

The perception and recognition of natural object shape from deforming and static shadows

J Farley Norman, Thomas E Dawson, Shane R Raines

Department of Psychology, Western Kentucky University, Bowling Green, KY 42101, USA;

e-mail: Farley.Norman@wku.edu

Received 28 January 1999, in revised form 4 August 1999

Abstract. In this study of the informativeness of shadows for the perception of object shape, observers viewed shadows cast by a set of natural solid objects and were required to discriminate between them. In some conditions the objects underwent rotation in depth while in other conditions they remained stationary, thus producing both deforming and static shadows. The orientation of the light source casting the shadows was also varied, leading to further alterations in the shape of the shadows. When deformations in the shadow boundary were present, the observers were able to reliably recognize and discriminate between the objects, invariant over the shadow distortions produced by movements of the light source. The recognition performance for the static shadows depended critically upon the content of the specific views that were shown. These results support the idea that there are invariant features of shadow boundaries that permit the recognition of shape (cf Koenderink, 1984 *Perception* **13** 321–330).

1 The perception and recognition of natural object shape from deforming and static shadows

Informally, it has long been known that cast shadows contain perceptually useful information about the shape of three-dimensional (3-D) objects. For example, cast shadows have been included in works of art (paintings, mosaics, etc) for centuries, if not thousands of years, to make the depicted 3-D scenes more realistic (Gombrich 1995). In addition, the properties of cast shadows and of their relationships with the casting objects and light sources were studied and analyzed 500 years ago by Leonardo da Vinci (Vinci 1989). Despite this early scientific and artistic interest in cast shadows, there has been little systematic research into their usefulness for the perception and recognition of naturally shaped 3-D objects by modern-day investigators. The purpose of the current study is to revisit the issue and examine how human observers use cast shadows, both stationary and moving, with no other accompanying visual information to recognize the shapes of natural objects.

In the past, many investigators have examined related issues. For example, consider silhouettes, which are a special case of cast shadows (see Gombrich 1995, plate 19; also Grafton 1979). Silhouettes are made by placing a solid 3-D object in between a light source and a translucent projection screen, which is oriented at a right angle to the direction of illumination. The shadows of the object are cast by the light source onto the projection screen, and are then viewed or recorded by an observer from the other side. The sharp, clean edges of silhouettes result from the fact that the objects are typically placed as close as possible to the surface of the projection screen. In 1953, Wallach and O'Connell studied the deformation of silhouettes that occurs when the objects casting shadows are rotated in depth. They used a variety of moving 3-D objects, such as wooden blocks, truncated cylinders, wire-frame figures, and luminous rods. Wallach and O'Connell found that, when human observers viewed the deforming silhouettes, they typically reported that they perceived solid, rigidly rotating 3-D objects, rather than the distorting silhouettes themselves. They referred to this phenomenon as the 'kinetic depth effect' (ie perceived depth through motion). However, Wallach and O'Connell also found that this effect was limited to silhouettes that possessed linear boundary contours, sharp corners, or readily identifiable object vertices. For example, they concluded

(page 209) that “curved contours which are deformed without displaying a form feature which identifies a specific point along the curve are seen as distorting, often even if for some reason the shadow is seen as a 3-D form ... it is now clear that the perceived distortions of deforming curved contours are not related to the kinetic depth effect at all”.

More recent research has shown that the conclusions of Wallach and O’Connell concerning the deformations of shadows with smoothly curved boundaries were incorrect. For example, Todd (1985), Cortese and Andersen (1991), and Norman and Todd (1994) have all demonstrated that deforming curved contours can reliably lead to the perception of a rigidly moving 3-D object. This finding is important, because many naturally shaped objects have surfaces that are smoothly curved and lack the sharp corners found in man-made objects like those used in Wallach and O’Connell’s study. Despite its significance, the research of Todd (1985), Cortese and Andersen (1991), and Norman and Todd (1994) is limited in that relatively simple curved objects were used (eg ellipsoids). No researchers have yet investigated the perception and recognition of 3-D shape from the shadow deformations of more complicated, smoothly curved objects (see, however, Norman et al 1995, who examined the perception of 3-D shape from deformations of shading and deformations of shadow boundaries plus specular highlights).

At this point, it is important to note that most of the traditional computational analyses designed to recover 3-D structure from motion (eg Ullman 1979; Hoffman and Bennett 1986) would not be able to recover 3-D structure from the deformation of smoothly curved shadows over time. This is because these models require the measurement of projected velocity or position of specific, identifiable, points on an object’s surface across multiple views. One cannot establish such a ‘correspondence’ of how a single surface location moves over time from a sequence of smoothly curved cast shadows, since different points on the surface of the object project to the shadow boundary at different moments in time. Thus, the motions of the projected shadows of smoothly curved objects are qualitatively different from the projected motions of specific surface locations marked by texture, corners, etc (also see Todd 1985, and Norman and Todd 1994), and thus require different computational and/or perceptual mechanisms for their analysis.

From the previously reviewed psychophysical studies, it seems clear that, under the right conditions, human observers can obtain useful information about 3-D objects from silhouettes and cast shadows. There are many gaps in our knowledge of how this process takes place, but the simple fact that we can perceive at least some aspects of the shape of 3-D objects from shadows is certain. However, the exact relationship between cast shadows and perceived 3-D structure is bound to be far from simple. Leonardo da Vinci pointed out that the shape of the cast shadow does not often bear a close resemblance to the object casting the shadow (Vinci 1989, pages 105–106). Rather, the shape of the shadow varies with the orientation of the object relative to the observer, the orientation of the object relative to the positions of environmental light sources, the size and distance of the light sources, and the shape (curved vs noncurved) and orientation of the background surface on which the shadow is cast. At this point, we do not know whether the perception or recognition of a 3-D object defined only by its cast shadow is invariant over such complex transformations.

