Journal of Experimental Psychology:
Human Perception and Performance
1989, Vol. 15, No. 1, 3-27

0096-1523/89/$00.75

Shape From Shadows

Patrick Cavanagh
Université de Montréal, Montréal, Québec, Canada

Yvan G. Leclerc
Artificial Intelligence Center SRI International,
Menlo Park, California

The colors, textures, and shapes of shadows are physically constrained in several ways in natural scenes. The visual system appears to ignore these constraints, however, and to accept many patterns as shadows even though they could not occur naturally. In the stimuli that we have studied, the only requirements for the perception of depth due to shadows were that shadow regions be darker than the surrounding, nonshadow regions and that there be consistent contrast polarity along the shadow border. Three-dimensional shape due to shadows was perceived when shadow areas were filled with colors or textures that could not occur in natural scenes, when shadow and nonshadow regions had textures that moved in different directions, or when they were presented on different depth planes. The results suggest that the interpretation of shadows begins with the identification of acceptable shadow borders by a cooperative process that requires consistent contrast polarity across a range of scales at each point along the border. Finally, we discuss how the identification of a shadow region can help the visual system to patch together areas that are separated by shadow boundaries, to identify directions of surface curvature, and to select a preferred three-dimensional interpretation while rejecting others.

How does the visual system identify and use shadow information in a scene? In general, objects are not illuminated uniformly from all directions, and the directed nature of the light produces both shading and shadow cues to the object shape. Shading, the variation of reflected flux with the angle between the incident light and the surface normal (Ikeuchi & Horn, 1981; Pentland, 1982; Woodham, 1981, 1984), can give information concerning surface orientation in areas receiving direct illumination. On the other hand, a shadow area is blocked from direct illumination (Beck, 1972; Berbaum, Tharp, & Mroczek, 1983; Gilchrist, Delman, & Jacobsen, 1983; da Vinci, see Richter, 1970; Shafer, 1985; Yonas, 1979), and it is specifically the shape of a shadow that carries three-dimensional (3-D) scene information. The shadow's shape, however, is determined by several factors simultaneously: the direction of the light source, the shape of the object casting the shadow, and the surface relief on which it falls, as well as the relative positions of the light source, object, and receiving surface. In theory, it is possible that the visual system could use the shape of a shadow to recover one or more of these factors if it had sufficient information concerning the remainder, but in practice, it seldom does. It is often the case that the shape of the object casting the shadow is unknown or that the relief of the surface on which it falls is indeterminate. In these cases the visual system cannot solve the shadow correspondence problem. Because it cannot link each shadow area to the object that produces it, the visual system must have a simpler set of rules and criteria for identifying a shadow based solely on the properties of the shadow region and its surround. In this article, we investigate the low-level criteria that govern the recovery of 3-D shape from shadows.

The ability to identify shadows is critical to the correct interpretation of the scene. In order to identify object surfaces, the visual system must locate the borders between the various materials that make up the object and those that distinguish it from the surrounding objects. The candidate borders include both borders arising from actual material changes and those due to shadows. If shadow regions are not correctly identified, their borders will be taken to indicate material changes dividing a continuous surface into two regions. If a scene has too little or only just enough information to specify its three-dimensional organization, a single erroneously labeled border can force a complete reorganization of the interpreted surfaces. Shadows, therefore, have very high nuisance value, and the reliable detection of shadow areas is a fundamental problem facing the visual system.

Although shadow borders are a nuisance for object segregation processes, they may be useful for recovering the relief of the surface on which the shadow falls. In Figure 1a, for example, the trees casting the shadows are not all visible in the photograph; thus it is evident that the visual system does not actually solve the shadow correspondence problem in this case but instead engages in some informed guessing. First, the dark, horizontal stripes in this scene are perceived as shadows as opposed to strips of dirty snow, and the bends in the shadows are then interpreted as bumps on the snow's surface.

Shadows may also give information about the objects casting them, and in some instances, the object's shape may be provided by the shadow information alone (Figure 1b). Three types of contours are involved in these instances of self-

Part of this research was presented to the annual meeting of the Association for Research in Vision and Ophthalmology in Sarasota, Florida, May 1985.

This research was supported by Grants A8606 to Patrick Cavanagh and A4470 to Steven Zucker from the National Sciences and Engineering Research Council of Canada and by the U.S. Defense Advanced Research Projects Agency under Contract MDA903-86-C-0084.

The authors would like to thank Luce Marinier and Joane Boucher for technical assistance and Stuart Anstis, Greg Dudek, and Steve Zucker for comments on the manuscript.

Correspondence concerning this article should be addressed to Patrick Cavanagh, Département de Psychologie, Université de Montréal, Montréal, Québec, H3C 3J7, Canada.

