Perception, 2004, volume 33, pages 1319 – 1338

DOI:10.1068/p5323

Rapid processing of cast and attached shadows

James H Elder, Sherry Trithart, Gregor Pintilie, Donald MacLean
Centre for Vision Research, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada;
e-mail: jelder@yorku.ca; URL: http://elderlab.yorku.ca
Received 4 February 2004, in revised form 25 May 2004; published online 4 November 2004

Abstract. We used a visual-search method to investigate the role of shadows in the rapid discrimination of scene properties. Targets and distractors were light or dark 2-D crescents of identical shape and size, on a mid-grey background. From the dark stimuli, illusory 3-D shapes can be created by blurring one arc of the crescent. If the inner arc is blurred, the stimulus is perceived as a curved surface with attached shadow. If the outer arc is blurred, the stimulus is perceived as a flat surface casting a shadow. In a series of five experiments, we used this simple stimulus to map out the shadow properties that the human visual system can rapidly detect and discriminate. To subtract out 2-D image factors, we compared search performance for dark-shadow stimuli with performance for light-shadow stimuli which generally do not elicit strong 3-D percepts. We found that the human visual system is capable of rapid discrimination based upon a number of different shadow properties, including the type of the shadow (cast or attached), the direction of the shadow, and the displacement of the shadow. While it is clear that shadows are not simply discounted in rapid search, it is unclear at this stage whether rapid discrimination is acting upon shadows per se or upon representations of 3-D object shape and position elicited by perceived shadows.

1 Introduction

One of the greatest crime-solvers of the golden age of American radio had the habit of dematerialising into an ephemeral state as circumstances required. ‘The Shadow’ could be ‘seen, but not seen’. In some ways, this might apply just as well to the real shadows that are in the scene around you as you read this paper. They are evidence of things, but they are not things in themselves. As an observer you might wish to notice them so as to infer the locations, shapes, and identities of objects around you, but you would not want to confuse them with real objects, nor would you want them to distract your attention from the material objects of interest in the scene.

The ambivalent role of shadows in perception may partly explain why there is still much we do not understand about shadows, despite the many discoveries that have been made over the last twenty-five years (Mamassian et al 1998). These open questions include: What suffices for the perception of a shadow? How good are we at distinguishing different types of shadow? Are we sensitive to the global consistency of shadows in a scene? Which properties of shadows are most useful in making inferences about a scene? Are these inferences the result of high-level cognitive reasoning, or fast visual mechanisms operating in parallel across the visual field?

In this paper we consider two types of shadow: cast and attached. Cast shadows occur when one surface blocks light from reaching another. The manifold of light rays that just graze the obstructing surface sweeps out a (fuzzy) curve on the shadowed surface, defining the shadow boundary. The fuzziness or blur of the curve is determined by the angular subtense of the illuminant.

An attached shadow occurs when a surface curves away from a light source. Again, the manifold of light rays that just graze the surface defines the shadow boundary. Thus, attached shadows are 1-D features of 2-D shading fields. Whereas shading generally refers to modulations in surface luminance arising from variations in the surface normal vector with respect to the illuminant direction, attached-shadow boundaries occur specifically when these two vectors are orthogonal.

Here, we used psychophysical methods to study the rapid perception of both types of shadow. Our method was to create a very simple stimulus that can produce a compelling visual impression of a shadow, and then manipulate properties of the stimulus to determine which properties we are sensitive to. In particular, we examined: (i) the discrimination of attached and cast shadows; (ii) the detection of illumination inconsistencies from cast shadow cues; (iii) discrimination based on depth from penumbral blur; and (iv) discrimination based on depth from shadow displacement.

We employed a visual-search methodology that has been used in other studies of 3-D perception (Enns and Rensink 1990; Rensink and Cavanagh 1993, 1998; Sun and Perona 1996). In this standard method, observers are asked to find a unique target stimulus in a random array of distractor stimuli. The focus of most visual-search studies has been on the rate of increase in response time as a function of the number of items in the display (the 'search slope').

There are two main ways in which these search slopes have been interpreted. In the traditional view, zero or near-zero search slopes are taken to suggest that target and distractor differ in some fundamental 'pre-attentive' feature, whereas positive search slopes suggest the absence of such a discriminating feature, and the need for focused attention to discriminate the stimuli (Treisman and Gelade 1980).

A more recent view is that search slopes reflect target/distractor similarity when projected onto a normalised internal representational space used by the early visual system (Verghese and Nakayama 1994). Thus, low search slopes suggest that target and distractor project onto points that are quite distant in this space, whereas high search slopes suggest that the two stimuli project onto nearby points.

In either case, search slope can be taken to reveal the degree to which rapid visual processing is tuned to represent and discriminate the properties that distinguish target from distractor. In our experiments, we created simple target and distractor stimuli that differed only in their shadow properties, and used the visual-search methodology to infer the degree to which the human visual system is tuned to represent and exploit shadows and their properties.

One of the key challenges in the psychophysical study of shadows is distinguishing between the effects of shadows as interpreted properties of a 3-D scene, from the effects of shadows as contrast modulations in the 2-D image. Our solution was to make use of the physical requirement that shadows be darker than the surrounding region of the surface on which they are cast. Earlier work suggests that the human visual system respects this requirement: 3-D shape perception is largely eliminated when appropriate contrast rules are not obeyed at shadow boundaries (Cavanagh and Leclerc 1989). Thus, in order to discriminate the 3-D effects of shadows from the 2-D effects of contrast modulations, we organised our psychophysical experiments into pairs of conditions. In one, shadows were rendered as (physically consistent) dark contrast modulations, and in the other, as (physically inconsistent) light modulations. By comparing search speed across these two conditions, we can effectively subtract out the 2-D component of the discrimination and isolate the 3-D component of shadow processing.

Light shadows have been shown to be less effective than dark shadows in generating perceptions of motion in depth (Kersten et al 1996), and more disruptive to 3-D object recognition (Tarr et al 1998). Rensink and Cavanagh (1993, 1998) have shown that light shadows are processed differently from dark shadows in a visual-search task. Thus, we have some confidence that light shadows form a good 2-D control stimulus. On the other hand, in their striking 'ball-in-the-box' experiments, Kersten et al (1997) found that light shadows were quantitatively as effective as dark shadows in determining the perceived 3-D motion of the object casting the shadow. Thus, it is possible that the light shadows used in our experiments may elicit some 3-D response in the visual system. Any differential effects between dark-shadow and light-shadow conditions found in our experiments should therefore be considered conservative estimates of the 3-D role played by cast shadows in stimulus discrimination.

To summarise, we suggest that differences between target and distractor stimuli that are not particularly salient when the stimuli are perceived only as 2-D contrast modulations may become highly salient when the visual system interprets the stimuli as shadows and objects in a 3-D scene. Thus, if shadows do play a role in forming the representations upon which rapid discrimination is based, then we expect that search slopes will be smaller for dark-shadow stimuli than for matched light-shadow stimuli. If, on the other hand, the interpretation of shadows is a slower, high-level, cognitive process, we would expect little difference between reaction times for dark-shadow and light-shadow stimuli.

In the next section, we briefly review what is already known about the rapid inference of 3-D scene properties, particularly shadows. We then describe our methodology and results of five experiments examining the role of shadows in rapid visual processing. Finally, we discuss the implications of these results for our understanding of the visual representations underlying rapid scene perception.

2 Earlier work

2.1 *Estimation of 3-D scene properties*

The retinal image projecting from any given scene is highly dependent upon the direction of the illuminant. Gibson (1950) pointed out that the perception of shape from shading in particular must depend upon estimation of the illuminant position. Berbaum et al (1983) provided evidence that human observers can use remembered light-source information to solve the concave/convex shading ambiguity that arises when viewing simple curved surfaces. In a later study, they found that even light-source information that is provided indirectly, via the shading on a familiar, unambiguous object or by the presence of cast shadows, can determine the perception of simultaneously presented ambiguous surfaces (Berbaum et al 1984).