An example that illustrates the nontrivial effect of moving the source of illumination is shown in figure 1. It is important to keep in mind that silhouettes are a special case of cast shadows where the light source is positioned so that the direction of illumination is perpendicular to the surface upon which the shadow is cast (Gombrich 1995, plate 19). For an observer viewing the silhouette, the direction of illumination is coincident with the observer’s line of sight. In nature, light from environmental sources like the sun rarely comes from a direction along an observer’s line of sight. Most of the time, the light is obliquely oriented at various angles—this variation in light-source

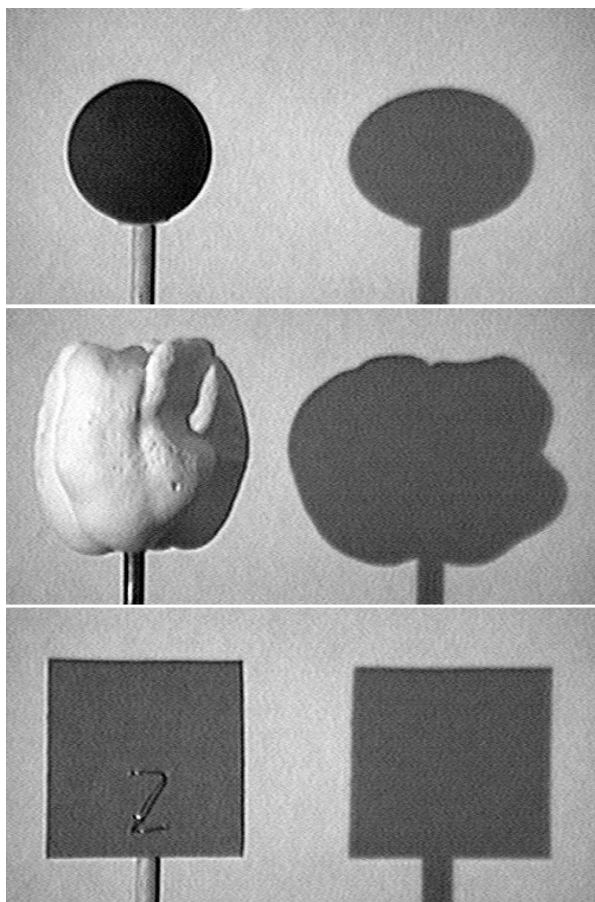


Figure 1. An example of the different types of shadow deformation that can occur for a simple curved object (top panel), a more complex, naturally curved object (middle panel), and a flat planar surface (bottom panel).

orientation has large effects upon the particular shape of the shadows that are cast by environmental objects. For example, figure 1 shows the shadows cast by three differently shaped objects when a light source is positioned 40° away from the line of sight. If the light source were located along the line of sight, the resulting silhouettes would have the same projected shape as the visible outer boundaries of the objects shown on the left side of the photograph. Notice that the effects of changing the light-source orientation are large and complex, depending upon the type and shape of the objects casting the shadows. First, consider the shadow of the spherical ball: the ball would produce a circular silhouette if it were frontally illuminated, but the shadow projected from a light source positioned away from the line of sight is shaped like an ellipse. In this case, the shadow resulting from oblique illumination is an affine (stretching) transformation of the silhouette cast from a frontal illumination. Notice that the transformation in the middle panel for a naturally shaped object (a bell pepper) is clearly not affine: the silhouette would be smoothly curved on the left and right sides if the light source had a frontal orientation, but illumination from an oblique angle produces an elongated shadow with a distinct 'notch' or indentation on the right side. Finally, notice that no stretching or nonlinear transformation occurs for oblique illumination of a flat cardboard square placed parallel to the projection surface (one can see this 'non-transformation' by holding a planar object parallel to a flat sidewalk on a sunny day). Leonardo da Vinci was right: there are many variations in the specific 2-D shape of a cast shadow that depend upon a wide variety of factors independent of the object itself. One question that has yet to be resolved concerns this issue: is the human perception

and recognition of shape invariant over such complex transformations as induced by movements of environmental light sources?

During the 1980s and 1990s, there has been a sustained interest in object recognition (eg Biederman 1987; Tarr 1995), although few of these studies have specifically dealt with the ability to recognize objects solely from silhouettes or cast shadows. Most of the studies have used photographs, solid outlines with internal shading, or line drawings with internal contours (eg Biederman and Ju 1988; Biederman and Gerhardstein 1993; Hayward and Tarr 1997). In some recent studies, however, Hayward (1998) and Tjan et al (1995) found that the ability to recognize objects from stationary silhouettes was essentially comparable to that achieved when internal contours, texture, and shading were present. These results suggest that static silhouettes contain much of the information that human observers need for recognition, at least for the type of objects used in the studies of Hayward and Tjan et al (simple volumetric primitives or 'geons'; aggregates of volumetric primitives, and animals). These studies are highly interesting and informative. Nevertheless, silhouettes represent only a special case of cast shadows in general. The results obtained with silhouettes do not necessarily extend to more general conditions of environmental illumination. It is also not yet clear how the recognition performance for static silhouettes would be affected if they deformed over time in response to the rotation of the actual objects in depth.

The purpose of the current experiment on the perception of cast shadows is fourfold: (i) to measure human observers' ability to discriminate and recognize naturally shaped objects from their cast shadows, (ii) to examine in more detail the perceptual effects of changes in the position of light sources, (iii) to compare discrimination and recognition performance for static and deforming shadows, and (iv) evaluate how recognition is influenced by the content of specific static shadows. In the past, there has been little research on object recognition in which naturally shaped objects have been used. Most of the previous studies have been devoted to the recognition of man-made objects with which observers are familiar (cars, airplanes, telephones, etc) or artificial objects made out of simple volumetric parts that are often familiar (cylinders, cones, etc). Given that our visual systems evolved in a natural environment, it would seem desirable to test observers' abilities with a set of highly similar, natural objects. In addition, to our knowledge, there has been no previous research into how human observers recognize the shape of 3-D objects from deforming cast shadows whose boundaries do not contain identifiable surface features. Thus, if observers can perform such a discrimination task accurately, it cannot be due to the operation of mechanisms implementing conventional 3-D structure-from-motion analyses. The purpose of the current experiment was to explore such issues.