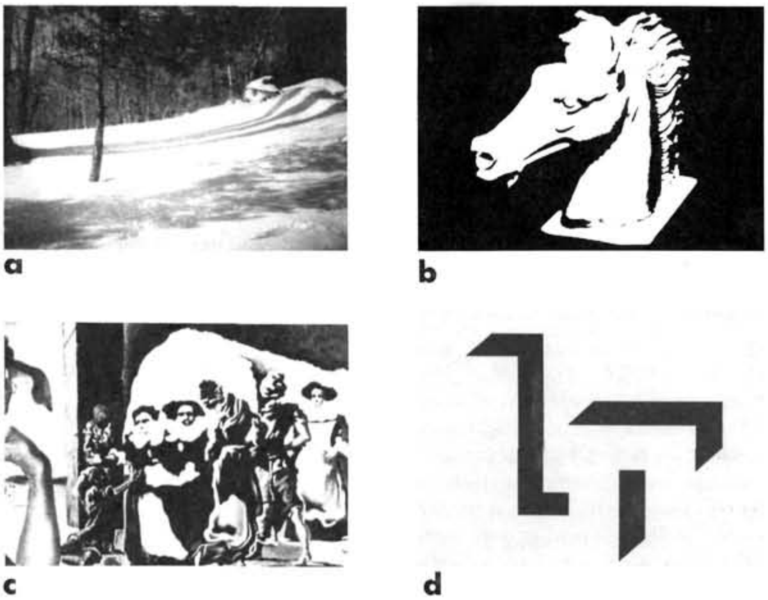


Figure 1. Panel a: The shadows of trees falling across snow. (The cast shadows provide information about the surface relief of the snow. Because the trees casting the shadows are not all visible in the photograph, the visual system cannot actually solve the shadow correspondence problem but must make informed guesses.) Panel b: A bust of Bucephalos, Alexander the Great's horse, having uniform reflectance everywhere on its surface. (The impression of three-dimensional shape is provided solely by the shadow information.) Panel c: A detail of Dali's "The Slave Market With Disappearing Bust of Voltaire." (Collection of Mr. and Mrs. A. Reynolds Morse. Reproduced courtesy of the Salvador Dali Museum, St. Petersburg, Florida.) Panel d: Familiarity with the shadowed object is not essential for depth from shadows. (Although the three-dimensional structure of this shadow object is not immediately evident for all observers, it is for many, and it can always be seen once it is pointed out [a long rod with a square cross-section bent at three locations to form a *u* with a bent righthand upright]. There may be a set of simple features that signal local structure in shadowed surfaces such as these.)

shadows: terminator contours, where the shadow is attached to the feature casting the shadow; cast contours, where the shadow of one part of the object falls on another part; and extremal or external contours, where the object's surface is normal to the observer's line of sight. In our experiments, we have not distinguished among these types of shadow contours, but at the end of the article, we discuss the possibility that the three types play different roles in retrieving 3-D shape.

The interpretation of shadow figures such as Figures 1a and 1b involves not only the low-level criteria that we will evaluate in our experiments but also high-level knowledge. Specifically, a major step in retrieving 3-D shape from shadows lies in deciding whether a particular area in an image is dark because it is a shadow or is dark simply because it has low reflectance (dark pigment). Because there is often insufficient information in the image to resolve this ambiguity, the visual system must be using knowledge about the objects potentially in the scene.

Pictures make the ambiguity between dark material and shadow even more evident because the dark areas in pictures like Figure 1b are just that: dark pigment placed on the page to produce the same pattern of light arriving at the eye that the real shadowed object would produce. The visual system prefers the 3-D interpretation of a familiar or simple figure (Figure 1d) rather than dark spotches on a flat surface, but nevertheless it registers, no doubt because of binocular disparity and texture cues, that these are spatial arrangements of pigment on paper, not real objects. Artists are naturally aware of the ambiguity between dark material and shadow; Dali, in particular, has exploited it in his painting "The Slave Market With Disappearing Bust of Voltaire" (Figure 1c), where two different 3-D interpretations are possible for the same image. In one, the dark areas are the dark material worn by the two women, and in the other, they are the shadows of the facial features of a bust.

We assume that the processes that derive shape from shadow pictures that we use in our experiments (presented on paper or cathode-ray tube [CRT] monitors) are the same as those that derive shape from real shadows. Some evidence suggests that the equivalence of shape-from-shading for real objects and their pictures may be established early in life or perhaps even be innate (Yonas, Cleaves, & Petersen, 1978).

Does the role of high-level knowledge evident in the bistable effect of Dali's painting imply that shadow-defined objects can be interpreted only if they are familiar? The shadowed images of familiar objects that we have examined—faces, cups, baby's booties, gloves, horses, and dogs—all produced compelling impressions of 3-D shape as long as the lighting direction captured important contours. Three-dimensional interpretations are not limited to familiar objects, however (Figure 1d). There may be a small set of local shadow contour shapes that the visual system "understands" so that unfamiliar surfaces may support 3-D interpretations as long as they make "sense" locally. We have not examined the possibility of locally interpretable shadow features in any detail. The shadow figures that we have used experimentally have all been those of familiar objects.

In order to determine the low-level aspects of a shadow region that make it acceptable as a shadow, we studied figures such as Figure 1b, where shadows are the only cues to the shape of the object. A shadow has two attributes of interest: (a) its shape, which in the images we have used is the only cue to both the object producing the shadow and the surface relief receiving it; and (b) its quality—the color, texture, brightness, motion, and binocular disparity of the shadow region. We examined these two aspects by presenting shadow figures defined solely by color, for example, or by color and luminance. In the first instance, we evaluated the contribution of shape descriptions conveyed by the various pathways of the visual system (e.g., color or motion) to the interpretation of shadows. In the second instance, we looked at the role of the physical constraints of natural scenes on the perception of shadows. For example, would the perception of a shadow figure be disturbed by inappropriate color relations between the shadow and nonshadow regions?

