition by Man and Machine*. MacMillan Press.

Knill, D. C., Kersten, D., & Yuille, A. (1996). Introduction: A Bayesian Formulation of Visual Perception. In Knill D.C. & Richards W. (Eds.), *Perception as Bayesian Inference* (Chap. 0). Cambridge University Press.

Knill, D. C., Kersten, D., & Mamassian, P. (1996). Implications of a Bayesian Formation. In Knill D.C. & Richards W. (Eds.), *Perception as Bayesian Inference* (Chap. 6). Cambridge University Press.

Knill, D. C., Mamassian, P., & Kersten, D. (1993). The geometry of shadows. Computer and Information Sciences Technical Report (**TR 93-47**). University of Minnesota.

Lowe, D. G. (1985). *Perceptual Organization and Visual Recognition*. Boston, MA: Kluwer Academic Publishers.

Mamassian, P., Knill, D.C. and Kersten, D.J. (1992). Interaction of cast shadow and change of size. *Perception*, **21**, Suppl. 2, 48. (Abstract). ECVF 1992, Pisa, Italy.

Mamassian, P., Kersten, D.J. and Knill, D.C. (1992). Spatial layout from cast shadows. *Investigative Ophthalmology & Visual Science*, **33**, 1332. (Abstract). Paper presented at ARVO 1992, Sarasota, Florida.

Metelli, F. (1975). Shadows without penumbra. In S. Ertel, L. Kemmler, & L. Stadler (Ed.), *Gestaltentheorie in der modernen psychologie* (pp. 200-209). Darmstadt: Dietrich Steinkopff.

Moore, C. and Cavanagh, P. (1996) Discriminating shadow and object regions in 2-tone images, *Investigative Ophthalmology & Visual Science (Supplemental)*, **37** (3), S193.

Nakayama, K., & Shimojo, S. (1992). Experiencing and perceiving visual surfaces. *Science*, **257**, 1357-1363.

Ramachandran, V. S. (1988). Perception of shape from shading. *Nature*, **331**, 163-166.

Rensink, R. & Cavanagh, P. (1993). Processing of shadows at preattentive levels. (Abstract)  
*Investigative Ophthalmology and Visual Science*, **34**, 1288.

Richards, W., Jepson, A., & Feldman, J. (1996). Priors, preferences and categorical percepts.  
In D. C. Knill & W. W. Richards (Eds.), *Perception as Bayesian Inference*. Cambridge, U.K.:  
Cambridge University Press.

Richards, W. & Jepson, A. (1992). What makes a good feature? (C.B.I.P. Paper 72). M.I.T.,  
Cambridge, Massachusetts.

Rock, I. (1983). *The Logic of Perception*. Cambridge, Massachusetts: M.I.T. Press.

Shafer, S. A. (1985). *Shadows and Silhouettes in Computer Vision*. Boston, Massachusetts:  
Kluwer Academic Publishers.

Tarr, M., Kersten, D., & Bühlhoff, H. H. (manuscript). Why the visual recognition system  
might encode the effects of illumination.

da Vinci, L. (1970). *Notebooks of Leonardo Da Vinci*. New York: Dover Publications, Inc.

Waltz, D. L. (1972). Understanding line drawings of scenes with shadows. In P. Winston  
(Ed.), *The Psychology of Computer Vision* New York: McGraw-Hill.

Wanger, L. R., Ferwerda, J. A. & Greenberg, D. P. (1992). Perceiving spatial relationships in  
computer-generated images. *IEEE Computer Graphics & Applications*, 11-58.

Yonas, A., Goldsmith L.T. & Hallstrom J. L. (1978). Development of sensitivity to informa-  
tion provided by cast shadows in pictures. *Perception*, **7**, 333-341.

## Figure Captions

---

Figure 1. Increasing the displacement between the cast shadows and the three foreground squares produces an impression of increasing depth (from left to right) relative to the background checkerboard. In this figure, the penumbra is fuzzy and is the same for all three shadows. In all of the experiments and demonstrations described in the rest of this paper, the shadows are sharp, with no penumbræ.

Figure 2. The square-over-checkerboard stimulus used in Experiment 1. Observers fixated the cross and judged whether the square gray patch appeared to move in depth when only its shadow moved. The shadow moved diagonally down and to the left and then back towards the square patch as shown by the arrow.

Figure 3. Three frames from animations made with the ball-in-a-box simulations for Demonstration 1. Panel **a** shows the left-most positions of the ball and shadow, for the first frame, in both horizontal and vertical shadow trajectory animations. Panel **b** shows the right-most positions for the diagonal shadow trajectory, and panel **c** shows the right-most position for the horizontal shadow trajectory, both for the last frames. The ball appears to rise above the box floor for the horizontal trajectory, but to remain on the floor for the diagonal trajectory. Experiments 2, 3 and 4 used the same basic arrangement.