These findings suggest that the perception of shape from shading involves the computation of a globally consistent visual scene. An even stronger constraint, suggested by Ramachandran (1988), is that the visual system assumes there is only one light source illuminating the entire visual scene. Ramachandran based this idea on his experiments with arrays of simple ambiguous shaded stimuli. He noted that, when stimuli with opposite directions of shading are mixed, there are two possible interpretations: (i) that the objects have the same sign of curvature but are receiving illumination from opposite directions, or (ii) that the objects have different signs of curvature and are receiving the same illumination. He found that observers invariably reported the latter, perceiving a mixture of object shapes, under a single illuminant. On the other hand, Cavanagh (1999b) has argued that there are many examples of impossible and inconsistent shading and shadows in visual art, yet these works of art seem not to disturb the casual observer.

Observers in Ramachandran's experiments also reported that when only one shape was reversed in shading, that shape seemed to pop-out by virtue of its different 3-D appearance; that reaction time is independent of the number of shapes in a display was later confirmed (Kleffner and Ramachandran 1992). These results and others suggest that 3-D scene properties including shape-from-shading and direction of lighting are computed rapidly and in parallel (Enns and Rensink 1990).

2.2 *The role of cast shadows in scene perception*

Yonas (1979) provided a striking example of how a cast shadow can influence the perceived location of an object relative to the surface on which the shadow is cast. Cunningham et al (1996) studied this quantitatively, using a stimulus consisting of naturalistically rendered rocks. Observers were asked to report whether one of the rocks was in front of the others; cast shadows were found to be helpful in speeding search for the foreground rock. Madison et al (2001) have shown that cast shadows are an effective cue for the judgment of surface contact. Kersten and colleagues have provided two powerful demonstrations of how moving cast shadows can influence our perception of object motion in depth (Kersten et al 1996, 1997).

It has also been suggested that cast shadows might help to disambiguate the shape of the surface on which they are cast (Cavanagh and Leclerc 1989). However, Mamassian et al (1998) have since demonstrated a case in which cast shadows fail to perceptually resolve a convex/concave ambiguity in the surface on which they are cast. While geometrically the shadows are only consistent with one interpretation, they turn out to have little effect on human perception.

While many studies suggest a helpful role for shadows in scene estimation, Rensink and Cavanagh (1993, 1998) have suggested that, to facilitate image interpretation, cast shadows may be rapidly discounted by the visual system. In a series of experiments, they found that visual search was slower when the feature distinguishing target from distractor could be perceived as a cast shadow. They argue that, while the visual system identifies shadows rapidly and in parallel across the visual field, once identified, the properties of these shadows (eg orientations) are made less readily accessible to the visual system, as though shadows were nuisance variables to be filtered out early in visual processing.

2.3 *Cast shadows and object recognition*

Tarr et al (1998) studied the effects of shadows and changing illumination conditions on the recognition of nonsense objects. They found that the addition of strong cast shadows had no impact on the reliability of recognition, even when observers were required to recognise objects seen under very different illumination conditions. Even more strikingly, the inclusion of cast shadows actually speeded the recognition response, even for recognition across different illuminations. Tarr et al concluded that rather than confounding object recognition, cast shadows serve to disambiguate 3-D shape, thus aiding recognition. Elaborating this result, Castiello (2001) has shown that object recognition is slower with incongruent lighting and shadows. However, the story is still not entirely clear. Using natural objects, Braje et al (2000) found no effects of shadows on either the reliability or speed of recognition.

To summarise, a substantial amount has been learned about shadows in the last twenty-five years, but there is still considerable debate and much we do not know. It seems clear that at least in simple scenes there is some degree of rapid global processing in perceiving shape from shading, and this appears to involve a (perhaps implicit) estimation of light-source direction. However, the degree to which the interpretation of more complex scenes involves global-consistency computations and light-source estimates is in doubt.

Cast shadows appear to have a role to play in the perception of relief (elevation relative to a ground surface), surface contact, and motion in depth, although their role in object recognition is less certain. In the next five experiments we explore in greater detail the role of both cast and attached shadows in the perception of simple scenes.

3 Methods

3.1 *Stimuli and apparatus*

Search stimuli were created with the Khoros software development environment from Khoral Inc. The stimuli were based on manipulations of a simple crescent shape, created by multiplicative combination of two image layers. Each of the layers contained a disk roughly 3.7 deg in diameter. In the first layer, pixels interior to the disk were assigned the value -1, while exterior pixels were assigned the value 0. In the second layer, interior pixels were assigned the value 0, and exterior pixels were assigned the value $+1$. The second disk was shifted by 26 min of arc relative to the first disk, in a direction 30° counterclockwise from vertical. Multiplying the two layers effectively caused the first disk to be partially occluded by the second disk, producing a crescent stimulus (figure 1, top). The stimulus was then scaled to a 256 grey-level gamut so that the background was mid-grey and the crescent was black.

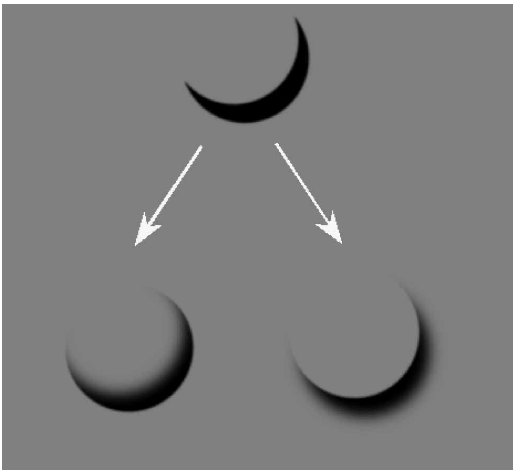

Figure 1. Creation of illusory objects by selective blurring. See text for details.

The crescent stimulus was modified by selectively blurring either layer with a Gaussian blur kernel. Convolution of the first layer with a Gaussian blur kernel causes the outer arc of the crescent to be blurred (figure 1, bottom right), while convolution of the second layer causes the inner arc to be blurred (figure 1, bottom left). The space constant (standard deviation) of the blur kernel was 22 min of arc in all experiments, unless otherwise noted.[1] This simple blurring operation can be seen to elicit powerful 3-D percepts of object shape and relief. The blurred inner arc of the stimulus on the left is seen as an attached shadow lying on a smoothly curved ball, while the blurred outer arc of the stimulus on the right is seen as the boundary of a shadow cast by a flat disk that is raised above a ground surface.[2] We call the resulting stimuli "illusory objects",

(1) This value was found to produce compelling percepts of 3-D shape. However, the precise value of the blur constant does not seem to be crucial: a broad range of blur scales near the chosen value produce qualitatively similar percepts.

(2) When printed, the blurred contours are often artifactually sharpened in an uneven manner. For the stimulus in the bottom right of figure 1, this can lead to the percept of a cylindrical object standing out from a ground surface, with part of the dark region interpreted as an attached-shadow region projecting from the side of the cylinder. This artifact is not present when the stimuli are viewed on a monitor. Under these conditions, the shadow appears to lie on flat ground surface, not on the side of a cylinder, and our subjects generally report the former percept. Note that the cylinder interpretation implies a non-generic (accidental) view, in which the illuminant direction and viewing direction lie in a common plane, with the illuminant from the top left and the viewing direction from the bottom right relative to the normal of the top of the cylinder. Otherwise we should see vertical isophotes on the side of the cylinder, ie part of the side of the cylinder should be in light and part in darkness.

since the contrast of the perceived surface is identical to that of the background. White-shadow control stimuli were produced by switching the sign of disk pixels in the first layer.

Visual-search displays were created within a 37 deg $\times$ 28 deg search window. Items were placed pseudorandomly in sequence. For each item, candidate locations were generated until a location that did not overlap a previously placed item was found.

We tested set sizes of 4, 9, and 14 items. We report the data as a function of the number of distractors in the display (for the target-present condition): 3, 8, and 13. Stimuli were displayed on a 17-inch 1024 $\times$ 768 pixel Sony Trinitron CRT. Observers sat roughly 50 cm from the display, in a dimly lit room.

3.2 *Procedure*

We employed a standard visual-search paradigm (Treisman and Gelade 1980). On each trial, the observer was presented with a search display containing a fixed number of search items, either all the same, or with one different 'target' item. The target appeared in the display with probability 0.5. Observers were asked to indicate, as rapidly as possible, whether they observed the target item to be present (left mouse click) or absent (right mouse click), while trying to keep their error rate below 10%. Incorrect responses were signaled by an audible beep.

