Perception, 2005, volume 34, pages 1301 – 1314

DOI:10.1068/p5418

Perceiving illumination inconsistencies in scenes

Yuri Ostrovsky§, Patrick Cavanagh#, Pawan Sinha§¶
§ Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology,
77 Massachusetts Avenue, Cambridge, MA 02139, USA; e-mail: psinha@mit.edu; # Department of
Psychology, William James Hall, 33 Kirkland Street, Harvard University, Cambridge, MA 02138, USA
Received 6 March 2002, in revised form 30 October 2004

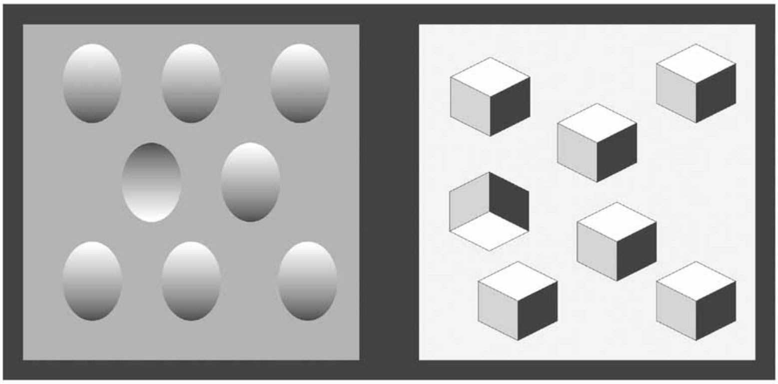
Abstract. The human visual system is adept at detecting and encoding statistical regularities in its spatiotemporal environment. Here, we report an unexpected failure of this ability in the context of perceiving inconsistencies in illumination distributions across a scene. Prior work with arrays of objects all having uniform reflectance has shown that one inconsistently illuminated target can 'pop out' among a field of consistently illuminated objects (eg Enns and Rensink, 1990 *Science* **247** 721 – 723; Sun and Perona, 1997 *Perception* **26** 519 – 529). In these studies, the luminance pattern of the odd target could be interpreted as arising from either an inconsistent illumination or inconsistent pigmentation of the target. Either cue might explain the rapid detection. In contrast, we find that once the geometrical regularity of the previous displays is removed, the visual system is remarkably insensitive to illumination inconsistencies, both in experimental stimuli and in altered images of real scenes. Whether the target is interpreted as oddly illuminated or oddly pigmented, it is very difficult to find if the only cue is deviation from the regularity of illumination or reflectance. Our results allow us to draw inferences about how the visual system encodes illumination distributions across scenes. Specifically, they suggest that the visual system does not verify the global consistency of locally derived estimates of illumination direction.

1 Introduction

When special-effects cinematographers create composite scenes (figure 1a), they go to great lengths to ensure that all objects are consistently illuminated (Fielding 1985; Brinkmann 1999). Differences in lighting directions across objects, they believe, would be immediately evident to the audience and reduce the realism of the scene. This intuition appears to be supported by formal experimental evidence. Visual-search studies have demonstrated that the visual system can efficiently and reliably spot an anomalously illuminated item in an array of identically lit three-dimensional (3-D) objects (Enns and Rensink 1990; Kleffner and Ramachandran 1992; see examples of their stimuli in figure 1b). Using accuracy measures in masked presentations of similar displays, Sun and Perona (1996a, 1996b, 1997) claim that fast, parallel processing of 3-D shape from shaded stimuli can occur in less than 80 ms. The cue for the target is ambiguous in these displays. It could be, as the authors claim, that the target appears to be of the same reflectance as the other objects, but is inconsistently illuminated. Alternatively, the target may appear to be illuminated by the same source as the other objects but is oddly pigmented. We will refer to both cues interchangeably without attempting to determine which is the critical factor, because our results show that, in fact, neither explanation is sufficient. The critical factor for the rapid processing of the oddly illuminated (or pigmented) target turns out to be the homogeneity of the 2-D and 3-D orientations of the distractors. When we allow the 3-D pose of the distractors (all identical cubes) to vary while maintaining consistent illumination, the oddly illuminated target becomes very difficult to detect. Odd illumination, on its own, in the absence of further image regularity, is not a feature that supports rapid search at all.

¶ Corresponding author

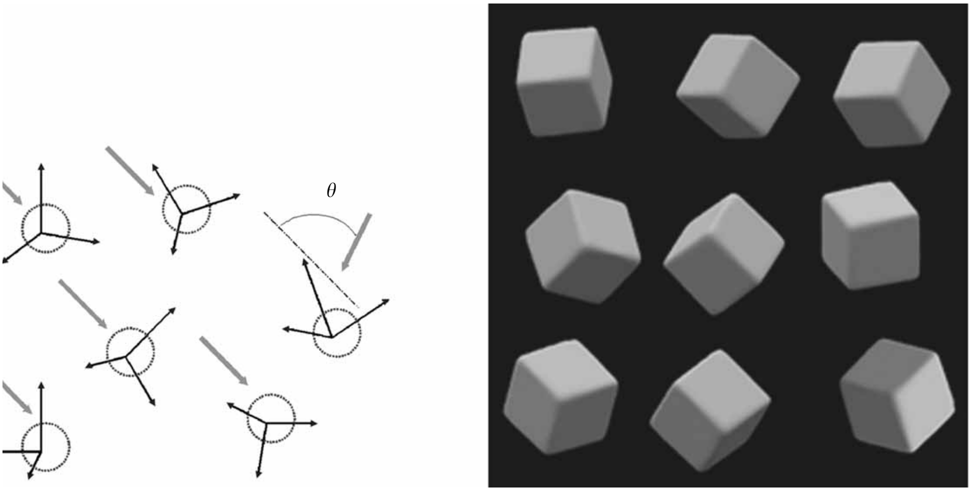
(a)



(b)

Figure 1. Scenes for which anomalies in illumination direction across objects are either expected to be or, in fact, are readily detectable. (a) A sample composite scene from the movie 'Jurassic Park II'. The dinosaurs and the humans are derived from separate scenes. Lighting directions for both are precisely equated. (b) Experimental displays which suggest that anomalies in lighting directions are perceptually very salient and 'pop out' pre-attentively either as an illumination inconsistency assuming uniform reflectances or a reflectance inconsistency assuming uniform illumination. This, and a number of other figures in this paper, can be seen in color on the *Perception* website at http://www.perceptionweb.com/misc/p5418/.

