

Moving Cast Shadows and the Perception of Relative Depth

Daniel Kersten, Pascal Mamassian & David C. Knill

Abstract

We describe a number of visual illusions of motion in depth in which the motion of an object's cast shadow determines the perceived 3D motion of the object. The illusory percepts are phenomenally very strong. We analyze the information which cast shadow motion provides for the inference of 3D object motion and experimentally measure human observers' use of this information. The experimental results show that cast shadow information overrides a number of other strong perceptual constraints, including viewers' assumptions of constant object size and a general viewpoint. Moreover, they support the hypothesis that the human visual system incorporates a stationary light source constraint in the perceptual processing of shadow motion. The system imposes the constraint even when image information suggests a moving light source.

DK and PM were supported by the National Science Foundation (BNS-9109514) and the Max Planck Society. DCK was supported by the Air Force Office for Scientific Research (AFOSR 90-2074) and NIH (EY09383-01A1). We thank Albert Yonas and Deborah Rossen for their comments and suggestions. Correspondence should be sent to: Daniel Kersten, N218 Elliott Hall, Psychology Department, 75 East River Road, Minneapolis, MN 55455, U.S.A.. Email: kersten@eye.psych.umn.edu.

This document is available as /pub/mpi-memos/TR-6.ps.Z via anonymous ftp from ftp.mpik-tueb.mpg.de or by writing to the Max-Planck-Institut für biologische Kybernetik, Spemannstr. 38, 72076 Tübingen, Germany.

1.0 Introduction

The relative displacement between an object and its cast shadow in an image provides an important source of visual information about the spatial layout of objects. Leonardo da Vinci elucidated the principle relating shadow displacement and the perception of relative depth in his notebooks: "...when representing objects above the eye and on one side--if you wish them to look 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) Artists regularly exploit this principle in static drawings and paintings of 3D scenes, and psychophysical research has shown the salience of static cast shadow information for judgments of depth (Yonas, 1978). Yonas et al. (1978) were able to show that the location of a cast shadow was able to influence the judged depth and height of an object above a ground plane in observers as young as three years old. The role of dynamic shadows in human perception, however, has received no scientific study. Because movement due to shadow boundaries is almost always present in the retinal image, understanding how the visual system processes shadow motion is a fundamental issue in vision. In this paper, we report a set of controlled experiments and phenomenal demonstrations which show:

- the relative motions of objects and their cast shadows in an image can produce remarkably strong percepts of 3D motion
- 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
- image features such as shadow darkness can be utilized, but are not necessary for the perception of depth from moving cast shadows
- support for a prior assumption of a *stationary light source constraint* by the visual system.

2.0 The Phenomenon

2.1 Experiment 1: Cast shadow motion is sufficient for the perception of motion in depth.

The first question is whether shadow motion is in fact used for the perception of relative motion in depth. Although it is reasonable to assume that an affirmative answer would follow given the evidence from judging static shadows in pictures, it is not necessarily the case for at least three reasons. First, the fact that a pictorial cue is useful for judgments of depth does not necessarily imply that variations of that cue will produce

the perception of motion in depth. The reason is that judgments based on static cues with long viewing times can involve conscious reasoning as well as perceptual processing. Second, the computational problem of identifying shadows is known to be very difficult. The real-time requirements of identifying shadows in motion may be even harder. Although processing of static shadows has received some study in computer vision (Waltz, 1972; Shafer, 1985), with few exceptions (Kender, J. R., & Smith, E. M., 1987) computer vision has ignored moving cast shadows. Third, if vision's primary function is to determine the identity and spatial layout of surfaces and objects, one could argue that variation of intensity in the image due to illumination might be discounted early given the processing overhead required. A related argument that the visual system discounts variations in illumination in order to determine surface color has been discussed since Helmholtz.