In this study we report the results of five experiments in which we examined rapid stimulus discrimination based upon shadow properties. Each experiment consisted of 4 – 8 conditions involving different stimulus manipulations. Each condition involved discrimination of a pair of distinct stimuli, and was matched with a second condition in which the target and distractor stimuli were exchanged, permitting an analysis of search asymmetries (Treisman and Souther 1985). Each condition consisted of 210 trials for each subject, 70 at each of 3 set sizes. Each condition was divided into two blocks of 105 trials each, 35 at each of the 3 set sizes.[3]

A randomised block design was employed to minimise order effects. Each experiment was split into two blocks. Each of the blocks consisted of 35 trials at each set size, for each condition within the experiment. The order of conditions within blocks and set sizes within conditions was randomised, as was the ordering of blocks across experiments.

To summarise, each observer completed all 5 experiments twice, in 2 separate blocks. The order of experiments was random in both blocks. Within each experiment, conditions were presented in random order, and within each condition, the 3 set sizes were presented in random order.

At the beginning of each experiment block, observers completed a training session in which they performed the task for one trial for each experiment, condition, and set size. At the beginning of each condition block, observers were presented with a display showing one target and one distractor placed side-by-side and labeled 'T' and 'D', so that observers had a good mental image of what they were looking for. Observers pressed a mouse button to initiate the 105 trials for that block.

Instructions to the observers referred to the stimuli as 'items'. No mention was made of shapes, shadows, or shading.

3.3 *Observers*

Ten student observers, naïve to the purpose of the experiments, participated in the study. All had normal or corrected-to-normal vision.

3.4 *Analysis*

Data for target-absent and target-present conditions were analysed separately. Error rates for the target-absent condition were generally lower than 5%. However, error rates for difficult conditions in the target-present case rose in some cases to over 20% with

[3] Note that reversal of target and distractor occurred across conditions. Within a set of 105 consecutive trials, the target and distractor remained fixed.

increasing set size. This is a known phenomenon with the standard visual-search method that has been shown to lead to negative biases in the estimation of search speed for difficult conditions (Elder and Zucker 1993). Since in this study our interpretations are primarily based upon comparisons between pairs of conditions, this bias serves to reduce the apparent power of our statistical tests, increasing the actual reliability of any conclusions we draw. Two-tailed t-tests have been used to test for statistical significance, at a significance level of 0.05.

As our main indicator of search performance we used the slope of the maximum-likelihood linear fit of the response-time data as a function of set size (Treisman and Gelade 1980) for the target-present trials. We also report the y intercept (extrapolated response time for 0 distractors).

For brevity, we report detailed data only for target-present conditions. However, we also analysed the ratio of the search slope for target-absent trials over the search slope for target-present trials. A ratio of 2 is consistent with a serial search process that on average processes half of the search items in target-present conditions and all of the items in target-absent conditions. This ratio can thus be a useful test for serial search. We summarise the results of this analysis in section 9.5.

4 Experiment 1: Blur discrimination control

Our primary approach to distinguishing 3-D shadow effects from 2-D contrast effects was to compare search times for dark-shadow stimuli and for contrast-reversed light-shadow stimuli. The purpose of our first experiment was to rule out more general light/dark asymmetries that might have confounded this inference in later experiments.

In these conditions, both target and distractor are simple crescent stimuli (figure 2). Discrimination is based upon a difference in blur. For either the target or distractor, both inner and outer arcs are blurred with a Gaussian blur kernel with a space constant of 7 min of arc, while the other stimulus remains sharp.[4] Note that, when both

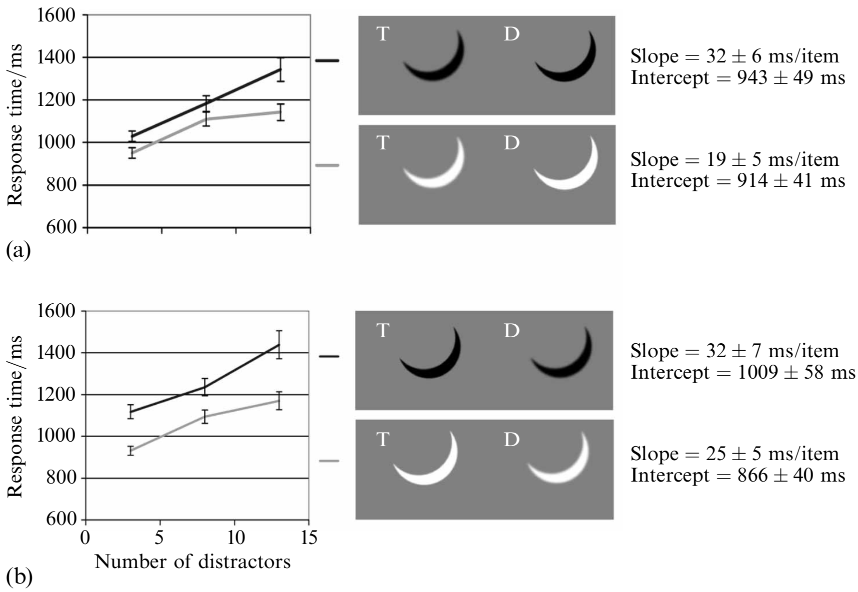

Figure 2. Experiment 1. Blur discrimination control—stimuli and results: (a) blurred targets, (b) sharp targets.

[4] A relatively small amount of blur was employed to avoid floor effects: finding a blurry target amongst sharp distractors or vice-versa is a relatively easy task.

inner and outer arcs are blurred, no strong 3-D percept is elicited: discrimination is likely to be based upon 2-D representations.

We tested both dark and light stimuli, with either the target or distractor blurred. The results are shown in figure 2. All search slopes were found to be positive ($p < 0.0002$), generally in the range of 19 – 39 ms per item. No significant target/distractor search asymmetries were observed ($p > 0.4$), and generally no significant differences in search speed for dark and light stimuli were observed ($p > 0.09$), except in one case (figure 2b): for sharp targets, the baseline reaction time (y intercept of the regression line) for light stimuli was significantly lower than that for dark stimuli ($p < 0.05$).

This difference will not confound the interpretation of our later results, since generally we will be trying to determine whether search is faster for dark stimuli than for light stimuli. Thus this slight difference in intercept will serve to increase the reliability of any statistical inferences we make in later experiments.[5]

We conclude from our first experiment that there is no general bias to discriminate blur for dark stimuli more rapidly than for light stimuli, when these stimuli are perceived as 2-D patches, and not as shadows in a 3-D scene.

5 Experiment 2: Cast and attached shadows

While blurring both arcs of the crescent stimulus does not elicit a strong 3-D percept, figure 1 shows that selective blurring of only one of the arcs generates distinctive 3-D percepts: of a curved 'ball' surface with an attached shadow, or of a flat 'disk' surface casting a shadow. The fact that such a minor manipulation could give rise to such powerful differences in perception suggests that our visual system is highly tuned to discriminating cast and attached shadows for the purpose of inferring 3-D scene properties. But is this 3-D discrimination a reflection of 'high-level' cognitive reasoning, or a more rapid, automatic process?

Our second experiment addresses this question. We tested discrimination of these cast-shadow and attached-shadow stimuli, using a Gaussian blur constant of 22 min of arc, in both dark and light contrast, and both target/distractor assignments. Our main prediction was that if the visual system can rapidly discriminate between cast and attached shadows, search slopes will be smaller for dark-shadow stimuli than for light-shadow stimuli.

Example search displays are shown in figure 3, and results are summarised in figure 4. The data clearly show a profound difference in the speed of search for dark-shadow and light-shadow stimuli. Search for the attached-shadow target in cast-shadow distractors was fast with dark shadows, with search slope not significantly different from zero ($p > 0.3$). Search slope for the cast-shadow target in attached-shadow distractors was small (5.7 ms per item) but significant ($p < 0.004$). Search for the corresponding stimuli in light shadows was much slower, both in terms of search slope ($p < 0.000001$) and intercept ($p < 0.004$).

Thus, it appears that cast and attached shadows can be discriminated very rapidly and in parallel across the visual field, presumably to support the rapid construction of a 3-D representation of the visual scene. The fact that search slopes are zero or near-zero suggests that this discrimination does not require high-level cognitive reasoning or focused attention.

Recall that the target and distractor stimuli differ only in whether the inner or outer arc is blurred. Thus the concept of 'inside' and 'outside' (figure/ground) seems crucial here. To test this, we created a new control stimulus, formed by conformally mapping the circular stimuli onto a straight line (figure 5). The removal of closure

[5] Although not significant, the data also show a trend to lower search slopes for the light stimuli than the dark stimuli. Again, this will serve to increase the reliability of any statistical inferences we make about the rapid processing of shadows.