In the scenes used in the experiments mentioned above, an unavoidable property of natural scenes has not been explored: 3-D pose across objects (not to mention object identity) is not fixed, but is highly variable. In order to characterize human sensitivity to illumination or reflectance inconsistencies in more general settings, we have designed displays that differ from those used thus far in a key respect. As in many previous studies, our displays comprise several identical 3-D objects with all distractors illuminated from one direction and the target from a different direction. However, instead of assigning the same orientations in space to these objects, we randomize them (see figure 2a). This makes illumination direction the only reliable differentiator between targets and distractors. Or alternatively, if uniform illumination is assumed, it makes the target the only object differing from the others in reflectance. When distractors are identical in their 2-D appearance, pop-out may occur owing to simple pattern-matching



(a) (b)

Figure 2. Display design for our experiments. (a) Basic configuration of our displays. The objects are all identical, as in the previous studies, but their orientations in space have been randomized so that illumination direction is the only reliable differentiator between distractors and target (or alternatively, a reflectance difference assuming uniform illumination). (b) A sample display depicting one particular set of illumination directions (top and right illumination). See text for details.

strategies (which may in turn rely on familiarity with canonical representations of 3-D objects, such as cubes or spheres, to achieve high speed). Our manipulation reduces the effectiveness of pattern-matching strategies and forces subjects to rely solely on illumination direction or reflectance. In terms of either of these two variables, the display is homogeneous with only one odd item. Such conditions are sufficient for rapid search if either of those features can be assimilated independently into the parsing of a scene. Furthermore, by not constraining all the objects in the display to assume identical poses, our stimuli better represent real-world conditions.

Finally, to augment our visual-search experiments with cubes, we also asked observers to detect illumination inconsistencies in doctored real-world photos. Here again, the inconsistencies are hard to detect, supporting the finding that direction of illumination is not itself an explicit feature that can be rapidly identified.

2 Experiments

2.1 *Methods*

We chose the cube, a simple 3-D object with a history of use in this domain (Sun and Perona 1996a, 1996b), as the stimulus item. Figure 2b shows a sample display. Half of all our displays were fully consistent (all objects illuminated from the same direction), while in the other half, one cube (the target) was illuminated from an orthogonal direction relative to the distractors. A lighting direction was randomly chosen from eight different directions separated by 45° in the same plane in which the cubes were arranged: top (directly overhead), bottom (directly beneath), left, right, top-left, top-right, bottom-left, and bottom-right. For the inconsistent displays, the illumination direction of the target cube was chosen to be 90° clockwise or counterclockwise from the direction of illumination over the rest of the scene. As in the previous studies, subjects were asked to report whether the display contained an anomalously lit target, ie whether the scene had an illumination-direction inconsistency. We report the results of two separate experiments, one with fixed presentation times and the other self-timed.

The first experiment showed displays with set sizes of 4, 9, and 12 items for durations of 100 ms, 500 ms, or 1000 ms. (Each trial was pseudo-randomly assigned a set size and presentation time.) The second experiment (conducted immediately after the first experiment with the same set of subjects) consisted of displays identical to those of the first experiment, but with displays persisting on the screen until subjects made a response. The display was located approximately 50 – 70 cm from the subject, with stimulus items spanning approximately 2 deg, located at a maximum of about 10 deg from the center. (The results do not change significantly across a large range of viewing distances, suggesting that cube size is not a critical determinant of the results.) Subjects were free to look anywhere in the image during the task. Subjects indicated whether the display was consistent or inconsistent in illumination direction by pressing one of two keys on a computer keyboard. Seventeen naive subjects (aged 18 – 45 years) participated. They were recruited from the MIT subject pool and received compensation for their participation.

In order to replicate results from earlier studies and also to have a baseline condition, we first tested subjects on displays with all distractors oriented identically. Subjects achieved ceiling level performance (about 90% correct detection of anomalous displays) under these conditions, even with just a 120 ms display time. Performance was invariant to distractor numerosity within a wide range (4 – 12), indicating a parallel search (Treisman 1985).

For the next set of experiments, we control for the confounds introduced in earlier studies by distractor homogeneity by randomizing cube orientations. Note that the distractors are now heterogeneous in their orientation but they are all homogeneous with respect to direction of lighting (with the target being the only cube with a different direction of lighting). This control dramatically changed the results. Figures 3a, 3b, and 3c show results from three different fixed display times and the self-timed conditions. The data demonstrate a remarkable inability on the part of the subjects to detect illumination inconsistencies even with long viewing durations. In contrast to the ceiling-level performance with heterogenous distractors (homogeneous in their orientations, as well as lighting), maximal performance with inhomogeneous distractors averaged 65% even for small set sizes. (Chance level performance is 50%.) Performance decreased with distractor numerosity and increased with display time, indicating a slow serial search strategy. In order to explicitly test for the significance of the effect of set size on performance, we performed a 2-factor ANOVA. The resulting p-value is $< 10^{-4}$, indicating a very significant effect of the set-size factor. Display time also had a significant effect on performance ($p < 0.03$ for the timed conditions, $p < 0.01$ for the self-timed condition). This pattern of results is clearly quite different from those obtained in previous studies.

On the basis of these results, we infer that subjects were quite insensitive to the illumination inconsistencies embedded in the experimental displays. The discrepancy between results from earlier experiments and our studies suggest that the use of a homogenous field of distractors may have rendered the task of spotting the illumination anomalies unnaturally easy. By the same token, however, our study may be criticized for making the task unnaturally hard—the heterogeneity of orientations may have a detrimental effect on visual-search tasks in general. There are at least two responses to this concern. First, as mentioned in the introduction, heterogeneity of object poses in a real-world scene is the rule rather than the exception. By incorporating this characteristic of natural scenes, our displays are rendered more ecologically valid than those with entirely homogenous arrays of objects. The results are, therefore, more likely to be reflective of our perceptual abilities in the real world.