Pathways and Shape Descriptions

Zeki (1978, 1980), van Essen, Maunsell, and Bixby (1981; see also van Essen, 1985, and Maunsell & Newsome, 1987), and others have claimed that the information in the visual scene is broken down into several separate representations, each specialized for a different attribute such as color or motion. Each area, in addition to performing specialized analyses, is also capable of representing two-dimensional (2-D) shape: a 2-D map of regions differentiated by the attribute in question. A simplified schematic of the pathways through these specialized areas is shown in Figure 2.

Gregory (1977, 1979) has suggested that each attribute may provide a rough map of the visual stimulus and that these maps may be aligned to the luminance representation, which he considers the master map. However, earlier studies have shown that several perceptual abilities can be supported by the shape information in a single representation such as color or texture in the absence of any luminance "master map." In particular, information signaled by stimuli having explicit contours was effective no matter which visual pathway was used (Cavanagh, 1985, 1987). Simple, 2-D letter shapes could be easily identified. Three-dimensional objects defined by complete contours in line drawings involving occlusion and perspective were interpreted in the same fashion whether represented by luminance, color, or texture. There was no indication that luminance information had any privileged role to play in these images.

Is the shape information conveyed by the various pathways also sufficient for the perception of depth from shadows? Shadow interpretation must be based strictly on shape in a stimulus such as Figure 1b because there is no other information available, no variation of surface reflectance, no texture, no shading, and no visible contours within the shadow area. To examine the ability of shape carried by these attributes to support the perception of shadows, we presented shadow figures defined by a single attribute at a time—color, motion, texture, binocular disparity, or luminance.

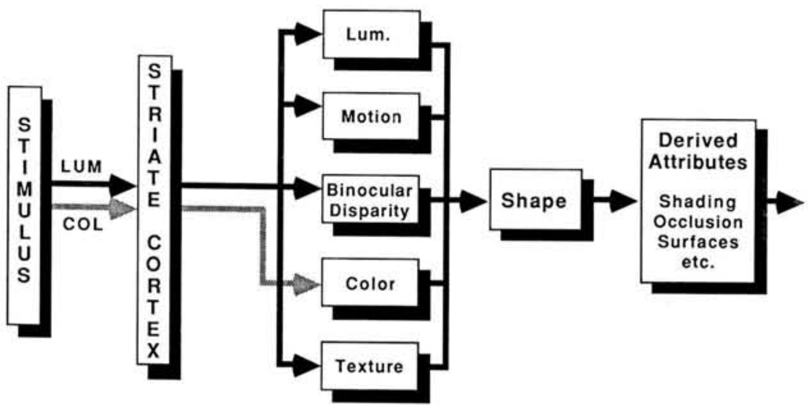


Figure 2. A schematic representation of perceptual pathways in the visual system. (Luminance and color pathways bring information from the retina to the striate cortex, where multifunction cells begin the analysis of orientation, motion, and binocular disparity. Following the striate cortex, information is routed to areas performing specialized analyses of various attributes: color [area V4, Zeki, 1978] and motion [area MT, van Essen, 1985], for example. Luminance, binocular disparity, and texture are other stimulus attributes that may receive independent perceptual analysis and that may also have separate areas of extrastriate cortex dedicated to their processing. Each of these specialized areas registers a two-dimensional representation of the attribute being analyzed, and all of these contribute to an overall representation of stimulus shape from which higher level attributes such as shading, occlusion, and surfaces can be derived.)

Constraints

Using the same figures, we also examined whether natural constraints play a role in the interpretation of shadows. Specifically, the physics of light restricts the changes of properties such as brightness and color that can occur across a shadow border. The most constrained shadow border involves a change only in illumination and not a change in material.

Shadows are generally classified as cast (an object's shadow falling on another surface) or attached (an object's shadow that falls on itself—a self-shadow), but this classification does not distinguish between the different types of borders bounding the shadow area. The contour of a cast shadow falls across and divides a continuous surface; however, a self-shadow contains both a cast contour, where the shadow falls on another part of the object's surface, and an attached contour, called a terminator, where the shadow is attached to the surface casting it. Even this distinction is not sufficient for our purposes because the strongest constraints apply specifically to shadow borders that fall across continuous surfaces. We label this type of border a *same-surface border.* A same-surface border may arise from a cast shadow, such as that on the lower surface of Figure 3a, or from the terminator contour of an attached shadow, such as the shadow edge around the middle of the sphere in Figure 3a. A joined-surface shadow border (Figure 3b) is a terminator contour that is attached to a sharp discontinuity in surface orientation. Because the discontinuity indicates a possible change in surface material, the strongest constraints cannot be applied here. An occluded-surface border (Figure 3b) occurs where the extremal contour of part of the object that is in shadow occludes the illuminated background surface.

To study the effect of constraints, we have concentrated on the same-surface border (Figure 3a). Because there is a single material surface being arbitrarily divided by a change of illumination, we do not expect changes in any property across the border except, of course, brightness. Specifically, there are five constraints.