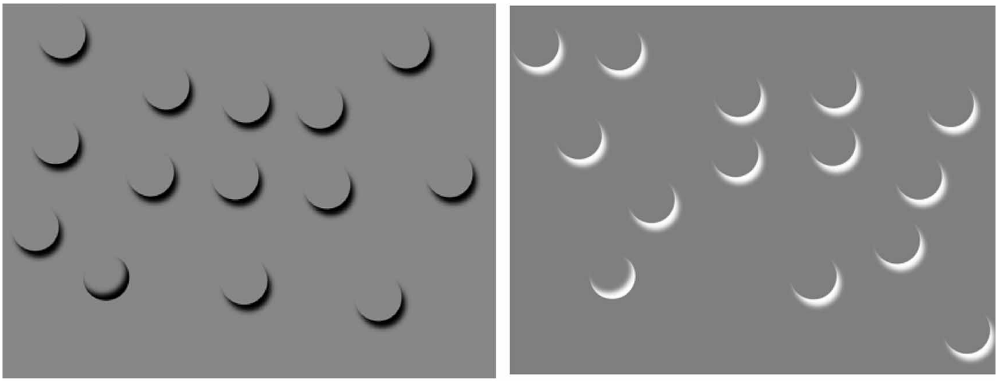

Figure 3. Experiment 2. Discriminating cast and attached shadows: example search displays.

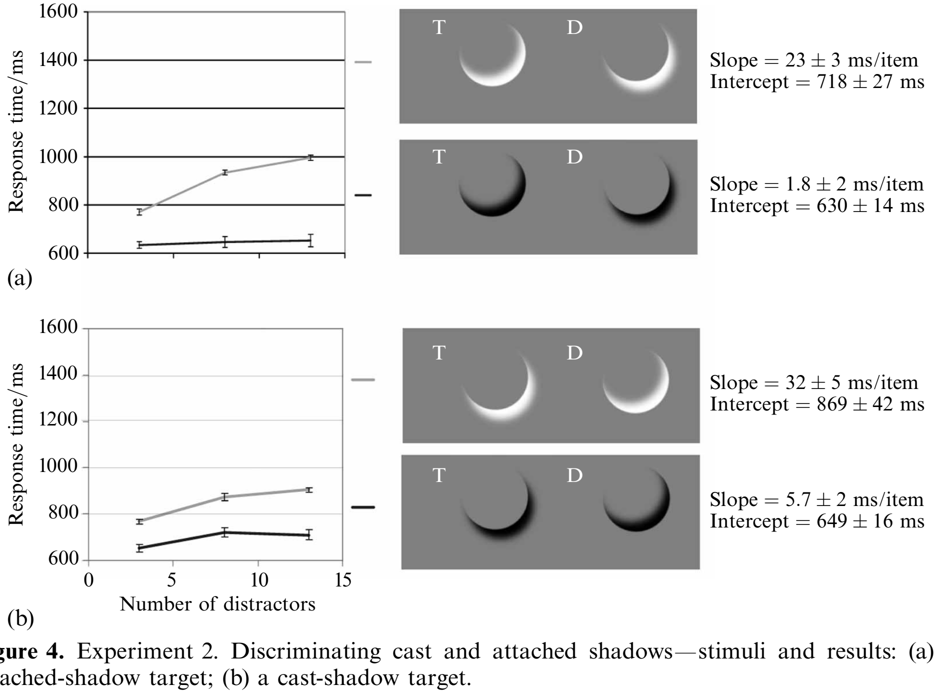

Figure 4. Experiment 2. Discriminating cast and attached shadows—stimuli and results: (a) an attached-shadow target; (b) a cast-shadow target.

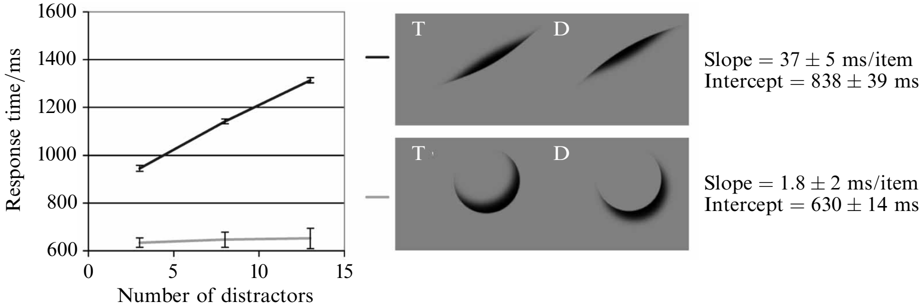

Figure 5. Experiment 2. Discriminating cast and attached shadows: control condition. When the attached-shadow and cast-shadow stimuli are conformally mapped onto a straight line, search becomes very slow.

information reduces the difference between target and distractor to a difference in 2-D orientation, and removes much of the phenomenological impression of 3-D shape. Correspondingly, both search slope and intercept rose significantly ($p < 5 \times 10^{-7}$). Thus, it is the combination of blur and closure information in these simple stimuli that is crucial to creating the powerful 3-D percepts that can drive rapid discrimination.(6)

6 Experiment 3: Illumination inconsistencies

Experiment 2 shows that shadows are not simply discounted in rapid scene perception, but that cast shadows and attached shadows may be rapidly discriminated to produce strikingly different 3-D percepts. What other properties of shadows might the visual system be sensitive to?

One issue of some debate is the degree to which the human visual system is sensitive to the consistency of shadows with a physically realistic illuminant. Cavanagh (1999a) has suggested that the visual system is rather insensitive to this information, citing numerous examples of illumination inconsistencies in visual art that appear to go unnoticed by observers. In our third experiment, we attempted to test our sensitivity to illumination inconsistencies, as revealed by the shadows cast by our simple illusory 'disk' stimuli.

In our first attempt, we simply measured search speed for discrimination of upright illusory disks from illusory disks rotated by 180° (figure 6). We found, however, that search was very fast for both dark and light shadows, with search slopes insignificantly different from zero ($p > 0.2$), or, in one case, significantly less than zero ($p < 0.04$). These results suggest a 'floor' effect; perhaps the stimuli are too easily discriminable on the basis of 2-D configuration (eg 'smiles' versus 'frowns'). We noted as well that inverted

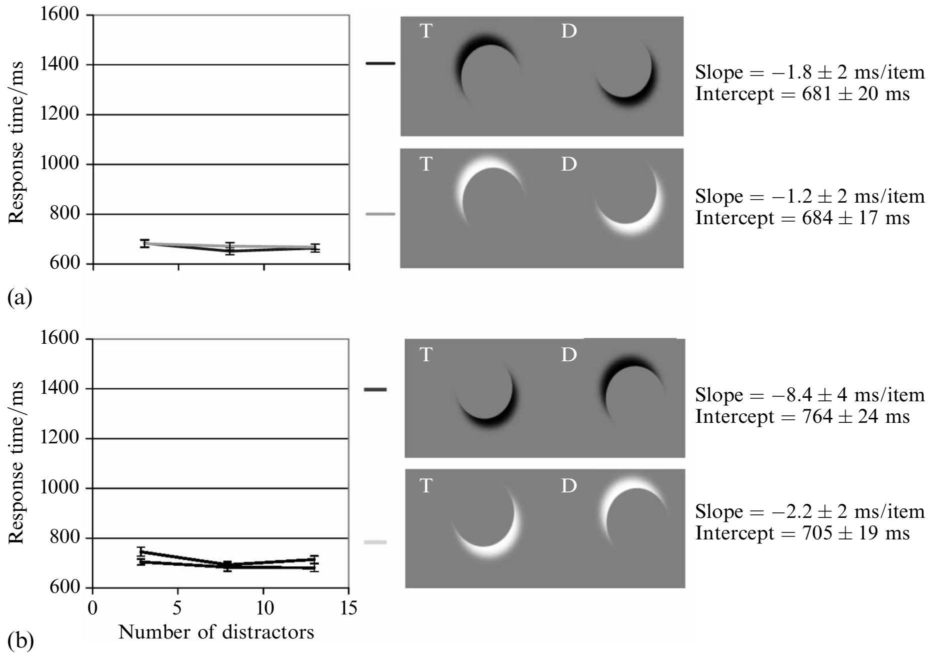

Figure 6. Experiment 3. Illumination inconsistencies for illusory objects—stimuli and results: (a) inverted targets; (b) upright targets.