2 Method

2.1 Apparatus

The shadows were created with a shadow-casting apparatus, similar to that used by Wallach and O'Connell (1953), where a physical light source is used to cast shadows of an object onto a translucent (ie not transparent) screen. In this arrangement, the resulting 2-D shadows are then viewed from the opposite side of the screen. Following Wallach and O'Connell's suggestion, to obtain the sharpest and clearest shadows (simulating outdoor conditions on a sunny day, no penumbra), we used the closest possible distance between our objects and the translucent screen (about an inch, or 2.5 cm, gap), and the maximum possible distance between the object and light source, in our case 3.6 m. The light source used to cast the shadows was a 250 W halogen lamp; an additional hood was placed around the lamp to direct the light towards the object and screen. A video camera was placed on the opposite side of the translucent screen, and the resulting cast shadows of the objects were digitally captured by an Apple Power Macintosh 8500/120.

2.2 Stimulus displays

The experimental stimuli were the digital images of the cast shadows of the objects, as captured by the computerized shadow-casting apparatus described earlier. The objects used to generate the shadows were plaster of Paris replicas of five ordinary bell peppers. The five were specifically chosen to have extremely similar sizes (eg vertical size for all of them was 8.0 cm), in order to prevent recognition based on overall differences in size, etc. 72 different shadows of these objects were obtained by placing them on a turntable located between the light source and projection screen and rotating the turntable in 5° angular increments. This process was repeated for three different orientations of the light source, 0°, 22.5°, and 45.0° from a direction perpendicular to the projection screen. Thus, there were 1080 (5 × 72 × 3) distinctly different images of the object shadows that were digitally captured and saved for later presentation. Example shadows of the five objects are shown in figure 2. It is important to keep in mind that these shadows would look entirely different for other combinations of object and/or light-source orientation. Figure 3 shows the effects of increasing the angle of illumination of the light source for the values used in this experiment. In this case (unlike the middle panel of figure 1), changing the angle of illumination led to a ‘stretch’ that affected the magnitudes of the curvatures along the shadow boundary contour, but did not alter appreciably the presence or absence of the shadow’s convexities and concavities. The final images of the cast shadows had a resolution of 640 × 480 pixels.

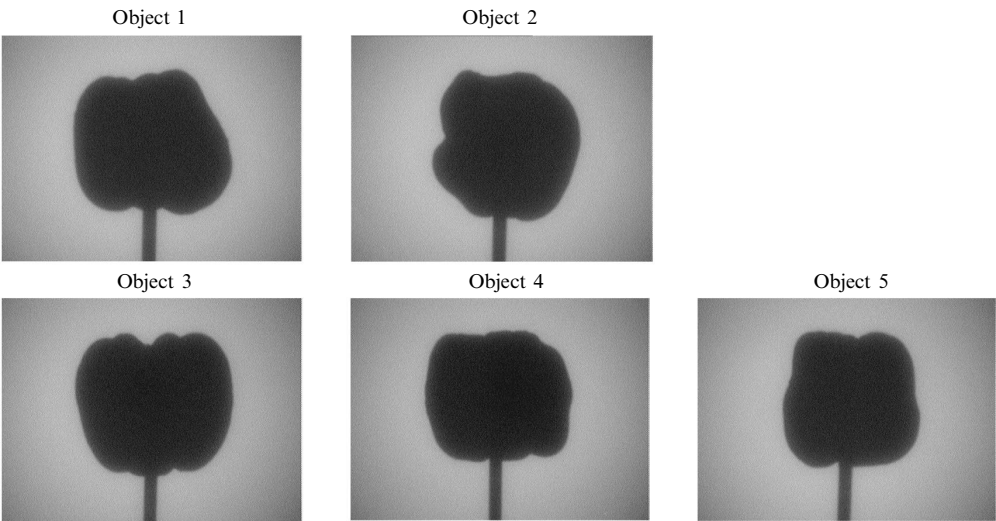


Figure 2. An illustration of the cast shadows used in the experiment.

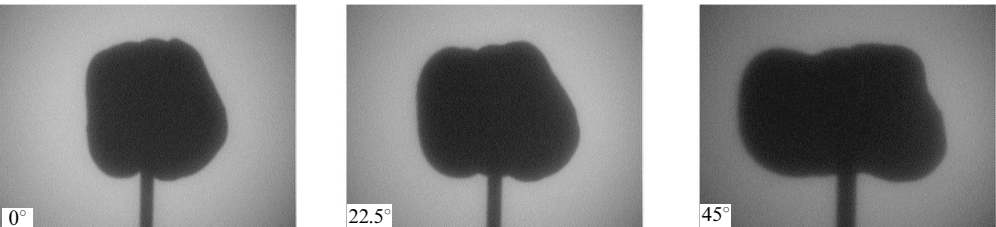


Figure 3. An illustration of the effects of changing the angle of illumination by moving the light source relative to object 1 and the projection screen.

2.3 Procedure

The digital images of the shadows that were captured by the Power Macintosh 8500/120 were transferred to a Power Macintosh 8600/300 that controlled the actual experiment and presented the shadows to the observers. Each of the observers participated in ten experimental sessions, each session consisting of a single block of 300 trials. Within each block, observers were shown 300 shadows (30 experimental conditions \times 10 trials per condition). The 30 conditions resulted from the orthogonal combination of 5 objects \times 3 angles of illumination \times 2 motion types, moving or deforming versus static. The order of the conditions within the block was randomly determined for each observer, as were the specific orientations of the static views (72 possibilities for each object and angle of illumination). The observers' task was to indicate from either the deforming or static shadows which object was depicted (1–5) by pressing the appropriate key on the computer keyboard. The observers could view each shadow, whether moving or static, for as long as necessary to recognize the object. In the case of the deforming shadows, full rotations of the objects were depicted (ie all 72 views were presented). After completion of the ten experimental blocks of trials, a total of 100 responses had been collected for each of the 30 conditions (3000 total trials per observer).

In order for the observers to make an appropriate response (ie object 1, 2, 3, 4, or 5) on each trial during the experiment, it was necessary that the observers learn, prior to the start of an experimental session, which object was which. In our case, we allowed the observers to participate in a practice set of trials immediately before the start of an experimental session. In these practice trials, the observers saw a series of 10 deforming shadows (5 objects \times 2 practice trials per object) at the 0° angle of illumination only, and were required to identify the object (1–5). If the observer was correct, he/she received feedback in the form of a short auditory beep. Successive blocks of 10 practice trials were run until the observer reached a criterion of 90% correct responses, then the experimental block of trials was initiated. All of the observers were able to reach this criterion, showing that it was possible to recognize natural objects from deforming shadows, at least under certain conditions.