Second, it is not the case that visual search in general is compromised to this extent by the heterogeneity of items in the distractor set. Recall that our distractors are

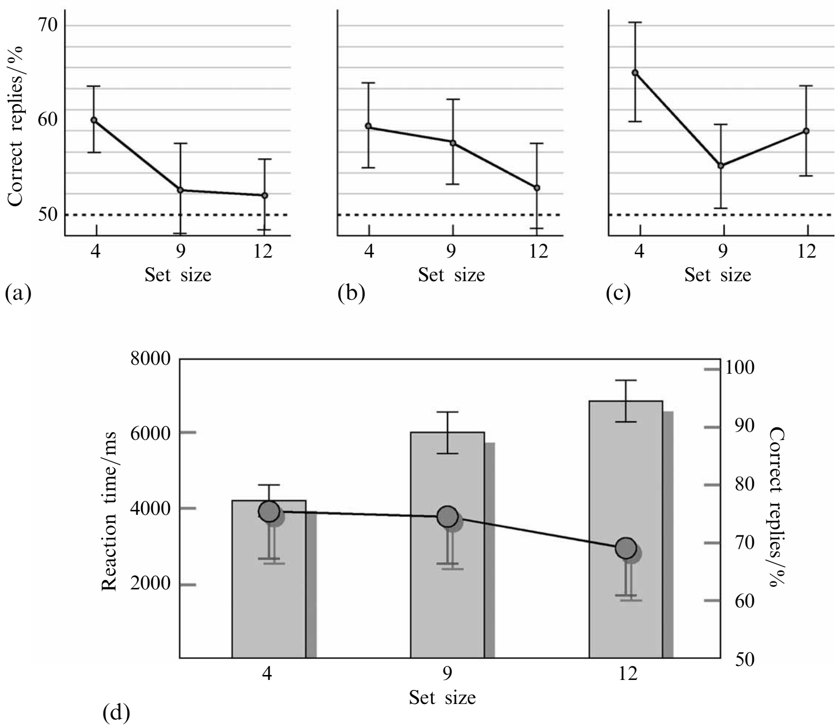


Figure 3. Experimental results. Each graph shows performance as a function of set size, parameterized by presentation time. (a) – (c) Performance on timed conditions: (a) 100 ms, (b) 500 ms, and (c) 1000 ms. Performance decreases with set size and increases with display time, as expected for slow, serial search tasks. (d) Reaction-time data (filled bars, left-hand vertical axis) and accuracy (filled circles, right-hand vertical axis) for the self-timed condition. Reaction time increases with set size, whereas performance remains the same or even decreases with set size, despite the increase in reaction time. Note that, although performance is better than on the faster, timed trials of (a) – (c), in absolute terms, it is still quite poor, illustrating that the task of detecting illumination or reflectance anomalies may be fundamentally difficult for our perceptual system. [Vertical bars in (a) – (d) show ±1.0 standard error, obtained after normalizing each subject's data to the same grand mean within each panel. This factors out the large individual variation in mean performance that is irrelevant to the discussion here.]

homogeneous in direction of lighting (or alternatively, in reflectance); the heterogeneity is on the task-irrelevant dimension of cube orientation. Do other attributes survive distractor heterogeneity on irrelevant dimensions similar to that used here? To address this question, we conducted an experiment wherein the relevant dimension was cube shape and the irrelevant dimension, as before, was cube orientation (figure 4). In the visual-search displays used in this experiment, the target differed from the distractors in the perceived rectilinearity of its internal junctions. In particular, the corner of the distorted cube facing the viewer was offset from its usual position, distorting the squareness of the faces. (The distortion was done in such a way that the outer boundary and the inner junctions were, independently, consistent with normal cubes). We found that, while distractor heterogeneity did slow down search, the detriment in performance was much less than that which we found for detecting inconsistencies in direction of lighting. Specifically, with a set size of 9 items, performance with presentation times of 200 ms, 500 ms, and 1000 ms averaged 68%, 71%, and 92%, respectively. In contrast to this, performance levels obtained while detecting lighting inconsistencies with similar set sizes and presentation times were just marginally above chance (shown in figures 3a, 3b, and 3c: 53%, 59%, and 56% for 9 items at 100 ms, 500 ms, and 1000 ms).

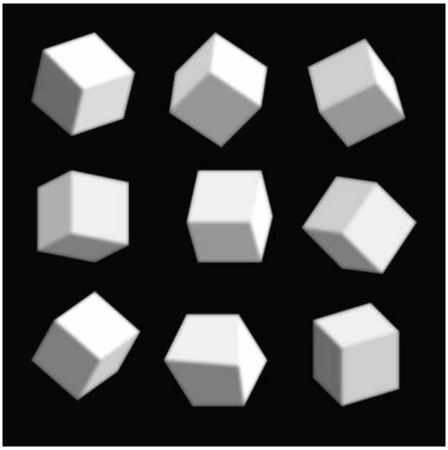


Figure 4. A sample display containing a cube whose shape is distorted. Such displays are much easier to search for inconsistencies than displays such as those in figure 2.

Since prior work had shown an asymmetry in performance in top-lit versus bottom-lit conditions, we also analyzed data as a function of primary lighting angle (figure 5) to be sure that the bottom-lit condition was not masking a pop-out effect in the top-lit conditions. For long presentation times, there was an accuracy advantage in top-lit conditions (distractors lit from directly overhead and 45° left of overhead), consistent with prior work, but performance was still poor and inconsistent with a pop-out interpretation. We show only the results for a presentation time of 1000 ms, as performance was at floor for shorter presentation times.