The computational difficulty lies in the fact that optic flow is determined by a complex interaction of causes. The form and evolution of optic flow is influenced by changes in the viewpoint of the observer, positions and shapes of the objects, and the illumination. Unlike the effect of shape, the effect of illumination on the image is not just local. Shadow boundaries are determined by the illumination, the casting object, the receiving object and the viewpoint. Unfortunately, there is no unambiguous local cue for a shadow edge. Nevertheless for human shape perception, static cast shadow boundary is useful for object shape perception as well as depth perception (Cavanagh, & Leclerc, 1989). How are shadows identified? Cavanagh (1991) argues, based on work with images of faces, that the identification of shadow boundaries and utilization of shadow information may in fact *follow* the recognition of the category that a shape belongs to. From this point of view, it is not unreasonable to suppose that judgments involving static shadows may require processes that are too slow to be useful in processing dynamic shadows for depth information. Yet, moving cast shadows are used routinely in cartoon animations and in video games; but does this merely enhance the realism of the pictures, or is this information useful for depth?

Figure 1 illustrates the well-known effect of shadow displacement on the perception of relative depth in static images: the closer an object is to its cast shadow in an image, the closer it appears in depth to the background surface. We created a motion analog of this demonstration, in which the shadow cast by a stationary square moves back and forth relative to the square (figure 2). We then ran a simple psychophysical experiment (Experiment 1) to test whether subjects would see the square move in depth (see figure 2 caption for details). When the shadow was rendered real-

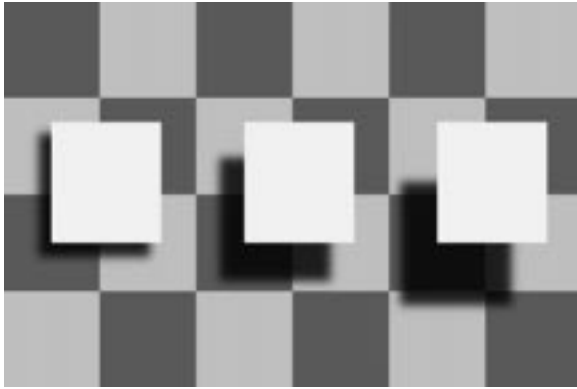


Fig. 1. Increasing the displacement between the cast shadows and the three foreground squares tends to produce an impression of increasing depth (from left to right) relative to the background checkerboard.

istically 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. Subjects who perceived the motion reported that the percept was phenomenally strong and immediate.

The result clearly shows an effect of cast shadow motion on observers' perception of 3D motion of an object. Moreover, a close analysis of the experimental stimuli reveals that for the observers who saw the motion in depth, the motion of the shadow overrode a number of 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 (related to object size constancy, cf. Gogel, Hartman, & Harker, 1957), and second, that the viewer is viewing the scene from a non-accidental, or general viewpoint (Biederman, 1985; Nakayama, & Shimojo, 1992). The assumption of object size constancy 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 size constancy constraint from an experiment, since the image size of an object is an inherent property of a stimulus. What is possible, however, 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.