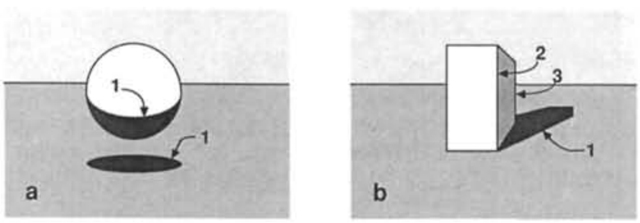


Figure 3. A same-surface border (1) may arise from a cast shadow such as that on the lower surface of Panel a or from the terminator contour of an attached shadow such as the shadow edge around the middle of the sphere in Panel a. (The terminator contour follows the points along the surface that are normal to the direction of the illuminant.) A joined-surface shadow border (2) in Panel b is a terminator contour that is attached to a sharp discontinuity in surface orientation. (Because the discontinuity indicates a possible change in surface material, the strongest constraints cannot be applied here. An occluded-surface border [3, Panel b] occurs where the external contour of a shadowed part of the object occludes the illuminated background surface.)

1. Brightness should decrease in the shadow region.
2. Only certain color changes can occur across the shadow border.
3. No change in either motion or depth is expected at the shadow border.
4. The nature and the contrast of the surface texture, if any, should be the same on both sides of the shadow border.
5. The shadow shape is constrained by the object casting the shadow, by the light source, and by the receiving surface.

Shadow Figures

How can we measure whether a particular area of a figure is perceived as a shadow or not? An observer's overt classification of an area as a shadow or nonshadow region may not be very reliable. In particular, our introspection about whether or not a particular region looks like a shadow may be only weakly related to whether our visual system has actually used the area as a shadow in recovering scene attributes. Gilchrist, Delman, and Jacobsen (1983) attempted to bypass this problem by measuring a scene attribute, surface brightness, that is directly affected by shadow interpretation. They introduced a small test patch at the potential shadow border so that it lay half inside and half outside the potential shadow. They then asked the observers to adjust the relative brightness of the two halves of the test patch until they appeared equally bright. The authors assumed that if the region was perceived as a shadow, the observer would correct for the reduced illuminant in the shadow region and show brightness constancy in the settings. We attempted a similar procedure in our figures but found very quickly that the presence of the test patch itself interacted with the interpretation of the shadow region. In fact, when the interpretation was somewhat ambiguous, it was the relative brightness of the two halves of the small test patch straddling the border that determined whether the region looked like a shadow or not.

We therefore chose a different procedure where a change in the interpretation of the shadow region produced a clear change in the perceived 3-D structure of the figure. We then asked observers to make judgments of high-level aspects of the figure, reporting when the 3-D structure of the image changed, while they adjusted selected parameters in the image.

The images in Figure 4 show three of the figures studied. For technical as well as theoretical reasons, the cues were limited to shadows rather than shading. We wanted to substitute attributes such as motion and texture for brightness, and because it is difficult to do this in a graded fashion, binary, or two-level, images were most appropriate. In addition, we needed a stimulus that changed significantly if the shadow regions were not interpreted as shadows. The faces and cup of Figure 4 are threshold images of real objects lighted by a single source, and they fill this requirement of two distinct organizations depending on the interpretation of the shadow areas.

Figure 4 shows positive and negative versions of a number of stimuli. Note that the shadow is correctly interpreted, and the 3-D structure of the figures is seen as long as the shadow regions are darker than the nonshadow regions. Most of the shadows in the first image (Figure 4a)—those cast by the nose,

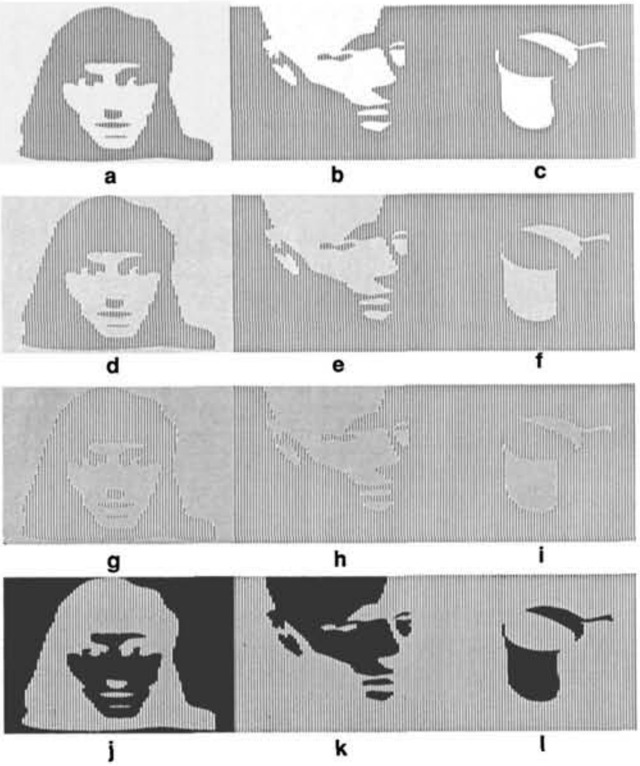


Figure 4. Binary images (i.e., two-valued, light and dark, in the examples here) of a woman's face (Panel a), a mannequin's face whose features are visible solely due to shadows (Panel b), and a coffee cup (Panel c). (Many of the contours of these figures are shadow contours, not object contours. All of the shadow contours are same-surface contours [of both cast and terminator varieties]. There are no sharp discontinuities producing joined-surface attached shadows and no extremal contours between shadowed object parts and lit backgrounds. Many of the object contours, both external contours and internal self-occlusions, are implicit, hidden in the shadow areas. The interpretation of these figures changes when an appropriate luminance difference is present between the shadowed and nonshadowed areas [top row] compared with when negative or no contrast is present [bottom and second from bottom row, respectively].)