It is crucially important to note, however, that the observers never received feedback during an experimental block of 300 trials. As a consequence, they never received feedback for any shadows resulting from the 22.5° or 45° angles of illumination. They also never received any feedback for any static views of any shadow of an object regardless of the angle of illumination.

2.4 Observers

The observers included two of the authors (TED and SRR), one other experienced psychophysical observer (HFN), and one additional observer (SRC) who was naive with regard to the purposes of the experiment.

3 Results

The results for two representative observers (TED and SRC) are shown in figures 4 and 5, in which recognition accuracy is plotted as a function of object, motion type (deforming vs static), and angle of illumination. All four observers showed essentially the same pattern of results. There was a large effect of motion (within-subjects factorial ANOVA, $F_{1,87} = 258.8$, $p < 0.001$), such that performance was nearly perfect for recognition of objects defined by deforming shadows, but was significantly reduced when static shadows were presented. Across all observers and conditions, recognition accuracy was 98% for deforming shadows and 64% for static shadows. Chance recognition performance for this task would be 20%. Consequently, even though performance was reduced when no motion was present, the observers were still able to identify the correct object in most instances. Other significant effects included a main effect involving the objects

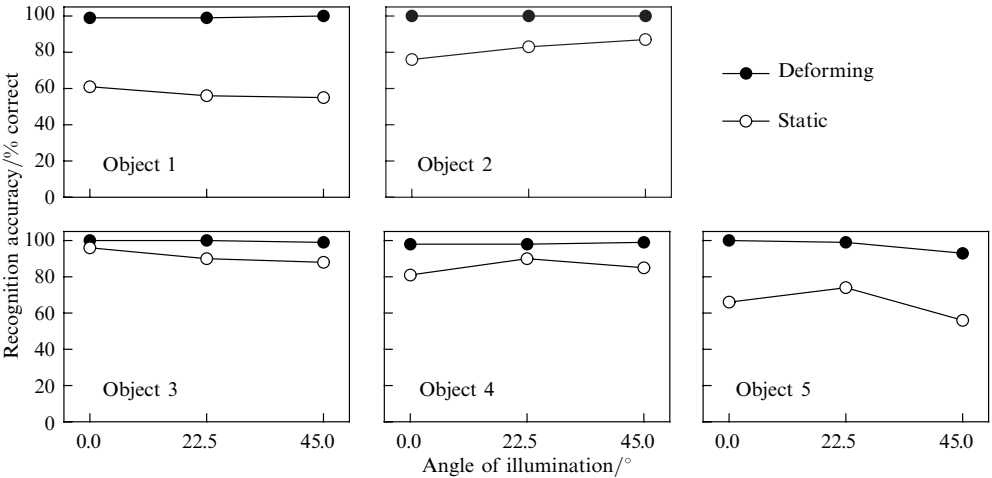


Figure 4. Recognition accuracy results for observer TED. The results are plotted as a function of object, presence or absence of motion, and changes in the angle of illumination.

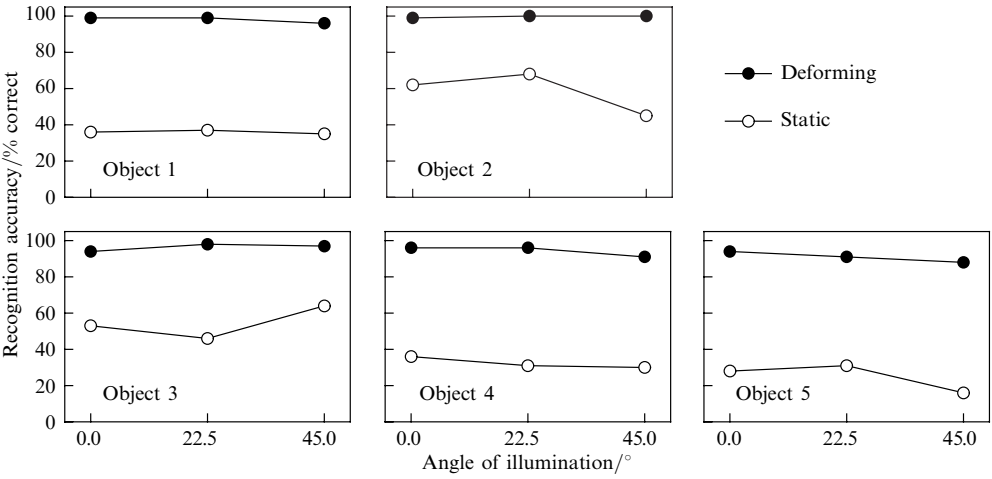


Figure 5. Recognition accuracy results for observer SRC.

themselves ($F_{4,87} = 9.0, p < 0.001$) and an object \times motion interaction ($F_{4,87} = 6.9, p < 0.001$). In general, object 3 was easiest to recognize, while object 5 was the most difficult. The object \times motion interaction was due to the fact that the absence of motion in the static displays made a large difference for recognizing some objects (eg objects 1 and 5) and less of a difference for other objects (eg object 3). All of these effects are evident from an inspection of figures 4 and 5. There were no other significant effects—for example, there was no main effect or significant interaction that involved the angle of illumination.

A more detailed analysis of the observers' recognition performance from the static shadows is shown in figure 6 for observers SRC and HFN. Recognition accuracy is plotted for object 1 as a function of the angle of illumination and orientation of the object. It is clear from these results that some static views were much more recognizable than others. Some views were recognized 100% of the time, while other views were never recognized. In addition, there is a striking similarity between the results of the two observers, suggesting that they were using the same features of the shadow boundaries to distinguish this object from the others. Indeed, the judgments of observers SRC and HFN were highly correlated (Pearson $r = 0.85$) for this object. Sizable correlations