2.2 Demonstration 1: 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. As in Experiment 1, the size of the object's

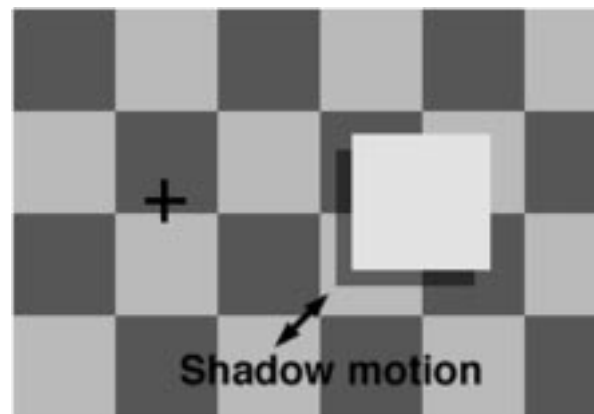


Fig. 2. Observers were asked to look at a fixation mark (+) placed on a checkerboard plane which subtended $6.6 \times 10^\circ$ of visual angle. Viewing distance was 500 mm. At a position 4.1° to the right of the fixation point, a foreground square was superimposed over a sharp shadow of the same size as the square. In a 500 msec. animated sequence, the shadow oscillated for one cycle through a 0.34° displacement from the foreground square. The foreground square remained stationary throughout the animation. Observers were asked to indicate whether the foreground square appeared to oscillate in depth or appeared to be stationary. 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^2 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. On 78% of the trials using dark shadows, observers reported seeing the foreground square as oscillating in depth--toward and away from the viewer. On only 40% of the trials using light shadows did subjects report seeing the square oscillating in depth (A Wilcoxon signed rank order test on the difference between light and dark shadows gave $p = 0.001$).