Figure 4a & b. Panel 4a shows a schematic time-lapse diagram of four frames from the animation used for Demonstration 2 (the non-linear motion). The shadow trajectory was artificially moved, leaving all other properties of the image, including the brightness of the ball and box, fixed. Observers reported the ball appearing to bounce at the third position from the left shown in the diagram. Panel 4b illustrates how shadow trajectories were varied in Experiments 2, 3, and 4. See the text for the actual values of the trajectory slopes used in each experiment. Solid

arrows indicate the trajectory of the ball (constant in all the animations), and dashed arrows indicate the trajectories of its shadow.

Figure 5. Three frames from animations made with the longer ball-in-a-box for Demonstration 2. Panel **a** shows the starting position. As in Fig. 3, panels **b** & **c** show the extreme right position of the ball for the horizontal and diagonal shadow trajectories, respectively. In these static images, the effect of the shadow on the apparent size of the ball is small, but noticeable. In the dynamic case with diagonal trajectory, the ball has the striking appearance of inflating as it moves from left to right. For the horizontal trajectory, the ball appears to remain the same size.

Figure 6. A side view of the box used in Experiment 2. The observer adjusted the height of a horizontal bar to the perceived height of the center of the ball relative to the bottom of the box (the cursor is shown in Figure 3b&c). The two cues were the position cast by the ball on the bottom of the box, and the changing size of the ball's image as the ball was moving in the box. (Drawing not to scale).

Figure 7. Perceived height as a function of object size change and shadow slope for four observers in Experiment 2. (a) Accumulation of the two cues. (b) Non-linear interaction of the two cues. (c) Significant interaction but no main effect of size change. (d) Size change vetoed by cast shadow cue.

Figure 8. The displacement  $\Delta S$  between an object and its shadow can be produced either by a change in light source position,  $\Delta L$  or by a change in depth of the object,  $\Delta D$ .

Figure 9. The top (a) and bottom (b) panels show the apparent ball height settings as a function of opacity for shadow contrasts of -1 and +1, respectively for observer GD in Experiment 3. There is no measurable effect of opacity for either contrast condition. Unnatural white



“shadows” (negative opacity), and reverse contrast “shadows” (negative contrast) were as effective as natural shadows. For a contrast of +1, and opacity of 0, the shadow is invisible, and the settings converge at an intermediate height. Error bars are  $\pm 1$  S.E..

Figure 10. Perceived height above the checkerboard floor of the ball, in the coordinates of the 3D simulated world, as a function of the shadow slope for Experiment 3. Data are shown for three subjects. Each point is the mean of 8 measurements. Error bars indicate  $\pm 1$  S.E. of the mean. As the shadow's trajectory slope goes from zero (horizontal) to one (identical to ball), the apparent peak height of the ball falls. The solid line shows the physically correct setting based on the light source direction used to render the scene. The dashed lines show fits to the data for a model in which it is assumed that each subject bases his or her estimate of the object trajectory on an implicit fixed light source position.

Figure 11. Perceived height of the ball above the checkerboard floor as a function of shadow slope and light source direction for Experiment 4. Dotted lines show results for a theoretical observer which accurately estimates the light source directions and uses these to calculate 3D object motion from cast shadow motion. The mean height estimates for the three groups of subjects are shown by the open symbols connected by solid lines. Subjects' mean response curves cluster around what would be predicted for a single intermediate, but inaccurate light source position. Error bars indicate  $\pm 1$  S.E. of the mean.

Figure 12. The bottom panel shows a time-lapse diagram of the approximate shadow positions for Demonstration 7. The top panels show frames 1, 7, 15, and 30 of a 30-frame sequence in which there is evidence from the shading on the ellipsoid and shadows cast by the vertical blocks that the illumination direction is changing as the ball moves from left to right. The shadow trajectory is similar to the non-linear one described in Figure 4a (Demonstration 2),

except that the shadow trajectory was changed by moving the light source. If the visual system could accurately take the changing lighting information into account, it would conclude that the ellipsoid is moving along a linear trajectory in the fronto-parallel plane; observers do not see this, rather the percept is of an ellipsoid starting near the observer (frame 1), moving back in depth (frame 15), and then towards the observer again (frame 30).

Figure 13. The position of the ball along the line of sight can be determined if the shadow location and light source direction are known. However, the shadow location is ambiguous. It could be a transparent surface floating above the floor, or even an opaque surface behind a transparent floor. If the shadow patch is identified as such, it should by necessity be on the receiving surface. Shadow identification may be achieved through a combination of local cues, as well as the global cue of correlated motion. Another constraint to place the shadow on the floor is the coincidence of the orientation of the shadow patch and the floor.