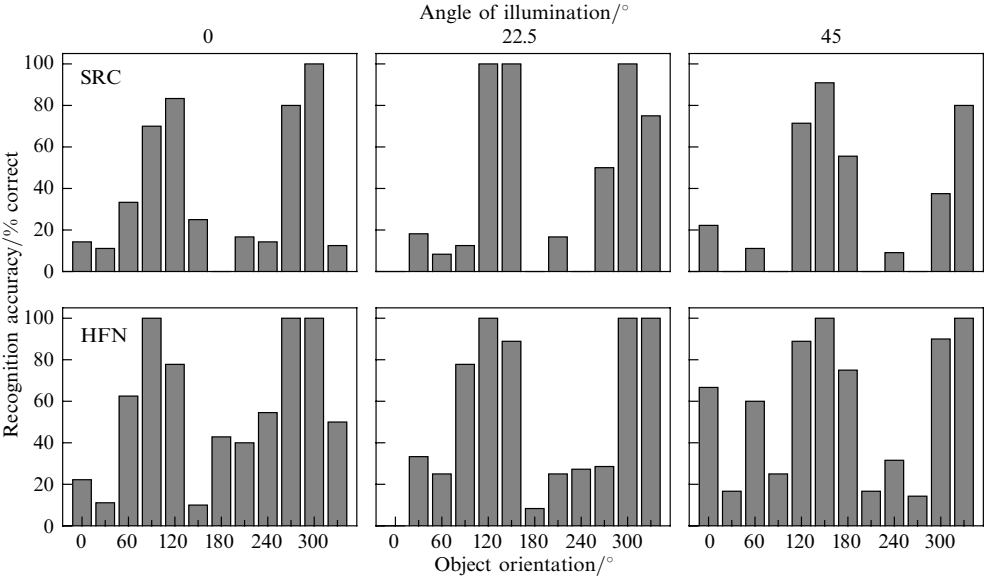


Figure 6. A comparison of the recognition performance exhibited by observers SRC (top) and HFN (bottom) for the static cast shadows of object 1. Accuracy is plotted as a function of the orientation of the object.

also existed for these observers for objects 2 ($r = 0.48$) and 5 ($r = 0.29$). Easily recognizable static views for objects 1 and 2 are illustrated in the left column of figure 7. These views would be recognized correctly by our observers, while the other views for the same objects (in the right column) would rarely, if ever, be recognized. Notice that the static shadows in the left column have prominent convexities/concavities along their boundaries. Object 1 has a prominent ‘bulge’ on the right, while object 2 has a ‘notch’ or deep indentation. There was no similarity or correlation between the judgments of observers SRC and HFN for objects 3 and 4 (Pearson r s of 0.02 and 0.04, respectively), suggesting that for these objects each observer was relying on a different set of boundary features for their identification.

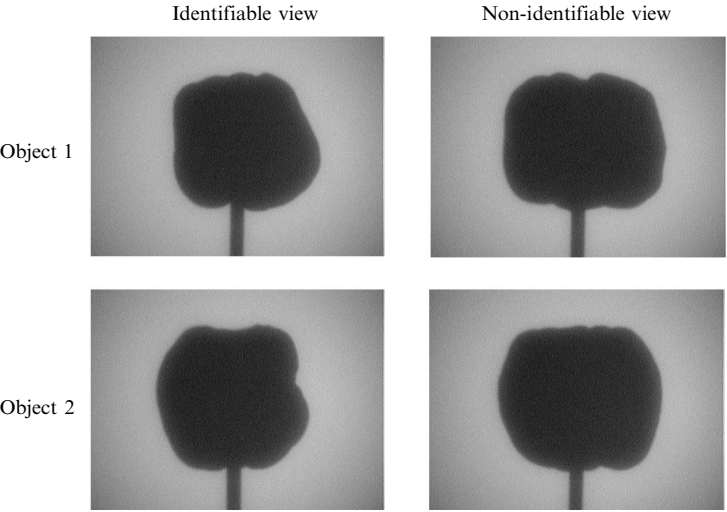


Figure 7. An illustration of representative identifiable and non-identifiable shadows for objects 1 and 2.

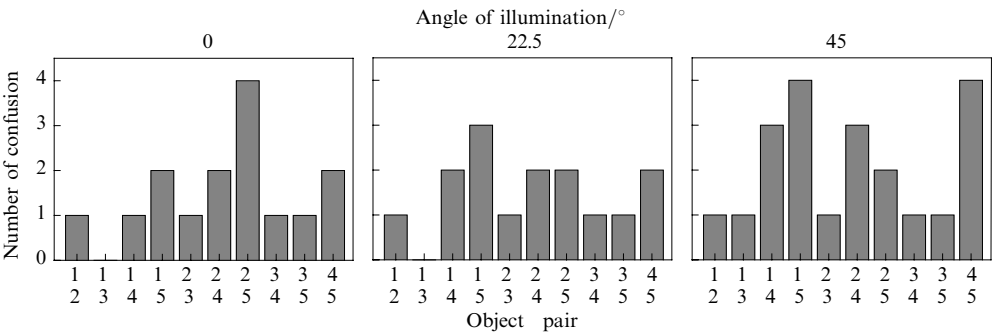


Figure 8. Results of the $-\ln \eta$ analysis conducted upon the responses obtained for the static shadows. The number of confusions plotted on the ordinate indicates the number of observers who had difficulty in discriminating (ie confused) each of the ten pairs of objects. Results are shown for each angle of illumination.

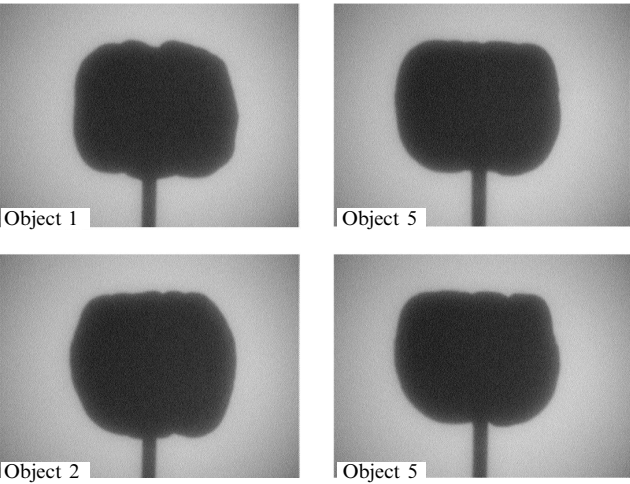


Figure 9. Confusable pairs of static shadows. The top row illustrates perceptually similar shadows of objects 1 and 5; the bottom row depicts perceptually similar shadows of objects 2 and 5.