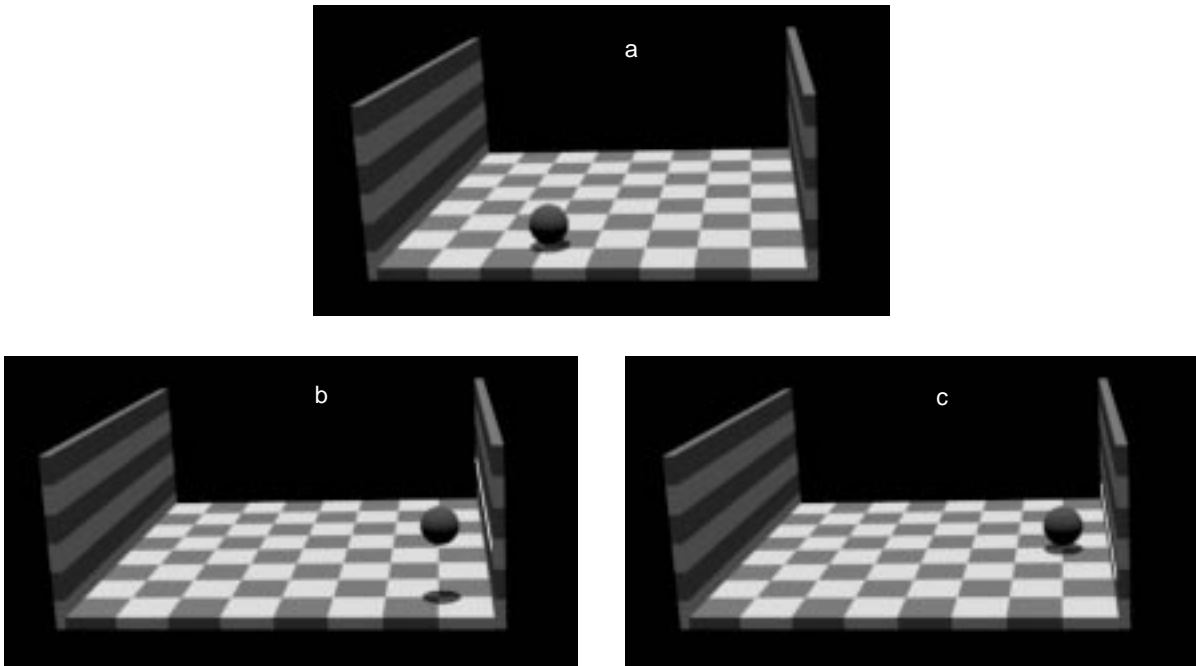


Fig. 3. Three frames from animations made with the ball-in-a-box simulation. In a simulated world, a ball was placed in a small 132 x 132 mm box and viewed from a point 355 mm from the center of the box with an elevation of 21.8° relative to the floor of the box. The viewpoint was offset slightly to the right, as shown. 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 and in the room for all the animations, except those used in Experiment 3, was generated by simulating a light source at infinity with a slant of 63.4° degrees relative to the floor of the box. In Experiment 3, we manipulated the shading on the ball as an independent variable. In all the animations, the ball moved in a linear trajectory in the image at an angle tilted by 21.8° from the horizontal. Its velocity 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 images shown here are copies of those used in the two animations for Demonstration 1. Figure 3a shows the left-most positions of the ball and shadow in both animations. Figure 3b shows the right-most positions in one of the animations and figure 3c shows the right-most position in the other. The demonstration animations were recorded on videotape, and observers were shown the taped animations. For the experiments (Experiments 2 and 3), however, the animations were shown on the screen of a Stardent GS2000 graphics computer. Subjects were given the task of adjusting a line along the right wall (shown in 3b and c) to match the apparent height of the middle of the ball at the right-most point of its trajectory. Subjects adjusted the height of the line by moving the computer's mouse and indicated a match by pressing the mouse button. The motion of the ball and its shadow continued throughout the course of a trial.

image, in this case that of a ball, remained fixed throughout the animation.

The first demonstration using this simulation (Demonstration 1) consisted of two different animation sequences: In the first, the ball's cast shadow followed a horizontal trajectory in the image (ending up at the position shown in figure 3b); in the second, it followed a diagonal trajectory identical to that of the ball's image (ending up at the position shown in figure 3c). 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 reported 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 in this experiment.

2.3 Demonstration 2: 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 2, everything was as with Demonstration 1, except that we tripled the length of the box in world coordinates (figure 4). 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. All of our observers reported seeing the ball apparently inflating and shrinking when the trajectory of the shadow matched the ball, but remaining fixed in size when the shadow trajectory was horizontal. In another study, we explicitly varied the image size of the ball together with the shadow trajectory slope and found a non-linear integration of the two sources of information in the perception of the relative position of the ball (Mamassian, Kersten, and Knill, 1992).

2.4 Demonstration 3: Moving cast shadow can produce the illusion of a non-linear object trajectory.

A third demonstration (Demonstration 3) further shows the sophistication of human 3D motion perception from relative shadow motion. We modified the animations used for Demonstration 1 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 5a). The ball's image moved in the same straight, diagonal trajectory as before. 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 motion in the image was smooth at that point

3.0 The Stationary Light Source Constraint

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 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 6), 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 and phenomenal demonstrations 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..

For static images of objects with cast shadows, the visual system must either assume a single light source illuminating all the objects in a scene or estimate the positions of different light sources illuminating the different objects. The phenomenal demonstration in

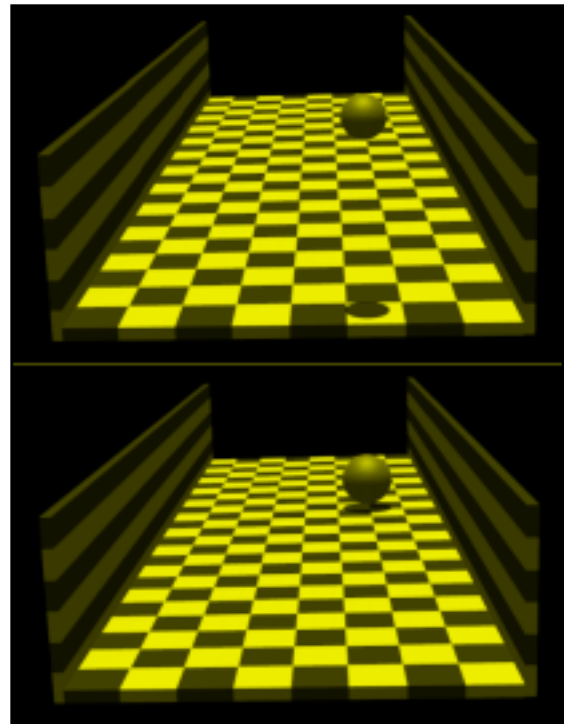


Fig. 4. The top and bottom panels show the extreme right position of the ball for the horizontal and diagonal shadow trajectories, respectively. In these static images, the effect of the shadow on the apparent size of the ball is small, but noticeable. In the dynamic case with diagonal trajectory, the 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.