Figure 1

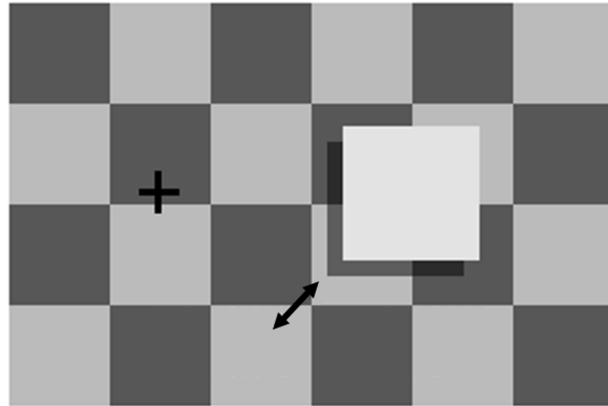


Figure 2

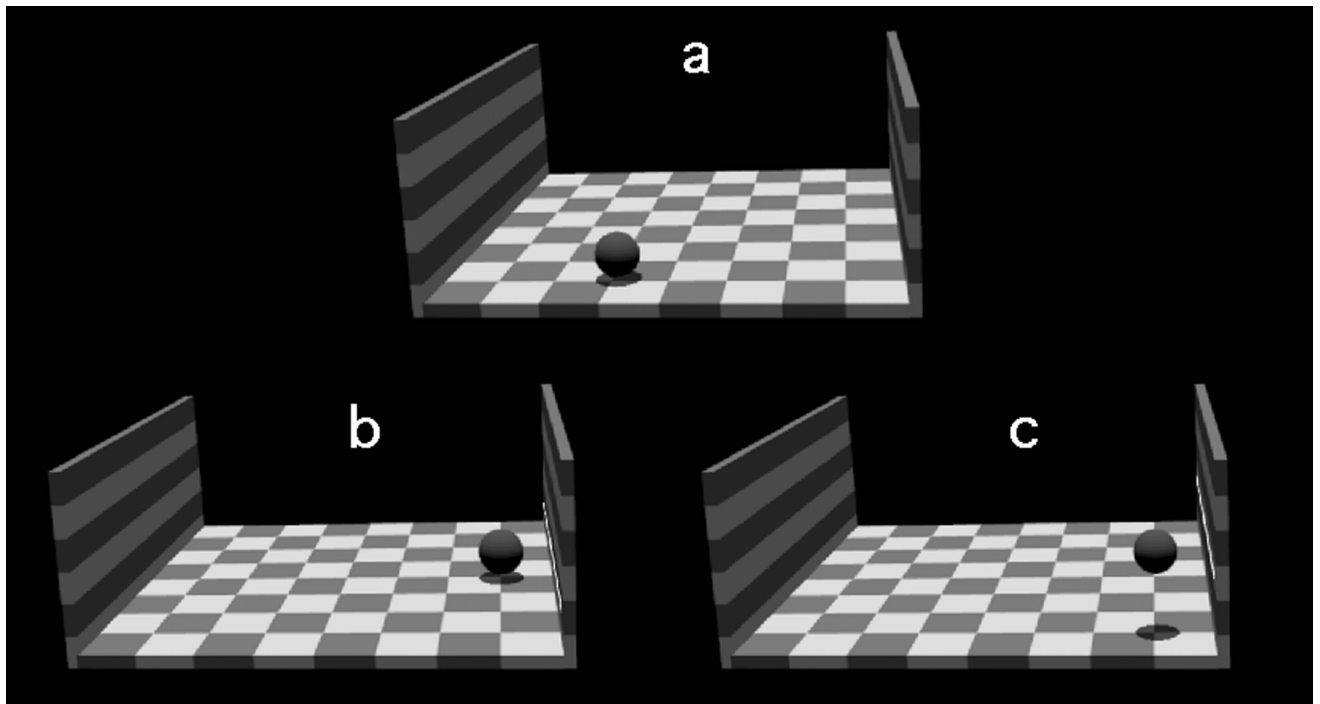


Figure 3

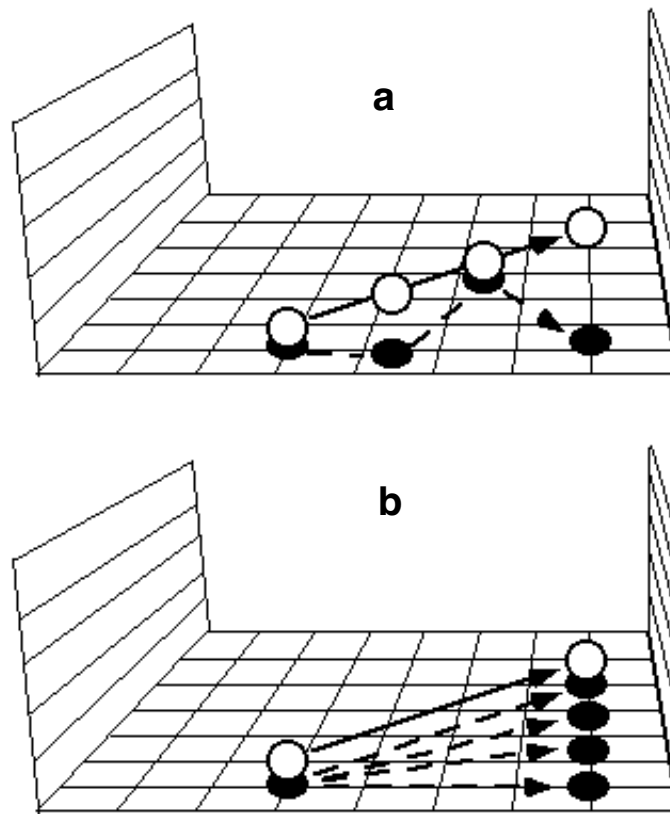