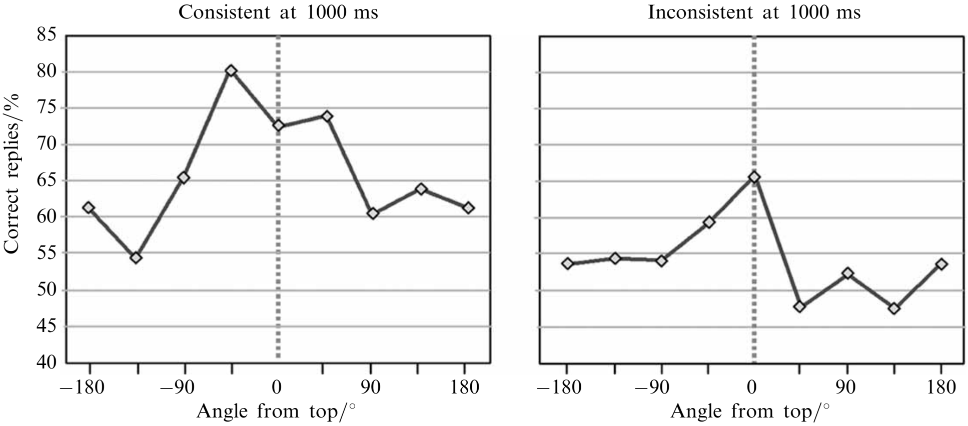


Figure 5. Accuracy as a function of primary lighting angle. Performance is poor at all lighting directions, but there is an advantage for top-lit versus bottom-lit distractors for long presentation times. No such advantage is observed for the shorter presentation times used in our experiments. (0° = directly above.)

As a next step, an important question that needs to be addressed is whether these results are specific to the experimental stimuli we used. There are two ways in which we have addressed this issue. The first approach involved using objects other than cubes to construct our search displays. These other stimuli (see figure 6 below) did not alter the pattern of results. Moreover, a separate group of subjects was asked to merely report the direction of lighting (choosing one of the eight direction possibilities mentioned above for cubes and the amoeboid shapes of figure 6, left) in an untimed, off-line task. Performance was nearly perfect, indicating that subjects did not have trouble comprehending the object geometry nor the lighting conditions. However, the poor performance obtained with multi-element displays suggests that subjects could not use this knowledge in a rapid, parallel manner.

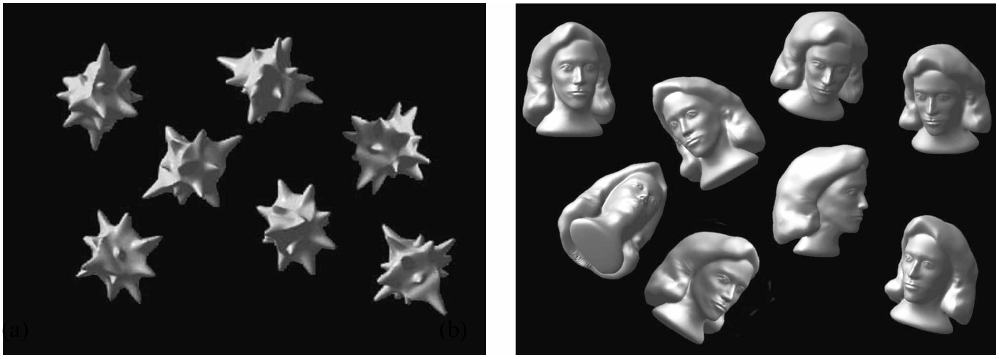


Figure 6. Examples of other stimuli used in experiments similar to experiment 1.

The second, and perhaps more informative, approach employed real-world scenes as the stimuli. To this end, we digitally modified images of real scenes to introduce illumination inconsistencies in them. The inconsistency between illumination directions within a scene averaged 90°. Some examples of the resulting images are shown in figure 7. (Unmodified versions are shown in figure 8.) A cursory examination of these scenes suggests that their illumination inconsistencies are not immediately evident—consistent with the results we obtained with the experimental stimuli. To verify these informal observations, we designed an experiment wherein subjects were shown image pairs consisting of one modified and one unmodified real scene in random order. Subjects had to indicate which scene in each pair had illumination inconsistencies.

The display was located approximately 50 – 70 cm from the subject, with images spanning 15 deg, on average. The experiment was run in three blocks, differing in the image presentation times (each image was shown for 1000, 2000, and 5000 ms in blocks 1, 2, and 3, respectively). A block comprised twenty-three image pairs. Each pair contained one image with inconsistent illumination and one unaltered image. The temporal ordering of the images in a pair was random. An image pair was presented as follows: (1) The first image in the pair was shown for the allotted time (depending on the block). (2) A gray screen was shown for 200 ms. (3) The second image was presented for the same amount of time as image 1. (4) A gray screen was shown until the subject indicated which of the two images was inconsistent by pressing one of two keys on the keyboard.

Figure 9 shows subjects' performance as a function of presentation time. Just as with our previous experimental displays, subjects performed poorly even with extended presentation times. Notwithstanding the explicit instructions to look for inconsistencies of illumination direction, subjects were not significantly above chance at presentation times of 1 s. Their performance improved to 70% when presentation time was increased to 5 s. These results indicate that the inconsistencies of illumination direction do not 'pop out', but require a relatively slow scan of the scene. In fact, it is conceivable that illumination inconsistencies in real scenes may be even less evident than is suggested by our results. In our stimuli, though we were careful to avoid them as best as we could, there may have been some local image artifacts (such as edges and chromatic differences) arising out of the image-doctoring operation. These artifacts may allow subjects to distinguish between modified and unmodified images. Furthermore, subjects were explicitly told before the start of the experiment to look for illumination inconsistencies. Unprimed subjects can be expected to be less sensitive to the inconsistencies in scenes. Indeed, in preliminary tests with subjects who were asked to pick out 'doctored' from 'undoctored' images without explicitly being asked to look for illumination inconsistencies, we found performance to be at chance even at the longest (5 s) presentation times.