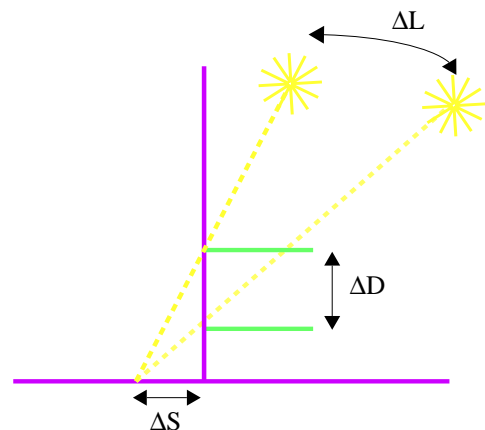


Fig. 6. A displacement ΔS between an object and its shadow can be produced either by a change in light source position, ΔL or by a change in depth of the object, ΔD .

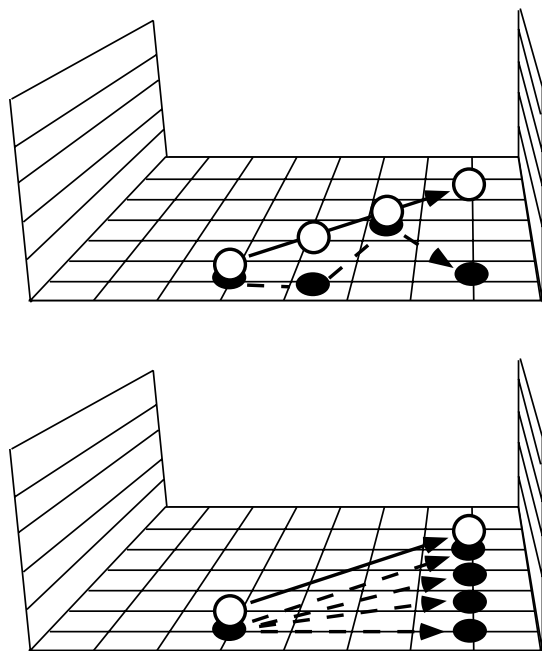


Fig. 5. Two schematic diagrams of some of the trajectories (in the image) followed by the ball and its shadow in the ball-in-a-box animations. Solid arrows indicate the trajectory of the ball (constant in all the animations), and dashed arrows indicate the trajectories of its shadow. (a) A time-lapse diagram of four frames from the animation used for Demonstration 3 (the non-linear motion). Observers reported the ball appearing to bounce at the third position from the left shown in the diagram. (b) The four different shadow trajectories used for Experiment 2. Each trajectory corresponds to a different animation used in the experiment.

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; Pentland, 1982; Ramachandran, 1988)). In order to explain the perception of motion in depth from moving cast shadows, we suggest that the visual system makes a different assumption about light sources: that the light source casting a shadow is fixed, at least on the time scale of the motion. We call this the *stationary light source constraint*. Such a constraint by itself supports only the qualitative perception of 3D object motion. In order to perceive the 3D motion of an object more exactly, the visual system must use image information or make assumptions about the exact position of the light source. Our discussion suggests two questions about the role of perceived light source position in the visual system's interpretation of cast shadow motion: First, does the system rely on a fixed light source constraint? Second, in making quantitative estimates of object motion, does the visual system estimate the light source posi-

tion based on image data, or does it rely on prior assumptions about light source position?

3.1 Experiment 2: A fixed light source constraint?

In order to study these questions, we designed a psychophysical paradigm to collect quantitative data on subject's perception of 3D motion from cast shadow motion. In the experiments, subjects viewed different ball-in-a-box animations and reported the height from the floor of the box to which the ball appeared to move at the right-most point of its trajectory (see the caption of figure 3 for a description of subjects' reporting method). We performed an initial, exploratory experiment to test whether subjects' performance could be fit by a model which based its estimates on a single, fixed position of the light source creating the ball's cast shadow. We tested four conditions, each corresponding to a different, linear shadow trajectory. The four trajectories had different slopes in the image plane, as shown in figure 5b. Figure 7 shows the results obtained for three observers. 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. 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.