Figure 4

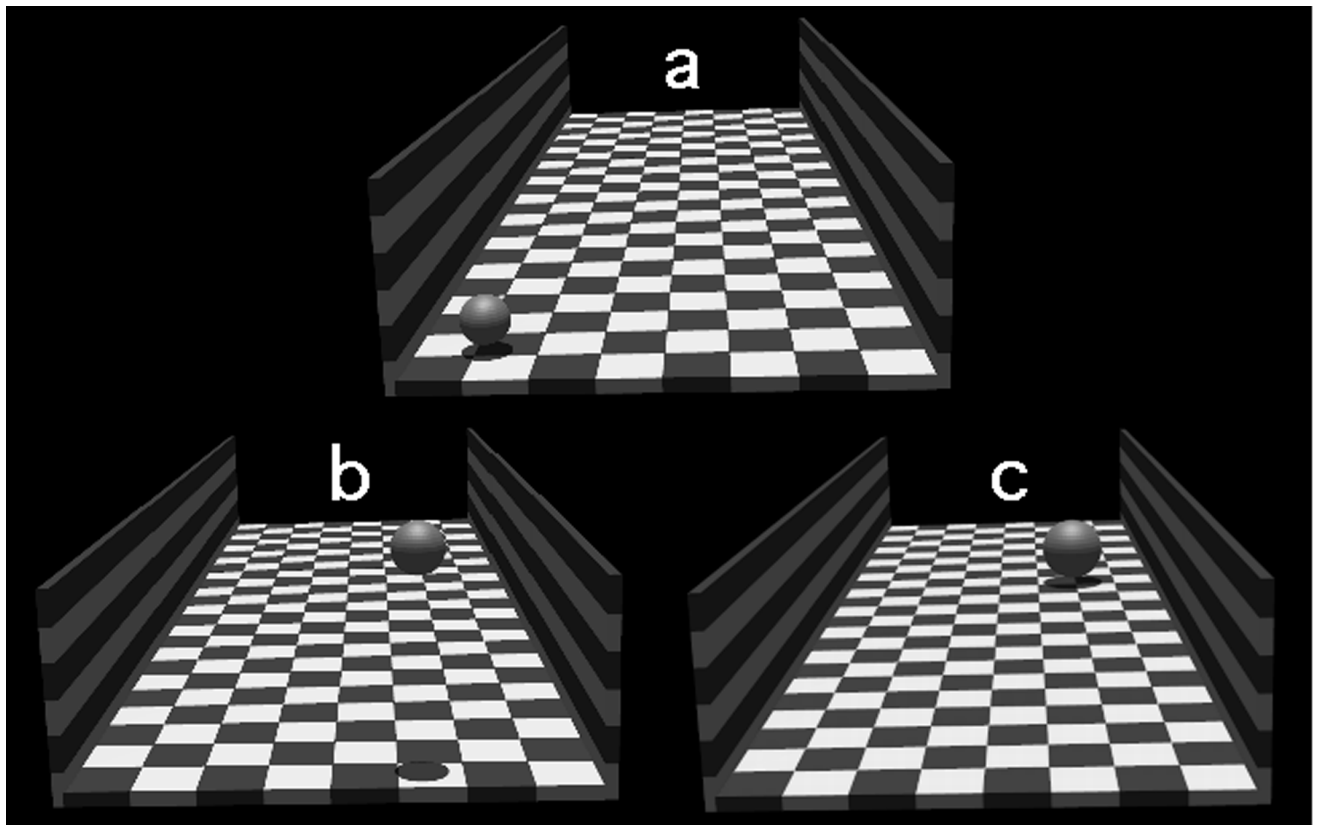


Figure 5

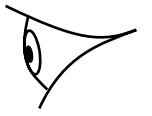
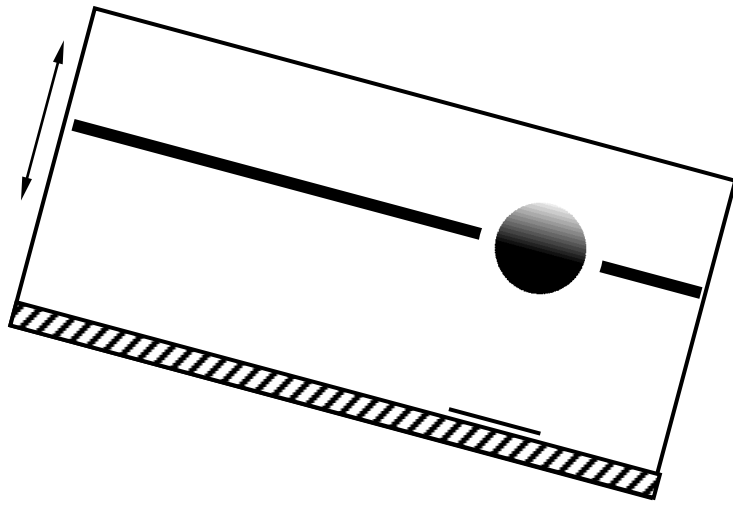