A final analysis was performed on the results obtained for the static shadows. Confusion matrices were obtained for all pairs of objects for each observer, and $-\ln \eta$ values were calculated (Luce 1963).⁽¹⁾ $-\ln \eta$ is the measure of discriminability or ‘psychological distance’ between two stimuli that is derived from Luce’s choice theory, analogous to the d' of signal detection theory. A $-\ln \eta$ value of 0.0 indicates that two stimuli are indiscriminable, while values of 4.0 or higher represent essentially perfect discrimination performance. Figure 8 provides histograms for each angle of illumination showing how many observers confused (ie had low discrimination performance indicated by values of $-\ln \eta$ below 2.0) each of the 10 pairs of objects. Some pairs (eg objects 1 and 2, and 1 and 3) were almost never confused (see left column of figure 7), while other pairs (eg objects 1 and 5, 2 and 5, 4 and 5) were confused by nearly all of the observers. The static shadows of objects 2 and 5 were frequently confused by all observers at the 0° angle of illumination. The static shadows of objects

⁽¹⁾ The $-\ln \eta$ values were calculated according to the following equation:

$$-\ln \eta = \frac{1}{2} \ln \frac{(N_{a,a} + \frac{1}{2})(N_{b,b} + \frac{1}{2})}{(N_{a,b} + \frac{1}{2})(N_{b,a} + \frac{1}{2})},$$

where $N_{a,a}$ is the number of responses of a when a was the actual stimulus, $N_{b,b}$ is the number of responses of b when b was the actual stimulus. Likewise, $N_{a,b}$ is the number of responses of a when b was the actual stimulus, and $N_{b,a}$ is the number of responses of b when a was the actual stimulus.

1 and 5 were frequently confused at both the 22.5° and the 45° angles of illumination, while the shadows of objects 4 and 5 were also highly confusable at the 45° angle of illumination. Confusable pairs of views (objects 1 and 5, 2 and 5) at the 22.5° angle of illumination are shown in the top and bottom rows of figure 9. Notice that just as the analysis suggests, the static shadows of objects 1 and 5 (top row), and 2 and 5 (bottom row) are perceptually very similar in appearance, despite the fact that there are numerous quantitative differences along their boundaries. It is important to note from this analysis, however, that about half of the object pairs had low confusabilities even when the static shadows were severely distorted at the 45° angle of illumination.

4 Discussion

The results clearly show that observers can recover significant amounts of information about 3-D shape from shadow boundary contours, especially when the shadows deform and change over time, even for a highly similar class of naturally shaped objects. This occurred despite the fact that the shadows were smoothly curved, and thus contained no projections of identifiable surface features whose motions could be analyzed by conventional structure-from-motion algorithms. The finding that observers could perceive rigid rotations in depth from such smoothly curved shadow deformations and then recognize the shape of individual objects confirms and extends the earlier research on shadow deformations by Todd (1985), Cortese and Andersen (1991), and Norman and Todd (1994).

Much of the research involving cast shadows in the past has not been concerned with the informativeness of shadows for the perception or recognition of shape *per se*, but has used them for other purposes. For example, cast shadows have often been employed to inform observers about the prevailing direction of illumination (eg Berbaum et al 1984; Erens et al 1993) or signal the motion of objects in depth (eg Norman and Todd 1994; Kersten et al 1996, 1997). In experiments designed to examine the perception of 3-D shape, cast shadows were not the sole focus of investigation. They were often used to supplement other forms of information, such as image shading (eg Mingolla and Todd 1986). The informativeness of cast shadows presented in isolation has not often been studied in its own right. Even the early pioneering research of Wallach and O'Connell (1953) on the kinetic depth effect in which the shadows of solid objects were used was a special case: in their experiment 1 they used solid objects that had sharp corners whose projected positions could be tracked over time, making them fundamentally no different from the patterns used in a conventional structure-from-motion experiment in which dots or texture are used to define the motions of the 3-D objects (eg Green 1961; Braunstein and Andersen 1984; Todd et al 1988; Norman and Lappin 1992). Indeed, in later experiments Wallach and O'Connell abandoned the shadows of solid objects and investigated the projected motions of wire-frame figures, luminous rods, and other figures with visible endpoints, etc. In the present experiment, we used natural objects that did not contain identifiable surface points, and nevertheless demonstrated that smoothly curved shadows with no other accompanying sources of visual information are sufficient for the accurate recognition of 3-D shape. As far as we are aware, this is the most conservative test yet of the ability of human observers to recognize objects from their cast shadows.

The present results confirm and extend the results of Hayward (1998) and Tjan et al (1995) who investigated the informativeness of stationary silhouettes. In our experiment, we also found that static silhouettes often led to accurate recognition. This was also true for the static shadows cast at oblique angles of illumination. The recognition performance for static silhouettes and shadows, however, was typically much less than that obtained for deforming silhouettes and shadows. At the same time, it is also clear that some orientations of the objects used in our experiment produced static silhouettes and shadows that were much more discriminable than others. On inspection,

these highly recognizable shadows invariably contained prominent convexities or concavities of the boundary contour. In contrast, views that did not contain any such significant features were invariably difficult to identify. Such convex and concave regions were shown by Koenderink and van Doorn in 1976 to correspond to elliptic (ie 'bump-like') and hyperbolic (ie 'saddle-like') regions on the surface of a 3-D object, respectively (also see Hoffman and Richards 1984; Koenderink 1984a, 1984b; Richards et al 1987; Singh et al 1999). Koenderink and van Doorn's finding is important, because all smoothly curved solid objects can be described or represented in terms of these qualitatively distinct types of surface regions, along with the parabolic (ie 'cylinder-like') surface regions that separate them. Our results are also consistent with those of Attneave (1954), who found that convex and concave regions of randomly shaped 2-D figures were especially informative for human observers.

In our study, we found no systematic effects of the angle of illumination, despite the fact that varying it produced large changes in the specific shape of the shadows that were cast. The change in light-source position led to an elongation of the cast shadows, plus a possible catastrophe (ie the emergence of new convex/concave regions or the disappearance of old convex/concave regions; cf Arnold 1984; Koenderink 1990; see figure 1, middle panel, for an example catastrophe), depending upon the specific relative orientation of the object and light source. Such catastrophes are informative, in that they signal the presence and shape of new surface regions not previously visible—whenever a catastrophe occurs, one learns more about the shape of the object that one is viewing.