The data from Experiment 2, while suggesting that the visual system uses a strategy in which it effectively accounts for light source position when interpreting cast shadow motion, does not directly answer either of the two questions we posed at the beginning of this section. We consider first the question of a fixed light source constraint and then turn to a consideration of whether and how the system estimates light source position. While the good fit of the fixed light source models to the data from Experiment 2 is consistent with the hypothesis that the visual system assumes a fixed light source constraint, observers could have

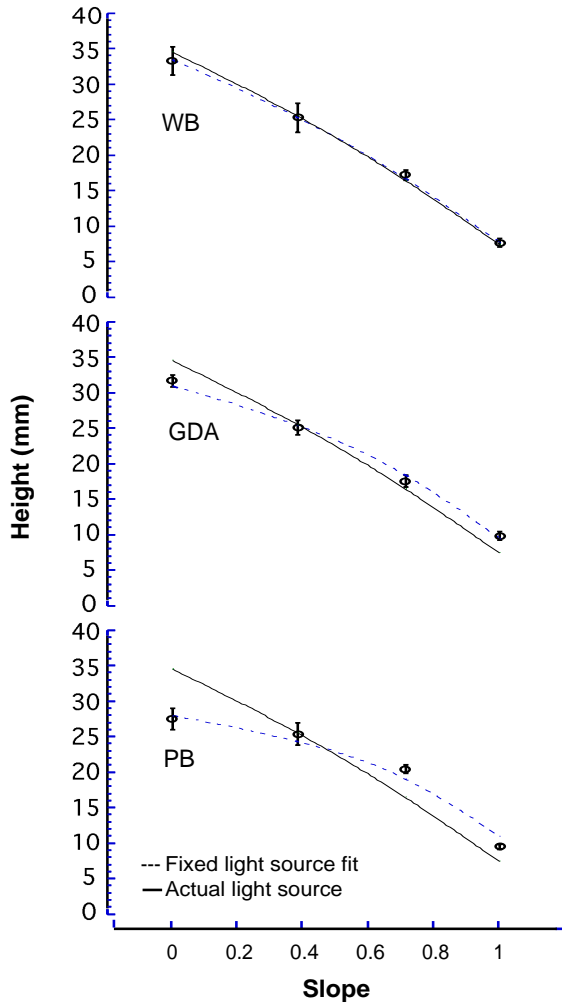


Fig. 7. Perceived height above the checkerboard floor of the ball, in the coordinates of the 3D simulated world, as a function of the shadow slope. Data are shown for three subjects. Each point is the mean of 8 measurements. Error bars indicate 1 S.E. of the mean. As the shadow's trajectory slope goes 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 each subject bases his or her estimate of object motion on an some different fixed light source position. In terms of distance (mm) from the middle of the checkerboard floor and slant (deg) with the floor, the light positions used to fit the data were: (419 mm, 60.8°); (105 mm, 50.4°); and, (67 mm, 46.8°) for observers WB, GDA, and PB, respectively.

used information in the stimulus (e.g. the shading on the ball and on the walls of the box) to infer that the light source was fixed. A stronger test of the hypothesized constraint would be to test 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.

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

In order to answer this question, we made a number of 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. 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 demonstrations and experiments (moving diagonally in the image plane, with no change in size). Unlike in the previous demonstrations and experiments, we generated shadows for these animations by rendering the scene with ray-tracing from the light source; however, we simulated 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 (the "correct" interpretation given the way the animations were generated).