Figure 6

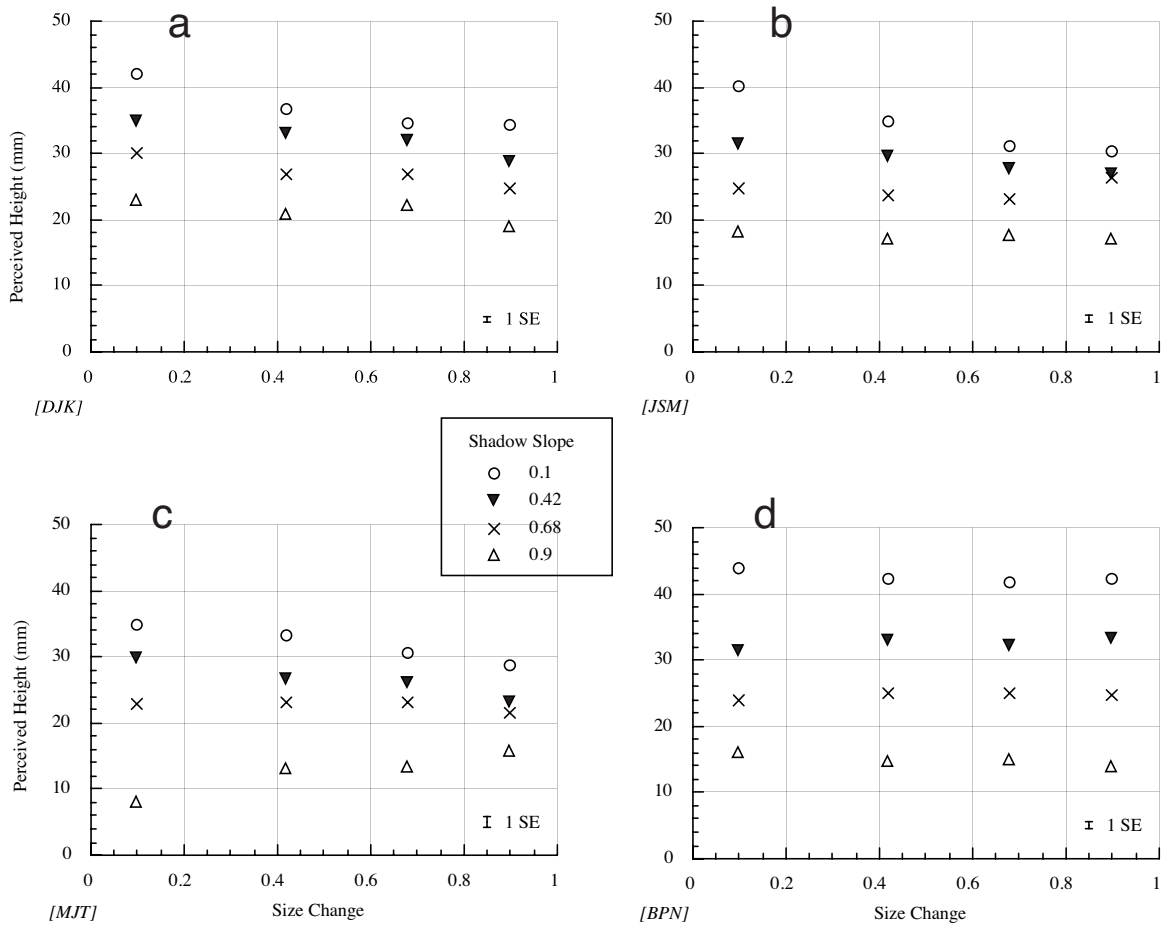


Figure 7

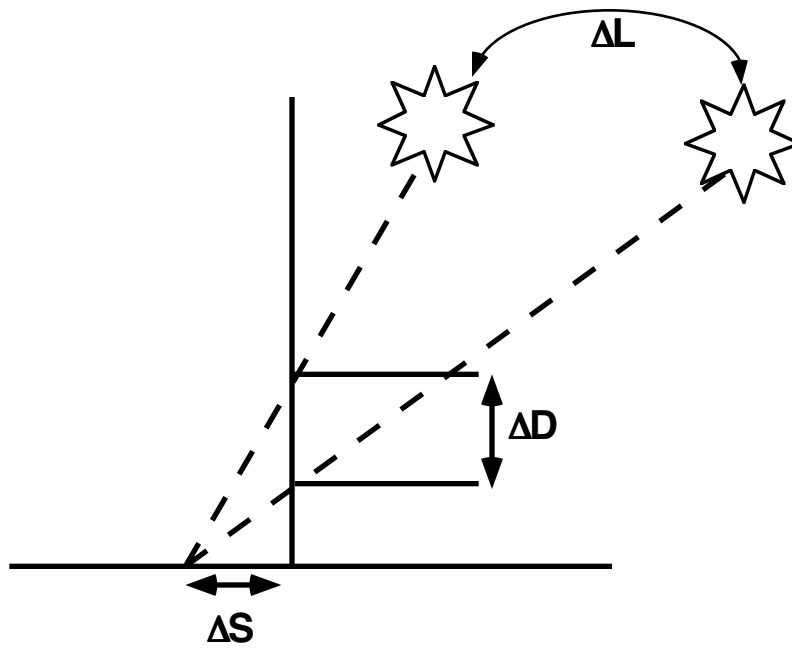