In our experiment, catastrophes of the shadow boundary contour also occurred over time in the moving conditions, because of the rotation of the objects in depth. For our smoothly curved objects, there were two general types of boundary alteration; they are depicted in figure 10 (cf Koenderink and van Doorn 1976). The sudden appearance of a prominent convexity in the shadow boundary signalled the presence of an elliptic surface region not previously visible (top row), whereas the sudden appearance of a prominent concavity similarly revealed the presence of a hyperbolic surface region (bottom row). Examples of these boundary changes over time for the shadows used in the present experiment are shown in figure 11. It is important to realize, however, that similar boundary catastrophes also occur in everyday life as observers and objects move relative to each other and to environmental light sources. In fact, from the types of boundary deformations identified in figure 10, observers could theoretically form a qualitative representation of the 3-D structure of a smoothly curved object in terms of the arrangement and location of its constituent elliptic and hyperbolic surface regions.

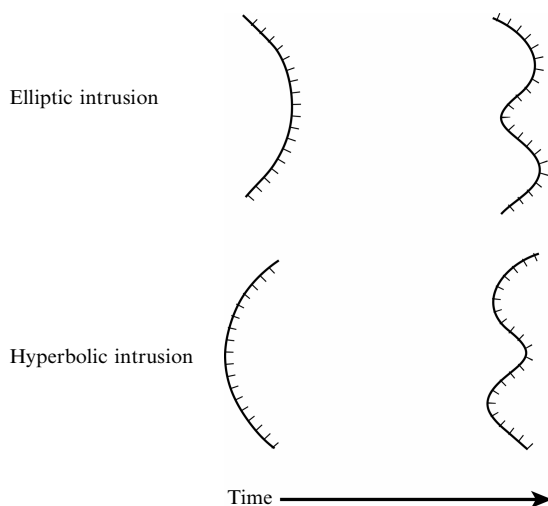


Figure 10. A schematic diagram illustrating the two types of boundary deformation commonly present in the deforming shadows used in the experiment. The solid lines represent the shadow boundaries, whereas the hatches indicate the sides of the boundaries that belong to the surface of the objects. The boundary deformation in the top row is characterized by the emergence of a convexity, as an elliptic (bump-like) surface region 'intrudes' into the boundary contour. In the bottom row, a hyperbolic (saddle-like) surface region intrudes into the boundary, forming a new concavity.

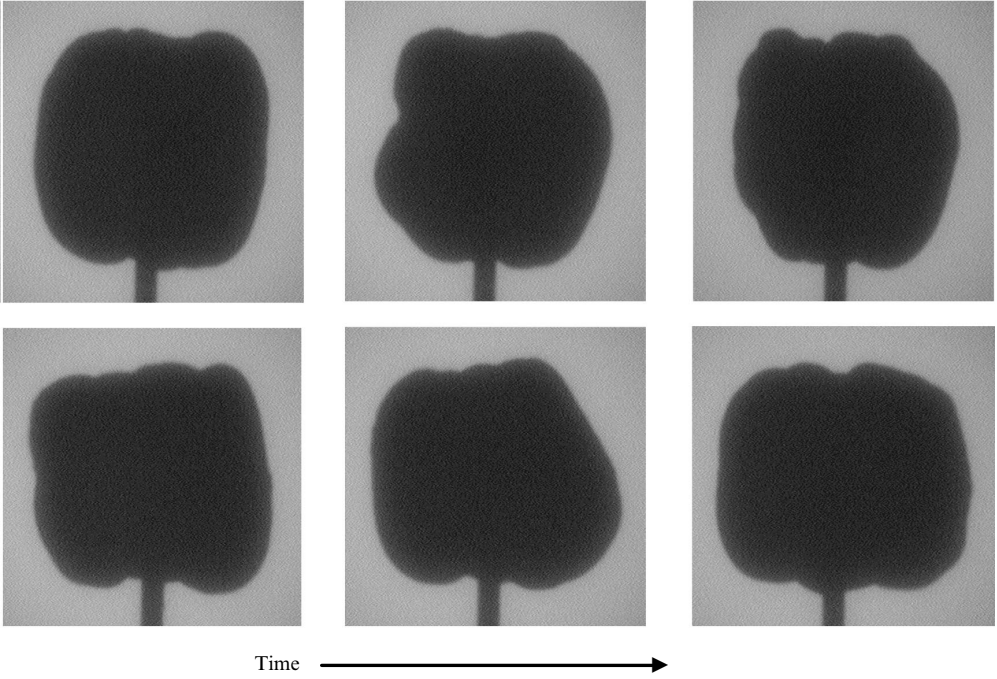


Figure 11. Examples of significant boundary changes or catastrophes that occurred in the moving cast shadows used in the present investigation. The top row illustrates the appearance and disappearance of a prominent shadow concavity (visible in the middle panel, left side of the shadow), whereas the bottom row illustrates the appearance and disappearance of a prominent convexity (middle panel, right side). As described in the text, such convexities and concavities signal the presence of elliptic and hyperbolic regions on the surface of a 3-D object casting the shadow.

Such qualitative representations are known as aspect graphs (Koenderink and van Doorn 1978; Koenderink 1984b; Van Effelterre 1994). The deformations of cast shadows of an object are especially informative if the object rotates through a complete 360° angle relative to an observer, since in that case all surface regions eventually project to and pass through the boundary contour at some moment in time. The entire qualitative structure of an object could thus be encoded in an aspect graph given only the information present in the deforming shadow boundary. From this point of view, one would expect that the viewing of deforming shadow boundary contours would lead to more accurate recognitions of objects than static shadow contours. In our experiment, the presence of deformation in our displays did lead to much higher rates of object identification and recognition. Much further research needs to be done, however, to evaluate whether human observers actually do represent 3-D shape in such a qualitative fashion.