(6) The sharp contours bounding the perceived objects in our stimuli are not completely closed. However, earlier research suggests that, perceptually, contour closure is a graded property: partial closure can be very effective in speeding the perception of shape (Elder and Zucker 1993).

disk stimuli with dark shadows could also be perceived as concave depressions in the ground surface, flipping the assignment of the blurred contour from cast shadow to attached shadow.

To overcome these problems we modified the disk object in the foreground layer to differ in contrast from the background (figure 7). This served, phenomenologically at least, to induce the completion of the shadows behind the occluding disks, making the 2-D 'smile' versus 'frown' cue much less salient. Reducing the salience of this 2-D cue reveals significant differences between search for illumination consistencies in dark and light shadows (figure 8). Whereas search slopes for dark shadows are not significantly different from zero ($p > 0.1$), search slopes for light shadows are significantly greater than zero ($p < 0.00002$), and significantly greater than search slopes for dark shadows as well ($p < 0.006$). Note that the introduction of non-zero contrast between foreground and background now removes the interpretation of the inverted

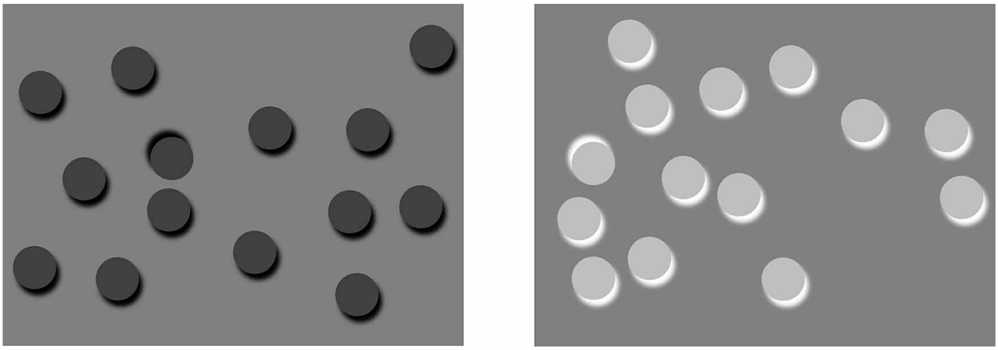

Figure 7. Experiment 3. Illumination inconsistencies for real objects: example search displays.

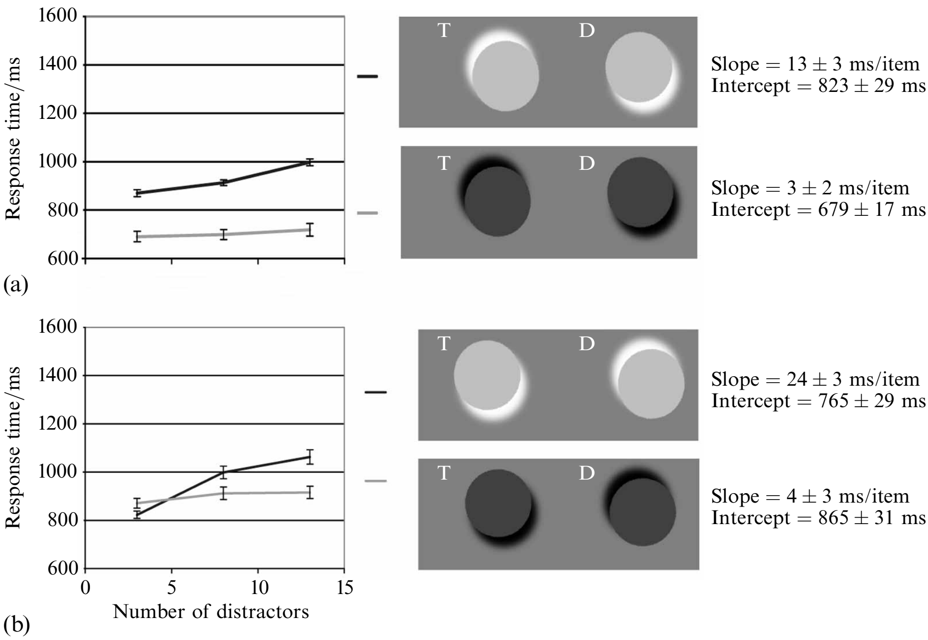

Figure 8. Experiment 3. Illumination inconsistencies for real objects—stimuli and results: (a) inverted targets, (b) upright targets.

stimuli as concave depressions, leaving the illuminant inconsistency as the most probable causal factor.[7]

Thus, we conclude that, counter to prior suggestions (Cavanagh 1999a), the visual system can rapidly detect illumination inconsistencies from cast shadows. This result extends earlier results showing rapid detection of illumination consistencies from shading (attached shadows) (Enns and Rensink 1990).

7 Experiment 4: Depth from shadow blur

In his shape-from-shading experiments, Ramachandran (1988) noted that the pop-out observed for reverse-shaded targets seemed to vanish when the smooth shading gradients were changed to abrupt luminance edges. This suggests that some amount of blurring may be a precondition for the perception of an attached shadow on a smooth surface. However, Enns and Rensink (1990) found illumination sensitivity in visual search for simple polyhedral objects with sharp attached shadows (illumination discontinuities). This suggests that the perception of 3-D shape depends upon consistency between contour and shading cues.

These results pertain to attached shadows. What is the role of blur in the perception of cast shadows? Under natural illumination conditions, cast shadows are often blurred. This blur is actually the shadow penumbra: that region of the ground surface receiving only partial illumination. Penumbral blur will arise for all non-point light sources, ie whenever the light source has some non-zero angular subtense. For example, the sun subtends roughly 0.5 deg at the earth's surface. This means that, when you are out on a sunny day, your head will cast a shadow whose penumbra also subtends rough 0.5 deg on your retina, corresponding to a blur kernel with space constant of roughly 15 min of arc, much more than the width of the point spread function of a 2 mm pupil (roughly 1 min of arc), and well above blur detection threshold (0.3 min of arc—Watt and Morgan 1983).

In general, the extent of penumbral blurring is primarily a function of two variables: the angular extent of the light source, and the distance between the casting object and the surface on which the shadow is cast. Penumbral blur increases as the light source is made larger or the casting object is lifted off the ground surface. This raises the possibility that the degree of penumbral blur might serve as a cue to object relief (elevation relative to the ground surface).

That the visual processing of shadows depends in some ways on blur is evident: in our basic stimulus (figure 1), blur is required in order to see shadows. Kersten et al (1996) found that the addition of penumbral blurring to moving shadows can enhance the impression of object motion in depth. Also note that while judgment of absolute relief is confounded by uncertainty in the extent of the light source, the relief of one object relative to another, on the assumption that they are illuminated by the same light source, might reasonably be estimated.

To test this idea, we modified the stimuli of experiment 3 (figure 7) to vary in their degree of penumbral blur by varying the space constant of the Gaussian blur kernel. We tested two levels of blur discrimination: blur of 0 versus blur of 22 min of arc, and blur of 22 min of arc versus blur of 44 min of arc; example search displays of the latter are shown in figure 9. We tested these two levels of blur discrimination for dark and light shadows, and for either target or distractor assignment, yielding a total of 8 conditions. Our prediction was that, if penumbral blur is used as a rapid cue to object relief, search slopes for dark-shadow stimuli should be smaller than for light-shadow stimuli.

[7] With a dedicated, precisely directed, collimated light source for each object, the image shown in figure 7 could be produced from a real 3-D scene. Thus, the scene is not technically impossible, just highly improbable.