Figure 8

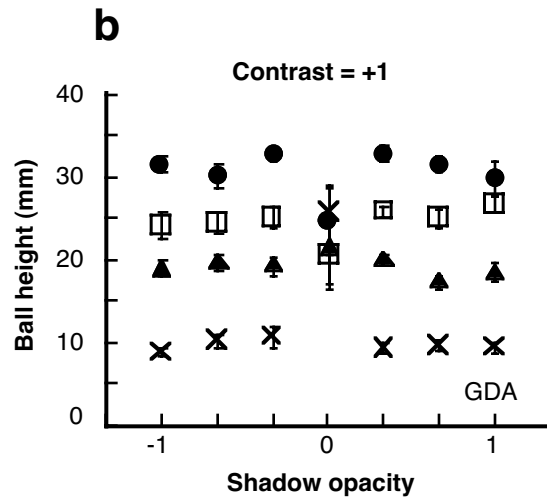
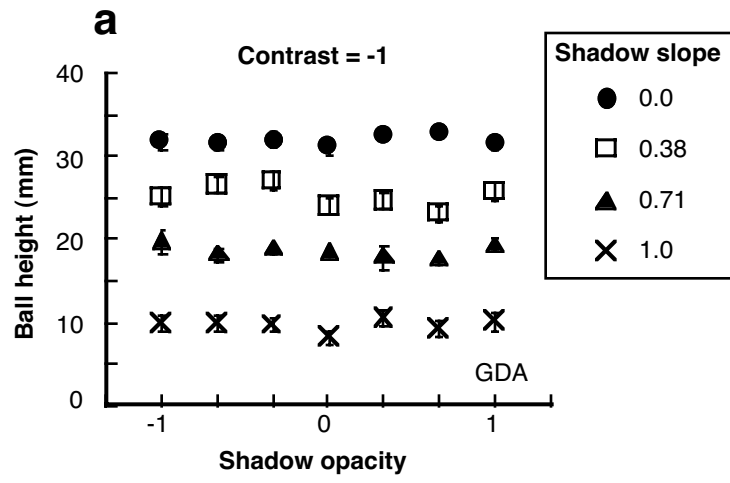