In summary, deforming shadows appear to contain a wealth of information to not only support the perception of how solid objects move in depth (eg rotation or translation in depth—Todd 1985; Norman and Todd 1994; Kersten et al 1996), but they also are sufficiently informative to permit human observers to recognize naturally shaped objects in the absence of all other sources of optical information.

Acknowledgements. We would like to thank James Todd, Hideko Fukuda Norman, and Suzanne McKee for their helpful comments and suggestions regarding this experiment. We would also like to thank two anonymous reviewers for their advice and assistance.

References

- Arnold V I, 1984 *Catastrophe Theory* (Berlin: Springer)
- Atneave F, 1954 "Some informational aspects of visual perception" *Psychological Review* **61** 183–193
- Berbaum K, Bever T, Chung C S, 1984 "Extending the perception of shape from known to unknown shading" *Perception* **13** 479–488
- Biederman I, 1987 "Recognition-by-components: A theory of human image understanding" *Psychological Review* **94** 115–147
- Biederman I, Gerhardstein P C, 1993 "Recognizing depth-rotated objects: evidence and conditions for three-dimensional viewpoint invariance" *Journal of Experimental Psychology: Human Perception and Performance* **19** 1162–1182
- Biederman I, Ju G, 1988 "Surface versus edge-based determinants of visual recognition" *Cognitive Psychology* **20** 38–64
- Braunstein M L, Andersen G J, 1984 "Shape and depth perception from parallel projections of three-dimensional motion" *Journal of Experimental Psychology: Human Perception and Performance* **10** 749–760
- Cortese J M, Andersen G J, 1991 "Recovery of 3-D shape from deforming contours" *Perception & Psychophysics* **49** 315–327
- Erens R G, Kappers A M, Koenderink J J, 1993 "Perception of local shape from shading" *Perception & Psychophysics* **54** 145–156
- Gombrich E H, 1995 *Shadows: The Depiction of Cast Shadows in Western Art* (London: National Gallery Publications)
- Grafton C B, 1979 *Silhouettes: A Pictorial Archive of Varied Illustrations* (New York: Dover)
- Green B F, 1961 "Figure coherence in the kinetic depth effect" *Journal of Experimental Psychology* **62** 272–282
- Hayward W G, 1998 "Effects of outline shape in object recognition" *Journal of Experimental Psychology: Human Perception and Performance* **24** 427–440
- Hayward W G, Tarr M J, 1997 "Testing conditions for viewpoint invariance in object recognition" *Journal of Experimental Psychology: Human Perception and Performance* **23** 1511–1521
- Hoffman D D, Bennett B M, 1986 "The computation of structure from fixed-axis motion: Rigid structures" *Biological Cybernetics* **54** 71–83
- Hoffman D D, Richards W A, 1984 "Parts of recognition" *Cognition* **18** 65–96
- Kersten D, Knill D C, Mamassian P, Bühlhoff I, 1996 "Illusory motion from shadows" *Nature (London)* **379** 31
- Kersten D, Mamassian P, Knill D C, 1997 "Moving cast shadows induce apparent motion in depth" *Perception* **26** 171–192
- Koenderink J J, 1984a "What does the occluding contour tell us about solid shape?" *Perception* **13** 321–330
- Koenderink J J, 1984b "The internal representation of solid shape and visual exploration", in *Sensory Experience, Adaptation, and Perception: Festschrift for Ivo Kohler* Eds L Spillmann, B R Wooten (Hillsdale, NJ: Lawrence Erlbaum Associates) pp 123–142
- Koenderink J J, 1990 *Solid Shape* (Cambridge, MA: MIT Press)
- Koenderink J J, Doorn A J van, 1976 "The singularities of the visual mapping" *Biological Cybernetics* **24** 51–59
- Koenderink J J, Doorn A J van, 1978 "How an ambulant observer can construct a model of the environment from the geometrical structure of the visual inflow", in *Kybernetik 1977* Eds G Hauske, E Butenandt (Munich: Oldenbourg) pp 224–247
- Luce R D, 1963 "Detection and recognition", in *Handbook of Mathematical Psychology* volume 1, Eds R D Luce, R R Bush, E Galanter (New York: John Wiley) pp 103–189
- Mingolla E, Todd J T, 1986 "Perception of solid shape from shading" *Biological Cybernetics* **53** 137–151
- Norman J F, Lappin J S, 1992 "The detection of surface curvatures defined by optical motion" *Perception & Psychophysics* **51** 386–396
- Norman J F, Todd J T, 1994 "Perception of rigid motion in depth from the optical deformations of shadows and occlusion boundaries" *Journal of Experimental Psychology: Human Perception and Performance* **20** 343–356
- Norman J F, Todd J T, Phillips F, 1995 "The perception of surface orientation from multiple sources of optical information" *Perception & Psychophysics* **57** 629–636
- Richards W A, Koenderink J J, Hoffman D D, 1987 "Inferring three-dimensional shapes from two-dimensional silhouettes" *Journal of the Optical Society of America A* **4** 1168–1175
- Singh M, Seyranian G D, Hoffman D D, 1999 "Parsing silhouettes: The short-cut rule" *Perception & Psychophysics* **61** 636–660

- Tarr M J, 1995 "Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects" *Psychonomic Bulletin & Review* **2** 55–82
- Tjan B S, Braje W L, Legge G E, Kersten D, 1995 "Human efficiency for recognizing 3-D objects in luminance noise" *Vision Research* **35** 3053–3069
- Todd J T, 1985 "Perception of structure from motion: Is projective correspondence of moving elements a necessary condition?" *Journal of Experimental Psychology: Human Perception and Performance* **11** 689–710
- Todd J T, Akerstrom R A, Reichel F D, Hayes W, 1988 "Apparent rotation in three-dimensional space: Effects of temporal, spatial, and structural factors" *Perception & Psychophysics* **43** 179–188
- Ullman S, 1979 *The Interpretation of Visual Motion* (Cambridge, MA: MIT Press)
- Van Efferterre T, 1994 "Aspect graphs for visual recognition of three-dimensional objects" *Perception* **23** 563–582
- Vinci L da, 1989 "Light and shade", in *Leonardo on Painting* Ed. M Kemp (London: Yale University Press) pp 88–115
- Wallach H, O'Connell D N, 1953 "The kinetic depth effect" *Journal of Experimental Psychology* **45** 205–217