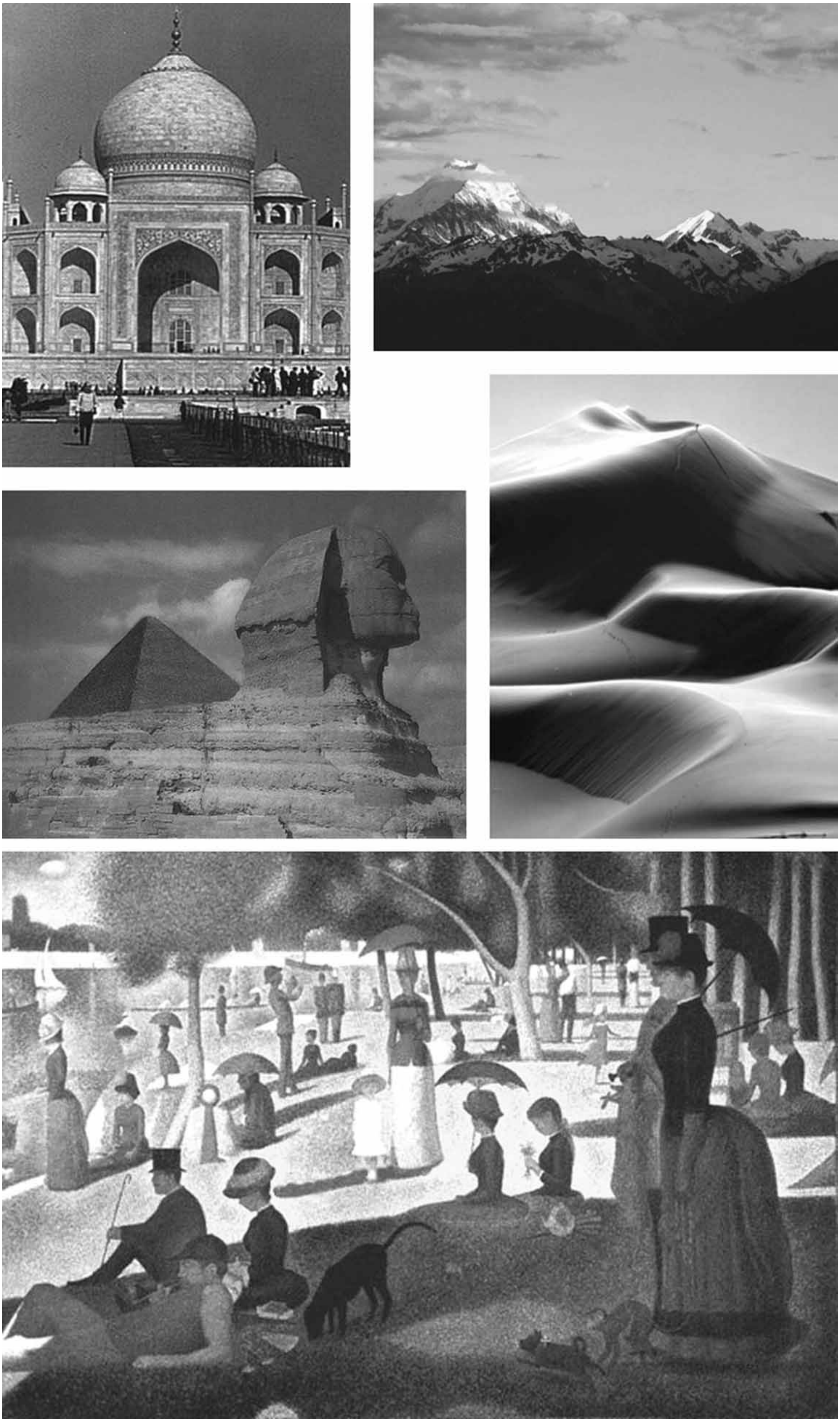


Figure 7. A few examples of scenes with digitally introduced illumination anomalies. Just as with the experimental displays shown in figure 2, the inconsistencies in these scenes are not perceptually salient.