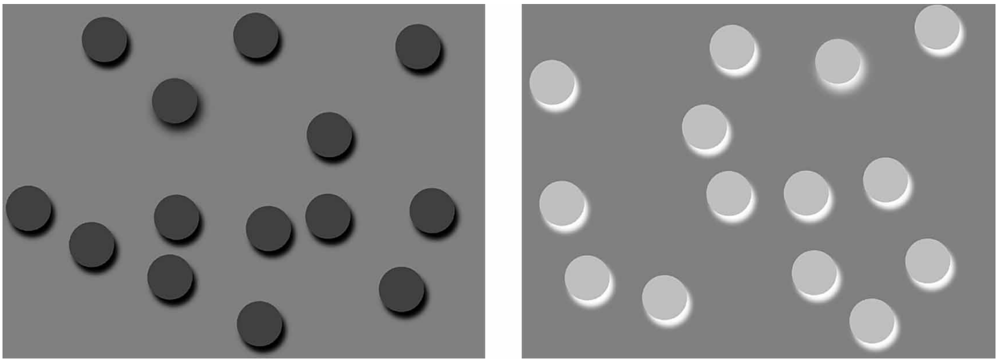

Figure 9. Experiment 4. Depth from shadow blur: example search displays. The blur constants used for these displays were 44 min of arc for the target and 22 min of arc for the distractors.

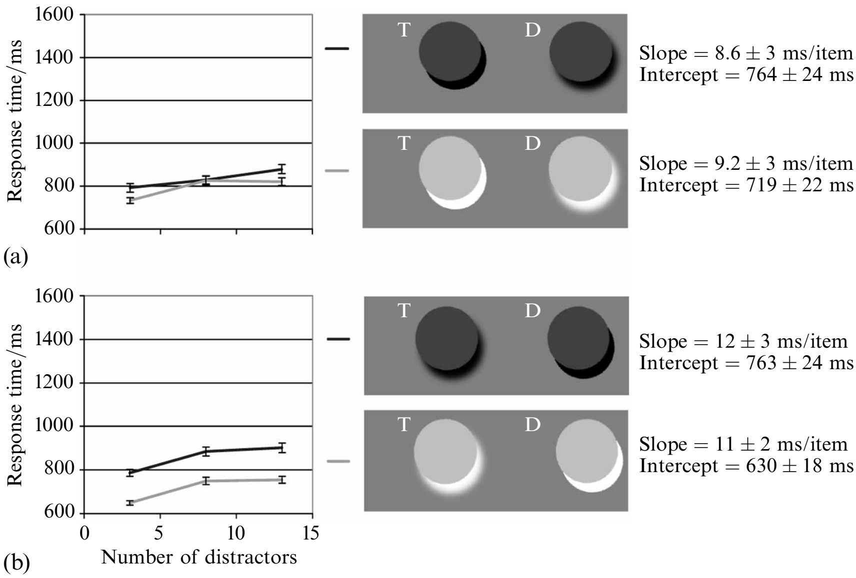

Figure 10. Experiment 4. Depth from shadow blur—stimuli and results: (a) target blur = 0 min of arc, distractor blur = 22 min of arc; (b) target blur = 22 min of arc, distractor blur = 0 min of arc.

The results of this experiment are summarised in figures 10 and 11. We found no significant differences between search slopes for dark and light shadows ($p > 0.09$). Thus, although it seems clear that we use shadow blur in some ways to rapidly process visual scenes, it is possible that blur is not used as a rapid metrical cue for object relief.

8 Experiment 5: Depth from shadow displacement

Blur is not the only shadow property that might be useful to estimate relative object relief. As an object is lifted off the surface, the shadow boundary does not only blur, it also displaces in the scene, and hence typically in the image. Since judgment of position might be considered a more fundamental visual capacity than judgment of blur, it is possible that our visual system may be able to use shadow displacement to judge the relative depth of objects casting shadows, even if it does not appear to use blur. Indeed, powerful evidence for this capacity has been demonstrated (Kersten et al 1996, 1997). Our goal here is to test whether this capacity exists in rapid scene discrimination.

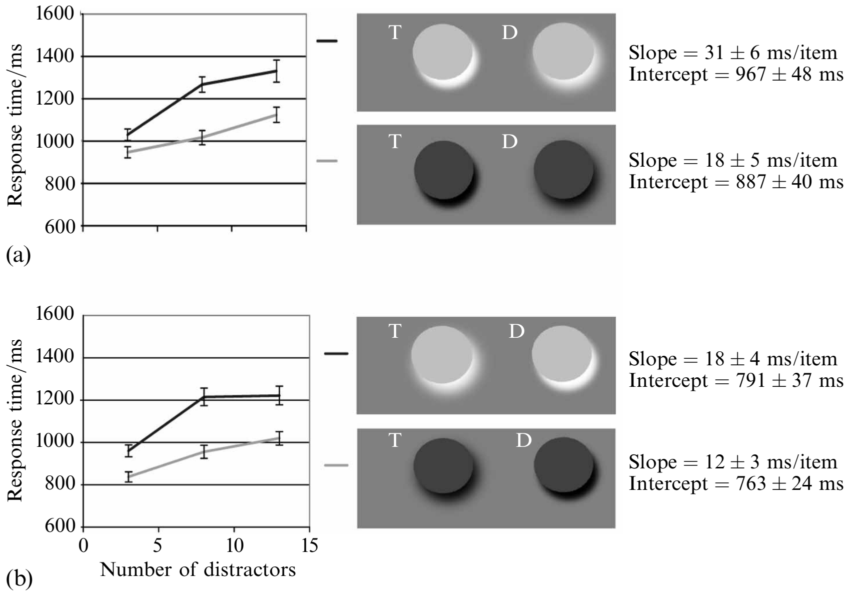

Figure 11. Experiment 4. Depth from shadow blur—stimuli and results: (a) target blur = 22 min of arc, distractor blur = 44 min of arc; (b) target blur = 44 min of arc, distractor blur = 22 min of arc.

To this end, we tested stimuli with shadow displacements of 26 min of arc against stimuli with shadow displacements of 44 min of arc, in both dark-shadow and light-shadow conditions. Example search displays are shown in figure 12, and results are summarised in figure 13.

Figure 13a shows a substantial asymmetry for dark shadows: search is faster when the target shadow is more displaced ($p < 2 \times 10^{-12}$). When it is the distractor shadow that is more displaced as in figure 13b, there is no significant difference in search time for dark and light shadows ($p > 0.3$). However, figure 13c shows that when the target shadow is more displaced, search is significantly faster for dark shadows ($p < 0.003$).

What could all of this mean? We think that there are in fact two factors acting at once in this experiment. The first is a 2-D factor: displacement of the shadow causes the overall stimulus to be larger, providing a strong 2-D size cue. Cavanagh and Arguin (1990) explored the role of stimulus size as a visual-search cue for stimuli defined by various attributes (luminance, colour, etc). Experiments with luminance-defined disks

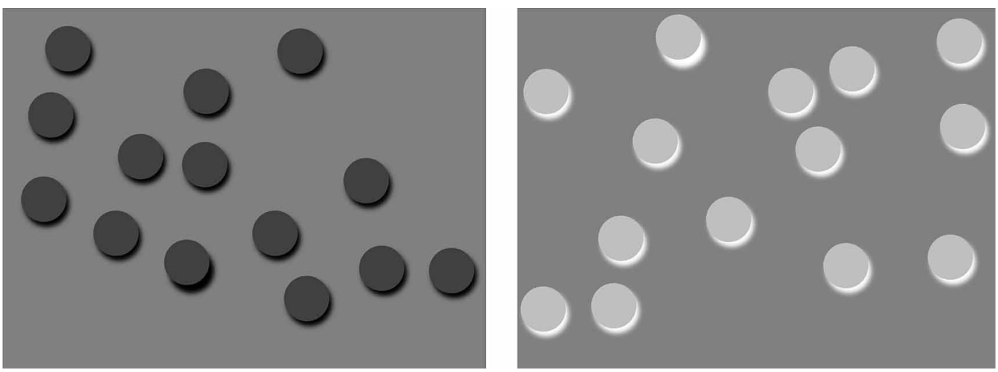

Figure 12. Experiment 5. Depth from shadow displacement: example search stimuli. Target shadow displacement = 44 min of arc, distractor shadow displacement = 26 min of arc.