eyebrows, and cheeks—are same-surface shadows falling on the skin of the face, which, except for the brightness change due to the shadow, should have uniform characteristics on both sides of the shadow borders. Note also that in this face, both shadow regions and regions of low reflectance—hair, eyebrows, and pupils—have the same brightness. Evidently our knowledge about faces allows us to segregate shadow and nonshadow regions in order to interpret these images. (Even though the dark areas contain both shadows and areas of low reflectance for Figure 4a, we will, for simplicity, refer to the dark areas of the images in Figures 4a, 4b, and 4c as shadow regions and the light areas as nonshadow regions.) The other two images we have studied have no areas of low reflectance; all dark areas are due to shadows. Our perception of the surface relief of the mannequin's face or the three-dimensional shape of the coffee cup is entirely dependent on the correct interpretation of the shadows. As previously mentioned, it may also be dependent on the familiarity of the stimuli (the structure of faces) or their simple local structure (a cup is a single concavity).

Although the three shadow figures that we used (Figure 4, top row) are generally seen as 3-D objects on first viewing, not all shadow figures are so easily interpreted. For example, the Mooney shadow faces, used to diagnose right parietal brain damage (Lezak, 1976; Mooney, 1957), are often difficult to perceive on first glance, and in Figure 1c here, many people do not see the bust of Voltaire until they are told what to look for. The difficulty and individual variability in the interpretation of shadow figures can be compared to that for random dot stereograms (Julesz, 1971), where some individuals initially take several minutes to perceive the 3-D shape. In our experiments, we did not address the question of the initial difficulty in getting to the shadow-based 3-D interpretation. Our goal was similar to that of the original stereogram experiments (Julesz, 1971) in that we determined the range of conditions for which a trained observer could extract the shadow-based 3-D interpretation from a known figure and the conditions for which this was impossible. In order to obtain reliable measurements with this procedure, it was essential that the observers could clearly discriminate the shadow-based from the alternate interpretation of the test figures and each condition began with a familiarization period to ensure that they could do so.

In our experiments the shadow figures were altered by adding color, texture, depth, or motion differences to the two areas of the image. Observers were then asked to adjust the figure contrast, the luminance contrast between the shadow and nonshadow regions. (This figure contrast is the dependent variable in our first five experiments.) In general, there were two quite distinct organizations of the figure visible at different figure contrasts: the "shadow figure," as seen in the top row of Figure 4, and an alternate organization seen when the shadow areas were not interpreted as shadows, as seen in the bottom row of Figure 4. Observers were asked to attend to high-level properties of the stimuli, such as facial expression, which were clearly affected by the low-level shadow cues, and to thoroughly familiarize themselves with how these changed as a function of the figure contrast.

Generally, when the shadows were not correctly interpreted, the regions themselves became segregated surfaces, giving the impression of 2-D cartoons (see Figure 4, bottom two rows or Figures 9a and 12a). For example, the light/dark borders of the face stimulus in Figure 4a, when incorrectly interpreted, appeared to mark a change in material, delineating hair from a mask-like face, or a button nose from the surrounding skin. When the shadows were correctly interpreted, however, the skin appeared continuous on both sides of the light/dark border, and the impression was that of a face with a definite expression.

The 3-D organization of the stimuli also changed noticeably when the shadows were correctly interpreted. The faces took on the structure of a face with protruding eyebrows, nose, and cheekbones. The cup appeared empty so that something could be placed inside without hitting any surface except the (invisible) bottom. When the shadow regions were not seen as shadows, the faces, besides lacking global organization, looked flatter. The cup appeared either to have a tilted surface inside it as if it were filled with sugar or to have lost its global organization altogether (Figures 4i and 4l) and looked like two unconnected pieces. Thus, observers adjusted the figure contrast to find the transition point between the two possible figure organizations. In general, the transition point could be located fairly easily, even in experimental conditions where it was quite difficult to see the figures.

How would the attribute used to present the figures—color, relative motion, stereo, or texture—affect the shadow interpretations? If 2-D shape defined by a particular attribute such as color or motion is capable of supporting 3-D shape from shadows, the shadow figure (the organization involving correctly interpreted shadows, top row, Figure 4) should still be recognizable with no luminance contrast between shadow and nonshadow regions, that is, at zero figure contrast. On the other hand, if the change of a particular attribute at the shadow border violates a shadow constraint, the shadow figure may not be recognizable no matter how much luminance contrast is also added.