Three demonstrations support the hypothesis that the visual system relies on a fixed light source constraint when interpreting shadow motion. For the first of the demonstrations (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). As in Demonstration 1, all observers reported seeing the ball as moving along different 3D trajectories in the two animations. 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);

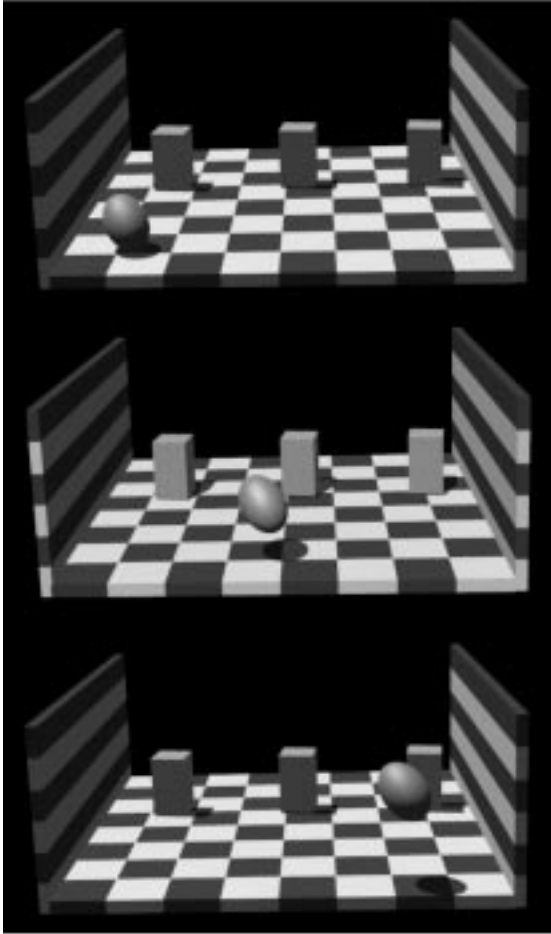


Fig. 8. From the top, the panels show frames 1, 15, and 30 of a 30 frame sequence in which there is evidence from the shading and cast shadows that the illumination direction is changing as the ball moves from left to right. If the visual system could take this information accurately into account, it would conclude that the football is moving along a linear trajectory in the fronto-parallel plane. It does not; rather the percept is of a football starting near the observer (frame 1), moving first forward and then back in depth (frame 15), and then towards the observer again (frame 30).

however, the phenomenon remained unchanged--observers still reported seeing different motions for the ellipsoid in the two animations. 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. The resulting animations 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. Finally, we generated an animation (Demonstration 7; figure 8) in which the motion of the light source caused a non-linear shadow motion which mimicked that of Demonstration 3, but with the objects of Demonstrations 5 and 6. When this animation was shown after the animation used in Demonstration 3, observers reported

that their percepts of non-linear 3D motion were the same for both animations. 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 strong evidence to the contrary.

3.3 Experiment 3: Is effective light source direction determined by prior assumptions or image data?

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 a final experiment (Experiment 3), we tested whether subjects' implicit light source direction is determined by the shading information on the ball or a prior bias. We ran the same ball-in-a-box experiment used for Experiment 1 with three different shading conditions for the ball, corresponding to three different, fixed light source positions (see figure 6 caption). 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 data (figure 6) showed a very small but significant effect consistent with observers' usage of shading information to indicate light source direction. The size of the effect, however, was far from what would be predicted theoretically, suggesting that in this experiment, a strong prior bias for a default light source position determined performance. 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.

4.0 Discussion

Our results raise a number of issues about the computations involved in the perception of 3D motion from shadow motion. Although we have shown that moving cast shadows are sufficient to produce apparent motion in depth, we have not delineated the specific properties that shadows must have to perceptually "link" them with their casting objects to produce the apparent motion. Experiment 1 showed that physically implausible light shadows could produce motion in depth, albeit with less frequency than dark shadows. In additional experiments using the ball-in-a-box simulation, we found that an object's cast shadow does not have to be physically reasonable--it can have the wrong contrast polarity or brightness -- for observers to see motion in depth. These results stand in contrast to those obtained for the interpretation of shadows in static images, which show that similar manipulations