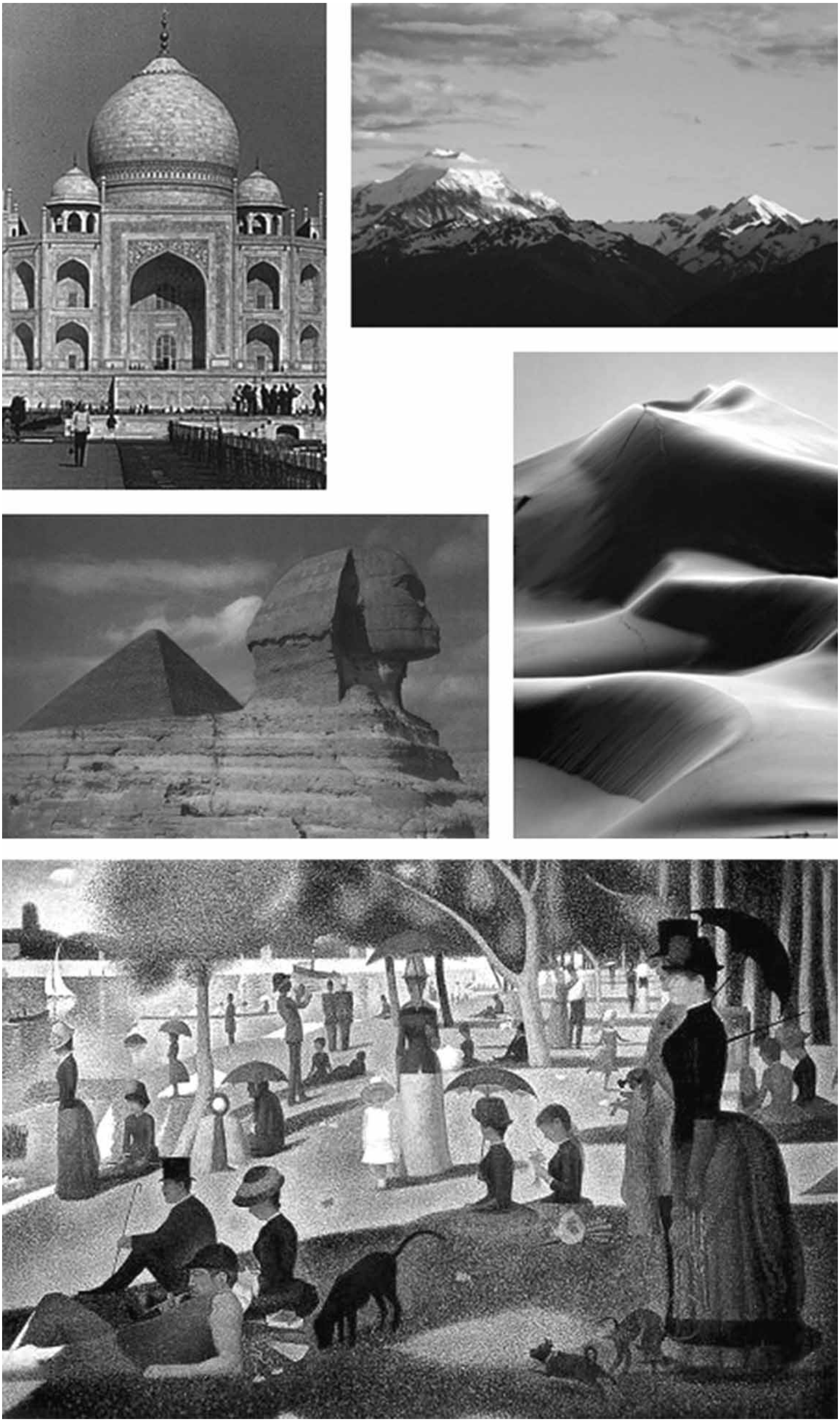


Figure 8. The same scenes as those shown in figure 7 in their original form (without illumination inconsistencies).

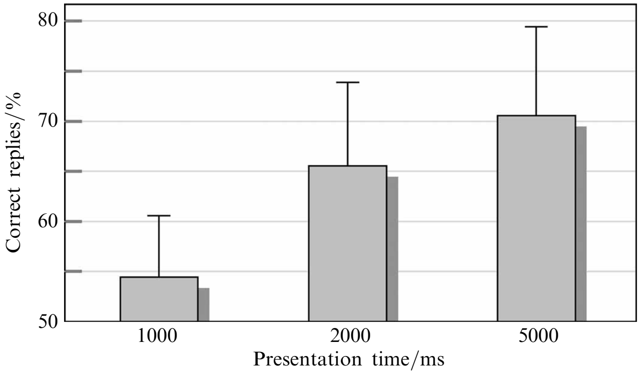


Figure 9. Summary of our results with real images. Despite being primed specifically to look for illumination inconsistencies in the images, subjects' performance was quite poor even with long presentation times. Chance level performance is 50%.

3 Discussion

Our results suggest that humans are quite insensitive to inconsistencies of illumination direction in the experimental displays we used and, more generally, in many real-world scenes as well. Artists have often exploited this insensitivity by choosing to depict illumination patterns in their paintings based on compositional aesthetics and social norms rather than constraining the patterns by physical laws (Gombrich 1995; Cavanagh 1999).

Besides the difficulty in detecting illumination inconsistencies, it can be argued that other factors might contribute to subjects' poor performance in the rapid-presentation condition (figure 3a). For instance, rapid presentation might compromise basic early-vision processes, including subjects' ability to perceive targets and distractors as 3-D shapes, or recover their reflectance distributions. However, Sun and Perona (1996b) have claimed that 3-D shape from shading is processed for multiple items within less than 80 ms. Mamassian et al (2003) briefly presented ambiguous shaded surfaces and recorded perceptual responses and ERPs. They found that the P1 response (a component peaking at about 100 ms in the ERP) showed a correlation to perceived 3-D shape. Based on this result, they claim that the shape of the shaded pattern must be disambiguated within the first 100 ms of stimulus presentation. So these results suggest that 3-D shape from shading should have been available even at our briefest presentation. Even if performance in our rapid presentation conditions was compromised to some extent, the critical outcome in our results was not related to the short duration. Performance remained poor (figures 3b and 3c) with longer presentation times and even with unlimited (self-timed) viewing. The problem at these durations cannot be impairments in the basic early vision processes involved in 3-D shape or reflectance perception; it must lie in the difficulty of detecting an illumination inconsistency when present in a heterogeneous scene. Further support for this argument derives from the results with manipulated natural scenes (figures 7, 8, and 9). After 1 to 5 s of viewing these scenes, shape and reflectance attributes are undoubtedly as fully processed as they ever will be, and yet the illumination anomalies are still difficult to find.

If the source of the difficulty is not the brief presentation, perhaps it is the heterogeneity of the distractors. But the variation is along a dimension that is irrelevant for the task and, in many displays, irrelevant variations have little or no effect on visual search. Imagine, for example, that the target in our displays was a red cube, whereas the distractors were the neutral gray of our current arrays. The red cube would certainly be detected rapidly and independently of the number of distractors, despite the irrelevant variation in cube orientation. If illumination direction is an explicit feature of each object, like color, then one oddly illuminated object in a field of uniformly illuminated objects should support rapid detection, even in the face of the irrelevant variations of object pose. Unlike color, however, illumination direction cannot be extracted independently of shape. Shape and illumination are co-dependent and the variation in shape might therefore interfere with the determination of illumination, either by slowing it down, or adding noise. To address this point, we ran a control where the target was a deviant shape (non-rectilinear) and the distractors varied in orientation as before. Here, with the target an odd shape, the search performance survived the distractor heterogeneity much better than it did when the target was the odd illumination, even though computing shape is equally important in both cases.