Figure 9

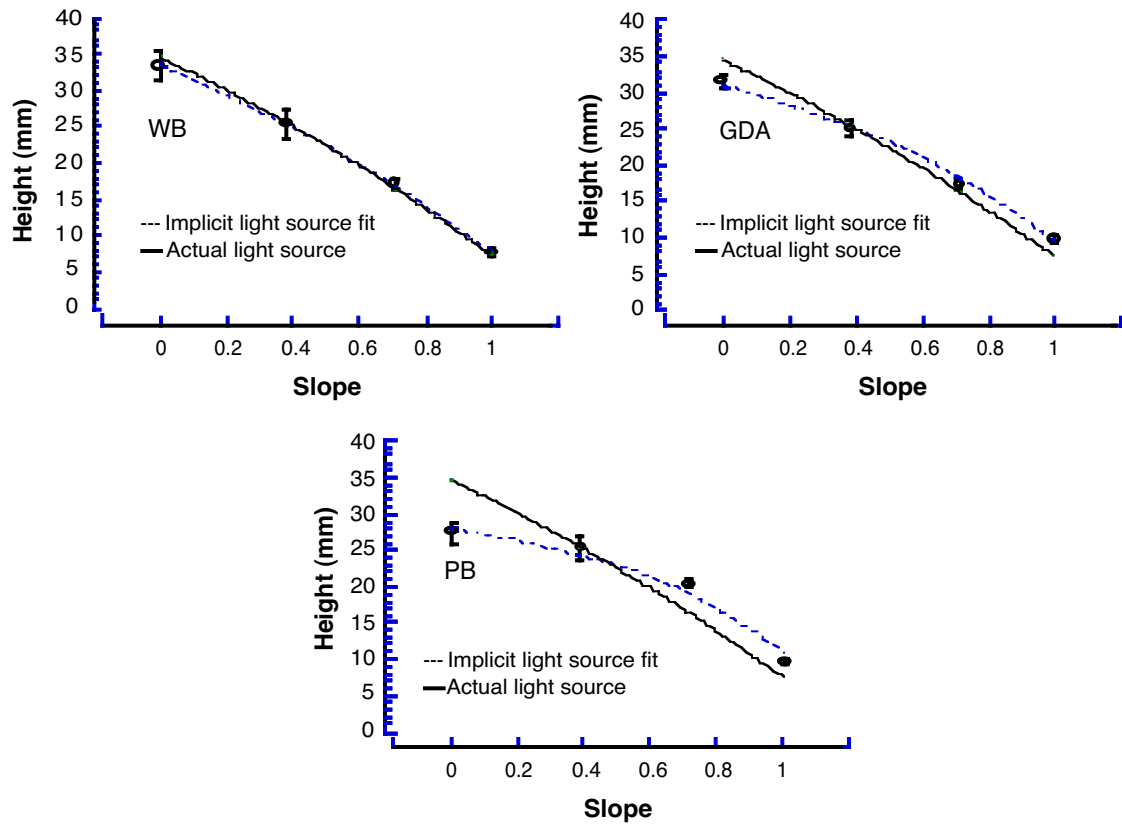


Figure 10



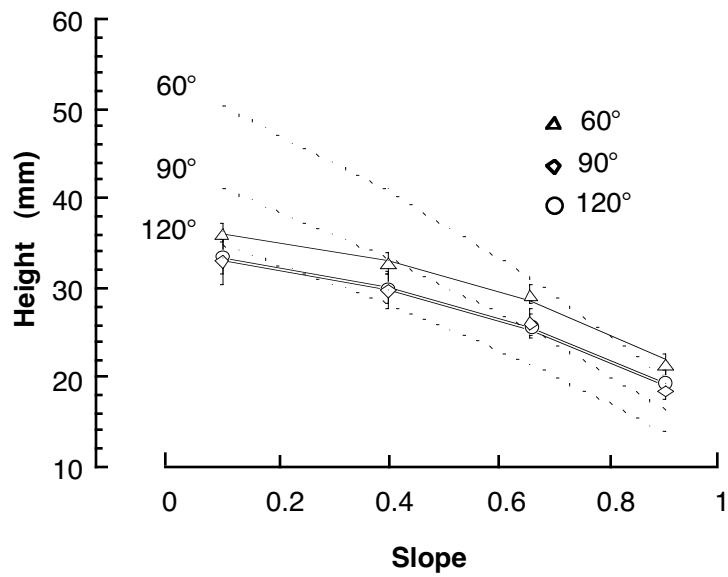


Figure 11