Experiment 1: Color

It is fairly common to have shadows that differ somewhat in color from their surround. Colored shadows can occur if there are two differently colored light sources, such as yellow sunlight and bluish skylight, or, as in Figure 5, two differently colored suns. In this case, the green shadow on the right is produced where the object has blocked the light from the red source but not the green. The light that falls in this shadow (green) must always fall in its surround as well, producing, in this case, a yellow (red plus green) surround. There are therefore constraints on the saturation of the colors in the shadows and the surrounds: A saturated green shadow can never have a saturated red surround, for example. These were the colors we used to examine whether this natural constraint had any influence on the perception of shadows in chromatic images.

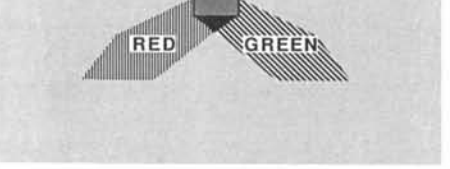


Figure 5. Colored shadows produced by two differently colored light sources. (Whatever color falls in a shadow must also fall in its surround. The shadow color is therefore more saturated than the surround color.)

We also tested images having the same color (red or green) in shadow and nonshadow regions to determine the figure contrast necessary for visibility in the absence of color differences at the shadow border.

Method

The images were presented on a computer-controlled monitor having 512 × 480 pixel resolution and a 30-Hz interlace raster. The Commission Internationale d'Éclairage (CIE) x and y coordinates for the three phosphors were determined by spectroradiometry to be 0.596, 0.346 for red; 0.293, 0.604 for green; and 0.149, 0.069 for blue. The screen size was 27 × 27 cm viewed from 1.93 m for an 8° visual angle display. Three stimuli were used: the woman's face (Figure 4a) subtending 8.0° × 8.0°; the mannequin's face (Figure 4b) subtending 6.0° × 7.5°; and the coffee cup (Figure 4c) subtending 5.0° × 7.0°. In all cases a fixation bull's-eye was placed near the center of the screen, within a uniform area (cheek or front portion of cup) to aid accommodation.

For the measurements with the chromatic figures, the light and dark areas of the original images were replaced with red and green. The nonshadow areas were filled with red and set to a luminance of 20 cd/m^2. The shadow areas were filled with green. The observer was asked to adjust the luminance of the green areas to the transition point where the 3-D shape of the face or cup just disappeared. In this and all the following experiments, the observer first adjusted the image over a sufficient range so that both organizations of the figure could be seen and became thoroughly familiar with both of them. Following this familiarization step, four settings were made: two starting from above the transition point and two starting from below, and the observer adjusted the figure contrast back and forth across the transition point until an acceptable setting was found. There did not appear to be much hysteresis in the location of the transition point, and the fact that the observer knew which figure was being presented in each condition may have contributed to this. The red and green areas were then exchanged: The nonshadow areas were filled with green, and the shadow areas were filled with red; red luminance was fixed at 20 cd/m^2. The observer again adjusted the luminance of the green areas to the transition point where the surface relief of the face or cup just disappeared. Four settings were made. The effective luminance of the fixed red area relative to the green varies from observer to observer because of individual differences in relative sensitivity (Cavanagh, Anstis, & Macleod, 1987), but it can be assumed to have the same value, say L_R for both sets of readings (red in nonshadow area and red in shadow area). If we assume that the threshold is the same, in contrast units, for the two arrangements of the red and green in the figure, then we can derive the threshold contrast from the two green luminance values—the first, L_{GH}, when green is in the nonshadow area and presumably more luminous than the effective red luminance; the second, L_{GL}, when it is in the shadow area and presumably less luminous. The threshold figure contrast, C_F, for the chromatic image can then be derived as

$$C_F = (\sqrt{L_{GH}} - \sqrt{L_{GL}})/(\sqrt{L_{GH}} + \sqrt{L_{GL}}). \quad (1)$$

For the single color (monochromatic) images, both shadow and nonshadow regions were shown in red for four settings and green for another four. The nonshadow region was fixed at 20 cd/m^2, and the observers adjusted the luminance of the shadow region until the surface relief of the face or cup just disappeared. Threshold figure contrast was taken as the average of the Michelson contrasts for the red and green images.

The 2 observers, PC (one of the authors) and LM, had normal color vision and normal or corrected-to-normal acuity.

Results

It was possible to perceive 3-D shape for the chromatic images of the faces and the cup when there was sufficient luminance contrast indicating that the presence of impossible shadow colors does not prevent the perception of shape from shadows. There were no systematic differences among the data for the three different stimuli. The results, averaged across the three stimuli, are shown in terms of figure contrast, the Michelson contrast between the mean luminance of the shadow and nonshadow areas (their difference divided by their sum). The figure contrast (Figure 6) at which the chromatic figures were correctly interpreted was similar to that necessary to perceive the faces or the cup when presented in a single color for one observer (LM), but additional contrast was required for the other observer (PC). Because the figure contrast necessary for the chromatic stimulus was greater than zero, we conclude that color alone cannot support the perception of 3-D shape from shadows. Given that the perception of the shadows in the chromatic figure required the same or somewhat higher figure contrast than it did in the monochromatic image, there may also be some interference from the colors. DeValois and Switkes (1983) have demonstrated that color can increase luminance contrast threshold by a factor of up to 2 or 3, and we (Cavanagh, Shioiri, & MacLeod, 1987) have recently found that there are individual variations in this interaction between color and luminance that are consistent with the differences we find in this experiment.