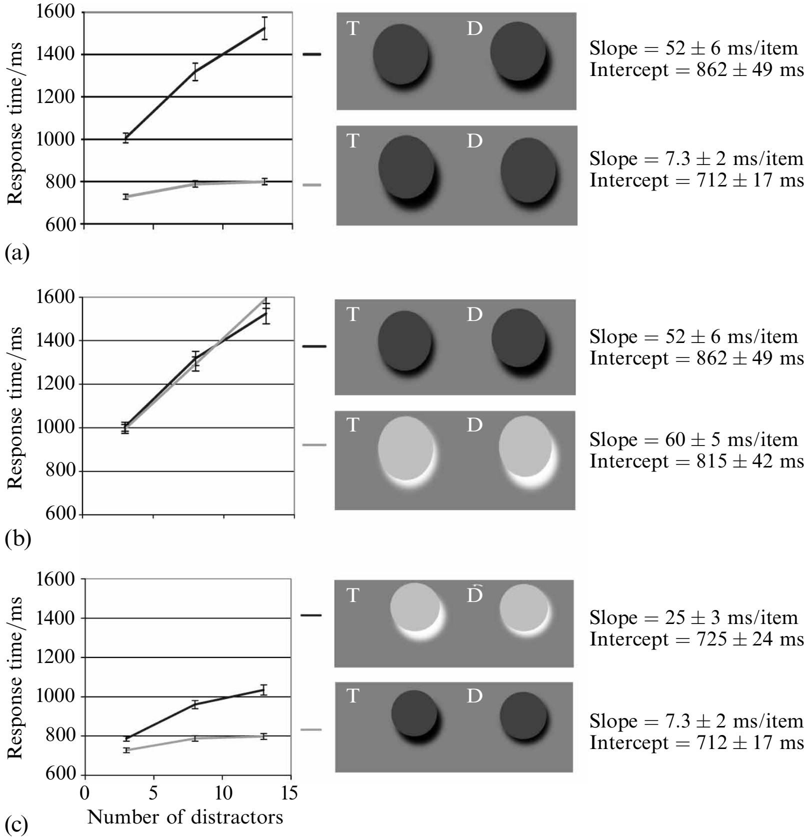

Figure 13. Experiment 5. Depth-from-shadow displacement: stimuli and results. Four conditions are shown in three separate plots: note that results for some conditions are repeated. (a) Search asymmetry for dark shadows; (b) equivalent discrimination performance for dark and light shadows when shadow displacement is less for targets; (c) faster discrimination for dark shadows when shadow displacement is greater for target.

revealed a strong search asymmetry: search for a large target in smaller distractors was substantially faster than for a small target in larger distractors. It is likely that this factor is active for both dark-shadow and light-shadow stimuli, explaining why search is faster in both cases when the target shadow has greater displacement than the distractor shadow.

However, this does not explain the difference in search speed between dark-shadow and light-shadow stimuli in the case when the target shadows are more displaced (figure 13c). We believe this to be caused by a second, 3-D factor: the greater displacement of the dark-shadow target brings the target forward perceptually relative to the distractors. We reason that search may be slower for the light-shadow stimuli in this case because this second factor is not acting.

If this analysis is correct, one might wonder why search for dark-shadow targets with less displaced shadows should be so slow (figure 13a). Our hypothesis is that here rapid 3-D shadow processing is actually working against the search task, since it is driving the target further toward the background relative to the distractors. This logic assumes that object salience is inversely related to egocentric distance. This is not an implausible idea, but thus far there is, to our knowledge, no solid evidence for this proposal in the literature. Notwithstanding, it does seem that the visual system can in some cases make rapid discriminations based upon cast-shadow displacements suggesting differential object relief.[8]

9 Discussion

9.1 *Discriminating cast and attached shadows*

In our second experiment we tested the visual system's ability to rapidly discriminate cast and attached shadows. This appears to be the first attempt to address this issue (Mamassian et al 1998). Our main finding is that search is much faster for dark shadows than light shadows, indicating a highly sensitive visual mechanism for discriminating cast and attached shadows.

This result is broadly consistent with earlier findings that shadows can have a profound impact on scene perception (Yonas 1979; Berbaum et al 1984; Cunningham et al 1996; Kersten et al 1996, 1997; Tarr et al 1998; Castiello 2001; Madison et al 2001). However, Rensink and Cavanagh (1993, 1998) have suggested that shadows are discounted early in visual processing, making the visual system insensitive to shadow properties. This seems to be incompatible with our results: how could the visual system be insensitive to shadow properties when it can discriminate different types of shadows rapidly, in parallel across the visual field?

These findings may not be as incompatible as they seem. A crucial question is: in our task, is the neural representation underlying discrimination a representation of different types of shadow, as distinct parts or features of the scene, or is it a representation of the different 3-D objects that these shadows make apparent? If the latter, then it may still be that the visual system is insensitive to properties of the shadows that do not substantially affect the shape and location of objects in the scene.

In support of this hypothesis, we note a slight asymmetry in our results with dark shadows. Whereas the search slope for detection of the attached-shadow target in cast-shadow distractors was not significantly different from zero (parallel search), it was significantly positive for the reverse task: detection of the cast-shadow target in attached-shadow distractors. Now suppose it were the case that observers were not discriminating the shadows per se, but rather the 3-D objects these shadows make apparent. It would then be more appropriate to describe this asymmetry in terms of the convex shape (the 'ball') made apparent by the attached shadow and the flat shape (the 'disk') made apparent by the cast shadow. Whereas the convex ball 'pops out' from the flat disks, the reverse is not the case. This makes some sense if it is true that objects closer to the observer are more salient than objects further away (section 8). Thus, it may be that, while shadows are important determinants of scene perception, the representation underlying rapid discrimination is of objects or surfaces, and not of shadows per se. Properties of shadows that are inconsequential in determining object shape and location may essentially be ignored by the visual system, as Rensink and Cavanagh (1993, 1998) have suggested. This idea certainly bears further investigation.

Whatever the exact underlying representation, this experiment has made clear that the human visual system can discriminate different types of shadow very efficiently. But what are the exact image features that allow this discrimination? In this experiment we are fortunate that the target and distractor are very simple and very similar, and so there are not many candidates. The only difference between the two stimuli is which of the two contours of the crescent figure is blurred: the inner arc or the outer arc. The control condition of figure 5 shows decisively that this concept of inside/outside is key: it is the conjunction of blur and the partial closure of the contours that determines the 3-D percepts.

[8] Since the disks are all of equal retinal size, disks that are perceived closer to the observer should appear smaller if size constancy is operating. Whether this occurs and if so how it affects object salience is an interesting open question.

The importance of closure in perceptual organisation has been known for many years (Wertheimer 1923/1938). Treisman and colleagues suggested that contour closure was a basic feature that could drive pre-attentive discrimination (Treisman and Gormican 1988). While this idea was supported by later work (Kovács and Julesz 1993), more recent evidence suggests that the property of closure per se may not have a substantial effect on contour detection (Tversky and Geisler 2003). What seems more clear is that the property of closure or partial closure is critical for the rapid formation of representations of 2-D object shape (Elder and Zucker 1993, 1994). The present results can be considered to support the extension of this principle to three dimensions, at least in conjunction with the additional contour blur features required to elicit the perception of shadows.

9.2 *Detecting illumination inconsistencies*
The results of experiment 3 suggest that the human visual system can rapidly (in parallel) detect cast shadows that indicate illumination inconsistencies. This fits well with earlier work (Berbaum et al 1984; Ramachandran 1988) suggesting that the visual system resolves the concave/convex shading ambiguity by a computation that enforces global consistency under a single-light-source assumption. It is also consistent with later work with polyhedral stimuli which showed that object shading indicating inconsistent or unique illumination could produce pop-out (Enns and Rensink 1990). Again, it is apparent from these results that shadows can only be 'discounted' by the visual system after they have been used to detect such inconsistencies.

But what then are we to make of Cavanagh's demonstrations that the human visual system is insensitive to inconsistent shadows in visual art (Cavanagh 1999a)? One possibility is that our ability to detect these inconsistencies is a function of scene complexity. Illumination in natural scenes is extremely complex, owing to the presence of multiple light sources of various geometries and significant inter-reflections from both visible and non-visible surfaces in the scene (Fleming et al 2003). It may be that the visual system attributes apparent inconsistencies in the illumination of complex scenes to light sources and surfaces that are not apparent. Another possibility is that we fail to observe these inconsistencies simply because of the many other things in the visual scene competing for our attention. It is well established that observers may fail to notice glaring changes or inconsistencies in visual scenes even upon extended viewing (Rensink et al 1997), unless they are cued to attend to these parts of the scene.