If the problem is not the computation of multiple 3-D shapes, perhaps it is the computation of illumination direction for each shape. As one reviewer pointed out, rapid computation of common illumination direction may only be possible for the multiple, identical distractors such as those used in previous studies (eg Enns and Rensink 1990; Sun and Perona 1996a, 1996b, 1997). Once there is heterogeneity of shape or 3-D orientation, no computation of common illumination may be feasible, even over long durations. That is, even if an illumination direction can be computed for each object, these multiple values might not support any accumulation into a group direction. We agree that this may be exactly what we have discovered: multiple identical items may be the one special case where computation of global illumination may be feasible. More natural scenes, even though illuminated with a single source, may not support the extraction of global illumination direction. In this case, the computation of scene illumination is an oddity itself, only available for the simplest of arrangements and not a general property available for objects in ordinary scenes. The rapid detection of the oddly illuminated target in previous studies may well have been based on other properties such as odd pigmentation or rapid analysis of canonical views, or properties yet to be identified. If the property is global illumination direction, however, the points raised here indicate that this property is available rapidly only in displays of identical items with unchanging 3-D orientation.

A recent study of visual search for an odd cast shadow (Rensink and Cavanagh 2004) adds to the evidence that illumination direction is not the key to rapid search, even with homogeneous distractors. In this study, all the search items were identical vertical posts with cast shadows falling behind and to the right. The shadow of the target fell at a sharper angle to the right than the distractors' shadows. In the baseline condition, the dark regions behind the posts appeared to be shadows. In the comparison conditions, several manipulations were used to change the interpretation of the shadow region into a non-shadow region which then appeared to be attached to the post to form a bent, two-part object. For example, the regions were colored white instead of black, or the entire array was inverted, or a contour was added around the shadow borders. None of these manipulations changed the shape of the post and its attached region; they only changed whether the attached region appeared to be a shadow or not. In all cases, search was at least twice as fast when the region was not taken to be a shadow. This result implies that, if illumination direction is the property driving the detection of the target in these displays, it is, even in this special case with homogeneous distractors, a rather inefficient cue. Rensink and Cavanagh (2004) did not claim that the shadows or direction of illumination were slow to compute. On the contrary, they claimed that they were analyzed rapidly in order to extract object contours and depth relation, and then discarded. If the same holds for the cube displays we have used here, then the problem is that illumination direction is only available during single-object processing, and discarded once the shape of the object is determined, rendering it inaccessible to subsequent visual search of the global scene. The goal of the analysis, after all, is to determine the objects' properties and discount the illuminant.

What might account for the insensitivity of the visual system to inconsistencies of illumination direction in more natural scenes? It is unlikely to be due simply to the visual system ignoring shading and shadows altogether. Studies (eg Hietanen et al 1992; Johnston et al 1992; Braje et al 1998, 2000; Tarr et al 1998) have shown that shading patterns are indeed encoded by the brain. In fact, even very young infants have been shown to be sensitive to information provided by shadows in pictures (Yonas et al 1978, 1979; Cameron and Gallup 1988). Also, previous studies have shown that the visual system can determine illumination direction for local image regions (Hagen 1976; Pentland 1982; Todd and Mingolla 1983). The contribution of our experiments lies in allowing us to address the issue of how these local estimates are combined across a scene. The results suggest that the visual system does not attempt to verify the global consistency of the local estimates. A corollary of this finding is that it is unlikely that the visual system encodes global illumination distributions (Langer and Zucker 1997).

There are two potential ecological roots of this 'deficiency'. First, there appears to be little adaptive advantage to be gained from having the ability to perform the verification of global illumination consistency. Local analysis typically suffices for key tasks like shape recovery (Erens et al 1993; Weinshall 1994). Second, in a single-source world with limited specularities and mutual reflections between objects, a fast, local analysis can suffice for recovering the needed information on illumination direction from a scene. According to this idea, our indifference to verifying global consistency of light directions may derive from the fact that our evolutionary history took place in an environment where a unitary light source (the sun) automatically enforced global consistency of local illumination patterns. Note that this idea precludes the hypothesis that our insensitivity to illumination inconsistencies is due to our willingness to tolerate multiple light sources. While such a hypothesis would account for the data, arguments against it include the accumulated body of work that indicates the bias of the visual system towards assuming a single light source (Ramachandran 1988; Kleffner and Ramachandran 1992; De Haan et al 1995); the lack of sensitivity observed even when the source is likely to be unitary (large-scale sunlit scenes in our stimulus set); and instances for which even the assumption of multiple light sources does not provide an adequate explanation (for instance, the inconsistency between the directions of shading gradient and shadow of the woman's skirt in Seurat's painting in figure 6, bottom panel).

How can we reconcile these experimental results with our subjective experience of noticing illumination anomalies in some old movies or poorly doctored (composited) images? Three factors may increase the perceptual salience of the anomalies in these cases. First, inconsistencies in other aspects of illumination besides just the direction (for instance, intensity and spectral content) may make them more easily detectable. Second, the existence of compositing-related artifacts, such as luminance, texture, or color edges, may signal the presence of inconsistencies. Third, familiarity with a scene may reduce the task of spotting illumination inconsistencies to one of novelty detection in images. The familiarity-based hypothesis would predict that inconsistencies in highly familiar scenes or objects would be perceptually salient. While we await a thorough test of this hypothesis, preliminary evidence does lend support to it. We experimented with images of human faces—a highly familiar shape for observers. Consistent with the hypothesis, anomalies in illumination of the kind shown in figure 10 are readily perceived (mean reaction time for distinguishing between anomalous and non-anomalous facial illumination was 300 ms). This result does not, however, distinguish between two possibilities. It may be the case that all lighting anomalies are perceptually salient so long as they are within one object, irrespective of whether the object is familiar or not. Alternatively, in order for the anomalies to be perceptually salient, the scene (showing a single object or multiple objects) may need to be familiar. Our preliminary results support the latter possibility. We measured reaction times with vertically inverted versions