In summary, color alone cannot signal shadow areas even though the 2-D shape is clearly visible—a luminance difference between shadow and nonshadow areas is necessary. An impossible shadow color does not suppress shape from shadows although for 1 subject the presence of color did interfere to some extent with the 3-D shape recovery. A saturated green shadow is impossible if the surround is a saturated red (the reverse is also impossible), but the shadow figures were easily recognized under these conditions as long as sufficient luminance contrast was present.

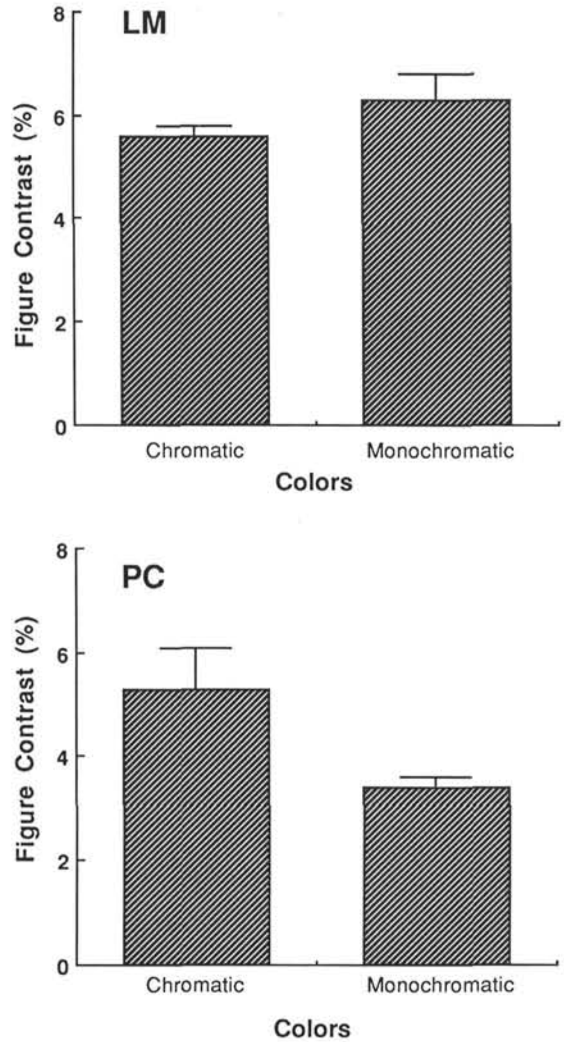


Figure 6. Figure contrast at which depth due to shadows just disappeared in the chromatic images (red/green) and monochromatic images (red/red and green/green) for Observers LM and PC. (Vertical bars show standard error [+1 *SE*].)

Experiment 2: Motion

When a cast shadow falls across a surface, its position on that surface is in a sense arbitrary, being determined by the object casting the shadow and the light source as well as by the shape of the receiving surface. We would not expect a change in the motion of the surface to be aligned with the shadow border. Relative motion between two regions is a strong cue to the depth relations between the two areas and often gives rise to compelling figure/ground organizations (Anstis, 1970; Braddick, 1974; Julesz, 1971; Ullman, 1979). There are therefore two reasons for relative motion between shadow and nonshadow regions to suggest that the border involves a change of material or surface and not a shadow falling across an unbroken surface: First, the probability that motion and brightness changes are perfectly aligned along an extensive border is exceedingly small, and second, the motion itself should give cues to the organization of the surfaces, in particular, that there are two surfaces moving relative to each other (Figure 7).

The shadow figures—the woman's face, the mannequin's face, and the coffee cup—were filled with random texture both in the shadow and nonshadow region. The texture of either region could be made to move while the texture in the other region remained stationary. The experiment involved manipulation of two different contrasts in the figures: the texture contrast, set by the experimenter, and the figure contrast, adjusted by the observer. The texture contrast is the Michelson contrast between the lightest and darkest texture elements within a single region, and both shadow and nonshadow regions were set to the same texture contrast for any given condition. The texture contrast was varied from 0% to 75%. The figure contrast, on the other hand, is the Michelson contrast between the mean luminance of the shadow region and the mean luminance of the nonshadow region. The observer could adjust the figure contrast until the surface relief of the face or cup was just visible (measurements were

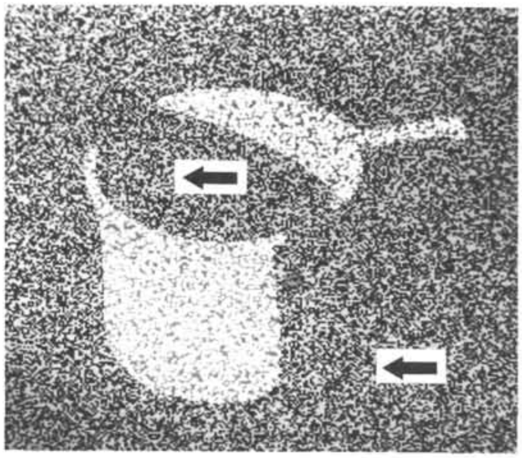


Figure 7. The shadow and nonshadow regions of the cup are distinguished by motion. (As indicated by the arrow, the shadow area would be moving, and the nonshadow area would be static in this example. A luminance difference between the shadow and nonshadow regions is also present here. Stimuli in the experiment included the cup, the woman's face, and the mannequin's face.)