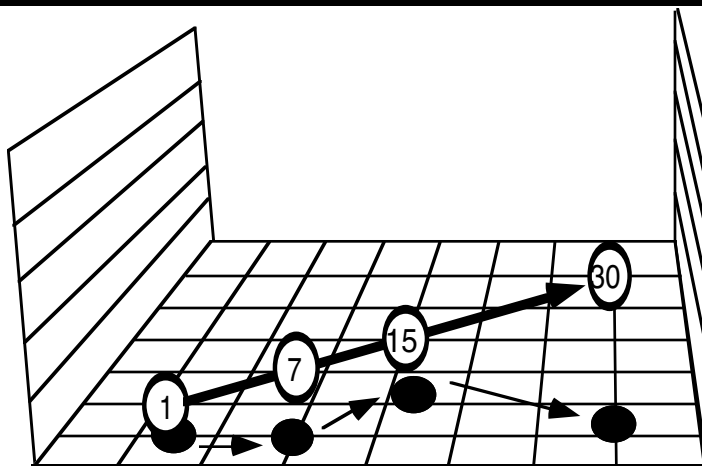
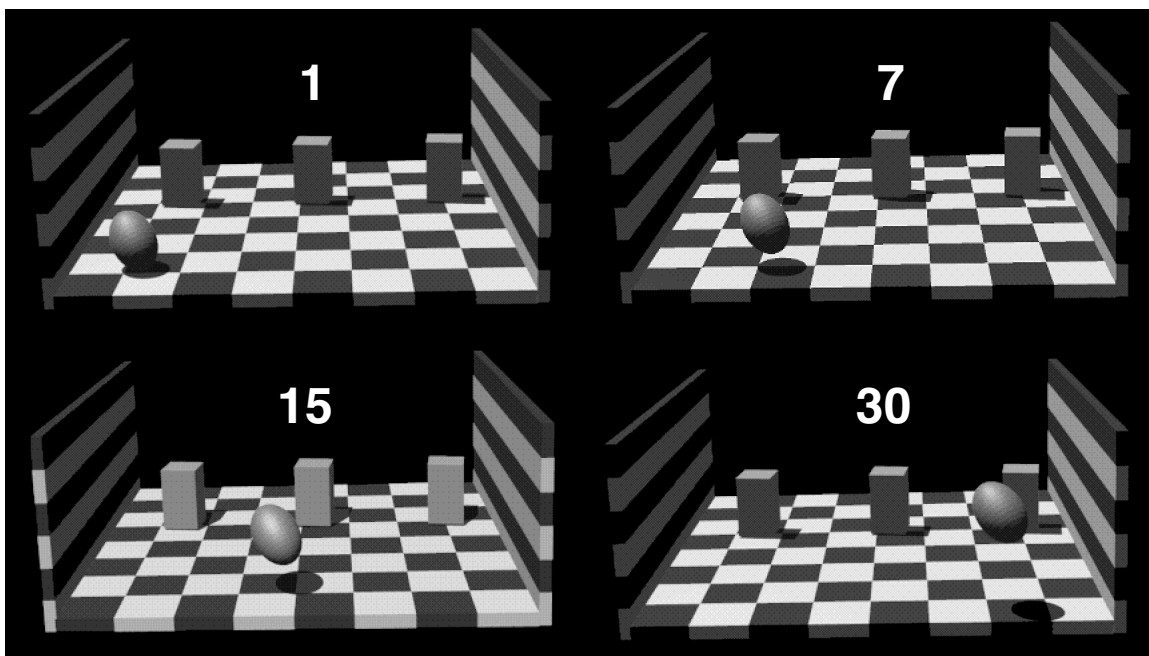


Figure 12

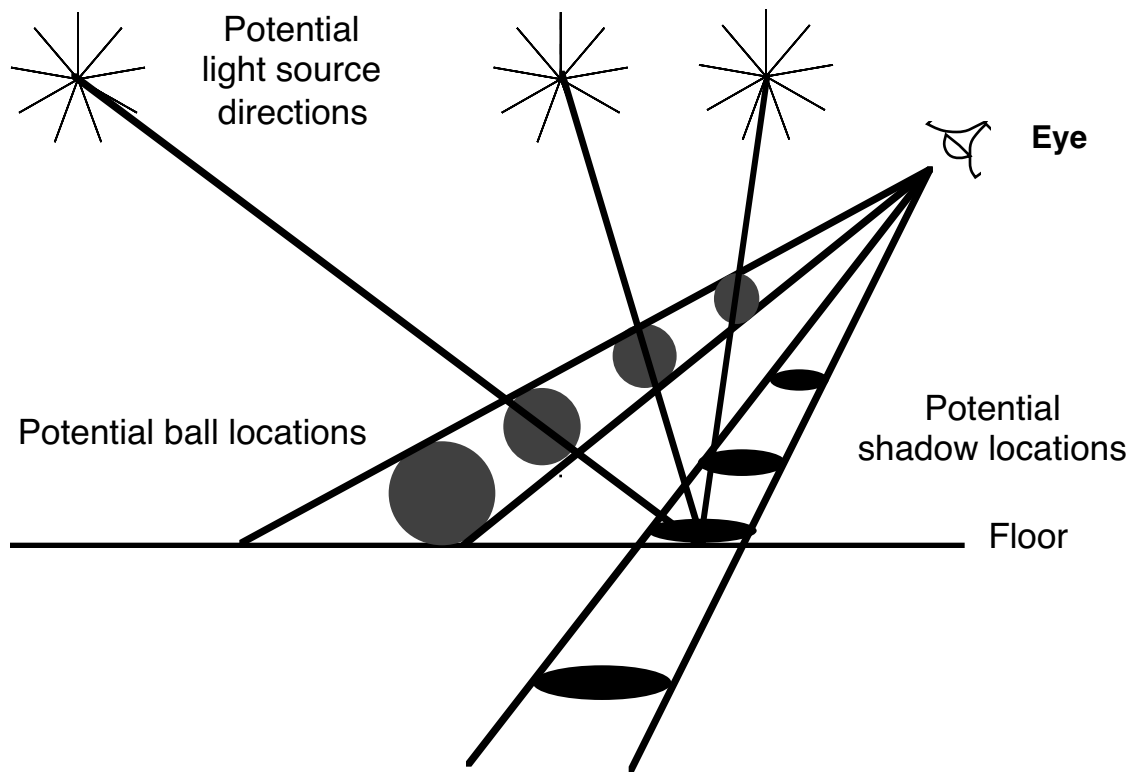


Figure 13