In their investigations of illumination inconsistencies with shaded polyhedral objects, Enns and Rensink (1990) found a strong search asymmetry consistent with a human bias toward lighting from above: search slopes for detecting a side-lit target object in top-lit distractor objects were far lower than for detecting a top-lit target in side-lit distractors. In our results, we find no such asymmetry in search slope (search is parallel both for detection of a bottom-lit target in top-lit distractors and the reverse). However, we *do* find an asymmetry in the baseline reaction times (y intercepts): search is faster for detecting the bottom-lit target than the top-lit target. Both of these results suggest that scenes are more rapidly encoded when shadows are consistent with typical illumination directions.

9.3 *Depth from shadow blur*
In our fourth experiment, we found no convincing evidence that the degree of penumbral blur determines the perceived relief of an object casting a shadow. This seems to be counter to the earlier finding that modulation of penumbral blur can enhance the perception of motion in depth (Kersten et al 1996). It is possible that the difference is due simply to statistical sampling error: we note that the difference in search slope between dark and light shadows approaches statistical significance for one of the conditions (detection of targets with less penumbral blur, figure 11).

9.4 *Depth from shadow displacement*

The main finding of Kersten et al (1996) was that smooth lateral shadow motion could elicit a powerful perception of object motion in depth. Cunningham et al (1996) also found reductions in response time for detection of a foreground object when cast shadows were included in the rendering. However, we note that since no light-shadow control was tested in the latter experiments, the effects of overall target size (including the object and its shadow) might have confounded the results. Indeed, in our fifth experiment, we found evidence for a substantial search asymmetry determined by object size (for both dark and light shadows). Nevertheless, we did still find evidence for a dark-shadow-specific displacement cue to object relief, and this is substantially consistent with the results of both Kersten et al (1996) and Cunningham et al (1996).

Perhaps the most interesting part of these results is the finding that cast shadows can both speed and retard search, depending upon whether they serve to bring the object toward the viewer or away from the viewer. This further supports the notion that task performance here may be based not upon discrimination of shadows per se, but upon discrimination of object properties (in this case, object relief) that are inferred from these shadows. However, this remains a conjecture until more is known of the effects of apparent egocentric distance on object salience.

9.5 *Parallel versus serial search*

For brevity, we report detailed data only for target-present conditions. However, we also analysed the ratio of the search slope for target-absent trials over the search slope for target-present trials. A ratio of 2 is consistent with a serial search process that on average processes half of the search items in target-present conditions and all of the items in target-absent conditions. We focused on experiments in which a significant difference between dark-shadow and light-shadow conditions was found (ie where there was evidence for discrimination based upon shadow percepts). For these conditions we found a mean search slope ratio of 4.8 for dark shadows (range of 2.3 – 11) and a mean search slope ratio of 2.2 for light shadows (range of 1.6 – 2.8). In the context of a parallel/serial dichotomous view of visual search, these results are consistent with a more sequential search process for light shadows and a more parallel search process for dark shadows.

10 Conclusions

At the beginning of this paper we raised a number of open questions about shadows. What suffices for the perception of a shadow? How good are we at distinguishing different types of shadow? Are we sensitive to the global consistency of shadows in a scene? Which properties of shadows are most useful in making inferences about a scene? Are these inferences the result of high-level cognitive reasoning, or fast visual mechanisms operating in parallel across the visual field?

Using psychophysical methods, we have in this paper attempted to answer, or partially answer, all of these questions. We found that a very simple combination of darkness, partial closure, and blur is sufficient to generate compelling percepts of shadows and shape, and that different types of shadow (cast and attached) can be generated by simple manipulations of contour blur. We found that these different types of shadow can be discriminated very rapidly and in parallel across the visual field. Beyond detecting shadows and discriminating cast shadows from attached shadows, we found evidence for rapid detection of global illumination inconsistencies and for rapid computation of depth from shadow displacement.

While it is clear that the visual computation underlying rapid discrimination and search involves interpretation and inference from both cast and attached shadows, there is some evidence that discrimination may be acting upon representations of 3-D object shape and position elicited by perceived shadows, and not by shadows directly. In this sense shadows may largely be perceived by proxy.

References

Berbaum K, Bever T, Chung C S, 1983 "Light source position in the perception of object shape" *Perception* **12** 411 – 416

Berbaum K, Bever T, Chung C S, 1984 "Extending the perception of shape from known to unknown shading" *Perception* **13** 479 – 488

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

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

Cavanagh P, 1999a "Pictorial art and vision", in *MIT Encyclopedia of Cognitive Science* Eds R A Wilson, F C Keil (Cambridge, MA: MIT Press) pp 648 – 651

Cavanagh P, 1999b "Top – down processing in vision", in *MIT Encyclopedia of Cognitive Science* Eds R A Wilson, F C Keil (Cambridge, MA: MIT Press) pp 844 – 845

Cavanagh P, Arguin M, 1990 "Effect of surface medium on visual search for orientation and size features" *Journal of Experimental Psychology: Human Perception and Performance* **16** 479 – 491

Cavanagh P, Leclerc Y G, 1989 "Shape from shadows" *Journal of Experimental Psychology: Human Perception and Performance* **15** 3 – 27

Cunningham R, Beck J, Mingolla E, 1996 "Visual search for a foreground object in continuous naturalistic displays: The importance of shadows and occlusion" *Investigative Ophthalmology & Visual Science* **34** S299 (abstract)

Elder J H, Zucker S W, 1993 "The effect of contour closure on the rapid discrimination of two-dimensional shapes" *Vision Research* **33** 981 – 991

Elder J H, Zucker S W, 1994 "A measure of closure" *Vision Research* **34** 3361 – 3370

Enns J T, Rensink R A, 1990 "Influence of scene-based properties on visual search" *Science* **247** 721 – 723

Fleming R W, Dror R O, Adelson E H, 2003 "Real-world illumination and the perception of surface reflectance properties" *Journal of Vision* **3** 347 – 368 (http://www.journalofvision.org/3/5/3/)

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

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

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

Kleffner D, Ramachandran V, 1992 "On the perception of shape from shading" *Perception & Psychophysics* **52** 18 – 36

Kovács I, Julesz B, 1993 "A closed curve is much more than an incomplete one: Effect of closure in figure – ground discrimination" *Proceedings of the National Academy of Sciences of the USA* **90** 7495 – 7497

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

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

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

Rensink R A, Cavanagh P, 1993 "Processing of shadows at preattentive levels" *Investigative Ophthalmology & Visual Science* **34** 1288 (abstract)

Rensink R A, Cavanagh P, 1998 "The influence of shadows on visual search", in *Proceedings of the 69th Annual Meeting of the Eastern Psychological Association* (Boston, MA: EPA), p. 89

Rensink R A, O'Regan J K, Clark J J, 1997 "To see or not to see: the need for attention to perceive changes in scenes" *Psychological Science* **8** 368 – 373

Sun J, Perona P, 1996 "Preattentive perception of elementary three dimensional shapes" *Vision Research* **36** 2515 – 2529

Tarr M, Kersten D, Bülthoff H, 1998 "Why the visual recognition system might encode the effects of illumination" *Vision Research* **38** 2259 – 2275

Treisman A, Gelade G, 1980 "A feature-integration theory of attention" *Cognitive Psychology* **12** 97 – 136

Treisman A, Gormican S, 1988 "Feature analysis in early vision: Evidence from search asymmetries" *Psychological Review* **95** 15 – 48

Treisman A, Souther J, 1985 "Search asymmetry: A diagnostic for preattentive processing of separable features" *Journal of Experimental Psychology: General* **114** 285 – 310

Tversky T, Geisler W S, Perry J, 2003 "Contour grouping: is there something special about closed contours?" *Journal of Vision* **3**(9) 121a (http://www.journalofvision.org/3/9/121/)
Verghese P, Nakayama K, 1994 "Stimulus discriminability and visual search" *Vision Research* **34** 2453 – 2468
Watt R J, Morgan M J, 1983 "The recognition and representation of edge blur: evidence for spatial primitives in human vision" *Vision Research* **23** 1465 – 1477
Wertheimer M, 1923/1938 "Laws of organization in perceptual forms", in *A Sourcebook of Gestalt Psychology* Ed. W D Ellis (London: Routledge and Kegan Paul) pp 71 – 88 (Translated by W D Ellis from 1923 *Psychologische Forschung* **4** 301 – 350)
Yonas A, 1979 "Attached and cast shadows", in *Perception and Pictorial Representation* Eds C F Nodine, D F Fisher (New York: Praeger) pp 100 – 109

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

PERCEPTION

VOLUME 33 2004

www.perceptionweb.com