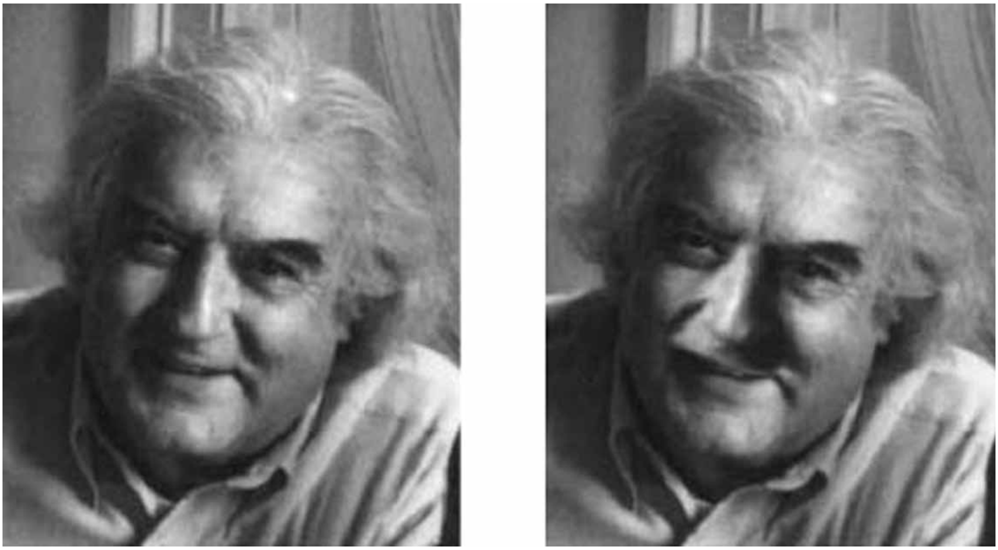


Figure 10. Illumination inconsistencies in faces. The inconsistent shadow of the nose in the image on the right is perceptually very evident. However, when these images are turned upside down, thus reducing the recognizability of the pattern, the perceptual salience of the inconsistency is reduced.

of our face stimuli. Given the lower level of familiarity observers have with inverted faces, the second possibility, but not the first, would predict that detection of illumination inconsistencies in inverted face images would be more difficult relative to upright faces. This is indeed what we find. Reaction times with inverted faces are nearly twice as long as with upright faces. The reader may verify the reduction in anomaly salience by turning figure 10 upside down.

In summary, our results show that observers are often remarkably insensitive to inconsistencies of illumination direction in experimental and natural scenes. This leads us to conclude that the visual system does not attempt to verify global consistency of estimates of the local illumination direction. These results bring up additional interesting questions, such as the roles of motion, albedo changes, and object relatability in facilitating the detection of illumination inconsistencies. Further, it is possible that total immersion in a full 3-D environment may facilitate illumination-direction awareness beyond our abilities with 2-D projections on paper or computer. Most of these issues can be investigated with the experimental paradigm we have presented in this paper.

Acknowledgments. We would like to thank Daniel Kersten, Tomaso Poggio, Ted Adelson, and Heinrich Bülthoff for useful discussion regarding this work. The research reported in this paper was supported in part by funds from the Alfred P Sloan Fellowship in Neuroscience, and the John Merck Scholars Award to PS, NIH EY09258 to PC, and an NIH Graduate Training Grant to the MIT Department of Brain and Cognitive Sciences.

References

Aks D J, Enns J T, 1992 "Visual search for direction of shading is influenced by apparent depth" *Perception & Psychophysics* **52** 63 – 74

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

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

Braun J, 1993 "Shape-from-shading is independent of visual attention and may be a 'texton' " *Spatial Vision* **7** 311 – 322

Brinkmann R, 1999 *The Art and Science of Digital Compositing* (San Mateo, CA: Morgan Kaufmann)

Cameron P A, Gallup G G, 1988 "Shadow recognition in human infants" *Infant Behavior & Development* **11** 465 – 471

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

De Haan E, Erens R G F, Noest A J, 1995 "Shape from shaded random surfaces" *Vision Research* **35** 2985 – 3001

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

Erens R G F, Kappers A M L, Koenderink J J, 1993 "Perception of local shape from shading" *Perception & Psychophysics* **54** 145 – 157

Fielding R, 1985 *The Technique of Special Effects Cinematography* (London: Focal Press)

Gombrich E H, 1995 *Shadows: The Depiction of Cast Shadows in Western Art* (London: National Gallery Publications)

Hagen M A, 1976 "The development of sensitivity to cast and attached shadows in pictures as information for the direction of the source of illumination" *Perception & Psychophysics* **20** 25 – 28

Hietanen J K, Perrett D I, Oram M W, Benson P J, Dittrich W H, 1992 "The effects of lighting conditions on responses of cells selective for face views in the macaque temporal cortex" *Experimental Brain Research* **89** 157 – 171

Johnston A, Hill H, Carman N, 1992 "Recognising faces: effects of lighting direction, inversion, and brightness reversal" *Perception* **21** 365 – 375

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

Langer M S, Zucker S W, 1997 "Casting light on illumination: A computational model and dimensional analysis of sources" *Computer Vision and Image Understanding* **65** 322 – 335

Mamassian P, Jentzsch I, Bacon B A, Schweinberger S R, 2003 "Neural correlates of shape from shading" *NeuroReport* **14** 971 – 975

Pentland A P, 1982 "Finding the illuminant direction" *Journal of the Optical Society of America* **72** 448 – 455

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

Rensink R A, Cavanagh P, 2004 "The influence of cast shadows on visual search" *Perception* **33** 1339 – 1358

Sun J, Perona P, 1996a "Early computation of shape and reflectance in the visual system" *Nature* **379** 165 – 168

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

Sun J, Perona P, 1997 "Shading and stereo in early perception of shape and reflectance" *Perception* **26** 519 – 529

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

Todd J T, Mingolla E, 1983 "Perception of surface curvature and direction of illumination from patterns of shading" *Journal of Experimental Psychology: Human Perception and Performance* **9** 583 – 595

Triesman A, 1985 "Preattentive processing in vision" *Computer Vision, Graphics and Image Processing* **31** 156 – 177

Weinshall D, 1994 "Local shape approximation from shading" *Journal of Mathematical Imaging and Vision* **4** 119 – 138

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

Yonas A, Kuskowski M, Sternfels S, 1979 "The role of frames of reference in the development of responsiveness to shading information" *Child Development* **50** 493 – 500

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

PERCEPTION

VOLUME 34 2005

www.perceptionweb.com