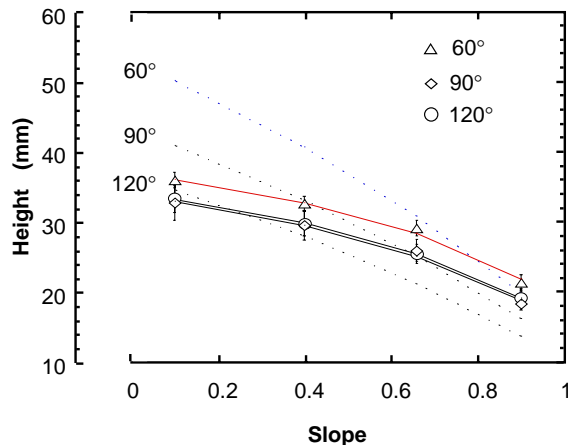


Fig. 9. Perceived height above the checkerboard floor of the ball for Experiment 3. 40 subjects were split into 3 groups (13, 13 and 14). Each group was shown animations like those used in Experiment 2 (i.e. having four different shadow trajectories) in which the ball had a different shading pattern, corresponding to being illuminated by a light source from one of three angles above the checkerboard: 60°, 90° and 120° (recall that the viewing direction was 21.8° above the checkerboard). All light sources were at infinity. The data predicted for an observer which accurately estimates the light source positions and uses these to mediate its inference of 3D object motion from cast shadow motion are shown by the dashed lines. 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 be predicted for a single intermediate light source position. Error bars indicate 1 S.E. of the mean.

of shadow brightness and contrast strongly interfere with shape perception Cavanagh, P., & Leclerc, Y. G. (1989). Moreover, the effect in our experiments is resistant to some significant deformations in the shape of the shadow. Replacing the ellipsoidal shadow of the ball with a square shadow, for example, does not reduce the effect.

The different results obtained for the interpretation of moving shadows and the interpretation of static shadows suggests that dynamic displays contain an important piece of information not available in static displays. The strongest candidate for such a piece of information is the correlation between the motions of objects and their cast shadows in dynamic displays. The nature of the correlated motion is related to the imposition of a stationary light source constraint. Assumption of 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. If the light 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. This suggests that the visual system may have

special mechanisms for detecting the correlated motions of objects and shadows. Such mechanisms would not only support the inference of depth from cast shadows, but they would also support the discounting of shadows as objects, in a way roughly analogous to the way the auditory system discounts echoes. Moreover, since the type of correlated motion we have described also provides useful information for perceptually linking disparate regions of an occluded object, the hypothesized motion detectors could subserve this important perceptual function as well.

A final issue raised by the demonstrations is 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 (Kersten, 1990; Rock, 1983). Assuming such a framework for visual system processing suggests a program of psychophysics which we refer to as a psychophysics of constraints. 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. This paper has presented an application of such a program of research to the perception of 3D spatial layout and motion from cast shadow information.

References

- Biederman, I. (1985). Human image understanding: recent research and a theory. *Computer Vision Graphics and Image Processing*, **32**, 29-73.
- Cavanagh, P., & Leclerc, Y. G. (1989). Shape from shadows. *Journal of Experimental Psychology, Human Perception and Performance*, **15**, 3-27.
- Cavanagh, P. (1991). What's up in top-down processing? In A. Gorea (Ed.), *Representations of Vision: Trends and tacit assumptions in vision research* (pp. 295-304).

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

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

Gogel, W. C., Hartman, B. O., & Harker, G. S. (1957). *Psychological Monographs*, **71**, 13-14.

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.

Mamassian, P., Kersten, D., & Knill, D. C. (1992). Spatial layout from cast shadows. *Association for Research in Vision and Ophthalmology*. Sarasota, Florida. **33**, 3191.

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

Pentland, A. P. (1982). Finding the illuminant direction. *Journal of the Optical Society of America*, **72**, 448-455.

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

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.

Yonas, A. (1978). Development of sensitivity to information provided by cast shadows in pictures. *Perception*, **7**, 333-341.

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