made on only one image at a time). If 2-D shape defined by motion could convey shadow information, we would expect that no figure contrast (0% contrast between the shadow and nonshadow regions) would be necessary for the depth from shadows to be visible. If, on the other hand, relative motion of the two regions interfered with the shadow interpretation, we would expect that additional figure contrast would be required when the textures were moving compared with when they were stationary.

Method

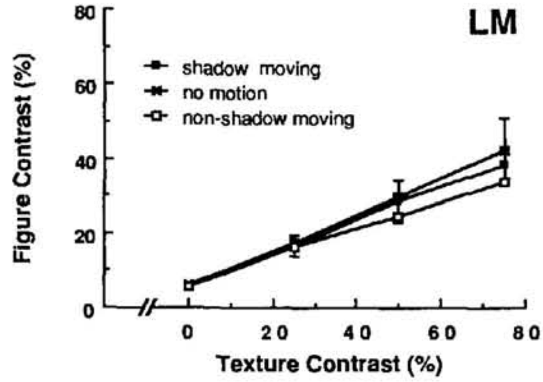
The stimuli—the woman's face, the mannequin's face, and the coffee cup—were identical to that of Experiment 1 except that they were presented in black and white, and a random texture of 256 × 240 checks, 50% light and 50% dark, each 0.03° square, filled the display area. The texture was set to 0%, 25%, 50%, or 75% contrast (Michelson contrast between the light and dark checks) in both shadow and nonshadow areas. Three separate conditions were used: texture moving in the shadow area, texture moving in the nonshadow area, no texture moving. Texture motion was 2.8°/s leftward. Although the textures moved, the areas within which the textures moved always remained stationary. A fixation point was used to prevent tracking of the moving textures. The mean luminance of the lighter, nonshadow areas was fixed at 40 cd/m^2, and the observer adjusted the mean luminance of the texture in the darker region until the 3-D shape from shadows just disappeared. The figure contrast at the transition point where the 3-D shape just disappeared is reported as the Michelson contrast between the mean luminance of the shadow and nonshadow regions.

Four readings were taken in each condition, and 2 observers, PC and LM, were used.

Results

When the texture contrast was 0% (i.e., untextured), the threshold figure contrast for shape from shadows was 3% to 5% (Figure 8), as it was for the single color stimuli of Experiment 1. When the texture contrast was increased and the relative motion became visible, it was evident that 2-D shape defined by motion alone could not support the perception of a shadow. Even though the 2-D shape of the figure could be seen clearly when the texture in the shadow or nonshadow region was moving and both regions had equal luminance (0% figure contrast), no impression of a face or of surface relief was visible until a luminance difference between shadow and nonshadow areas had been added in. In fact, much *more* figure contrast was necessary in these conditions than when there was no texture and no motion. The data revealed that this increase in threshold figure contrast was due to the presence of the texture, not the motion. The figure contrast at which the face or cup became recognizable when the texture in one or the other region was moving (Figure 8) was not substantially different from the necessary figure contrast when neither was moving.

The relative motion did produce a clear segregation of the shadow and nonshadow regions into two planes, with the static area appearing to float in front of the moving background. Although this did not interfere with the interpretation



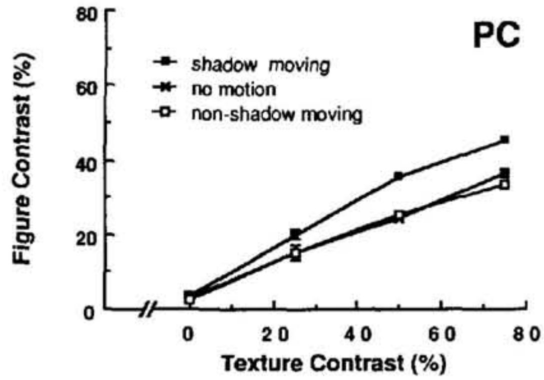


Figure 8. Figure contrast at which depth due to shadows just disappeared when the texture in the dark, or shadow, areas was moving, when the texture in the light, or nonshadow, areas was moving, and when neither was moving, as a function of texture contrast. (Observers LM and PC. Vertical bars show typical standard errors [±1 *SE*].)

of the shadows, it did produce a secondary effect of transparency as if the figures were etched on a transparent surface, with the static areas opaque and the moving areas transparent. In the front plane, the face or cup could be seen as an integrated figure with surface relief, but it was also possible to look through some areas of the figure to see a moving, textured background plane.

The change in motion at the shadow border violates the physical constraint that a single surface should not change its speed at a shadow boundary. However, the face could be seen whether the dark or light region moved or whether they both moved in the same direction at different speeds or in opposite directions (these last two conditions were verified separately). In all of these conditions, light and dark regions belonging to the same surface, such as the cheek, were moving at different speeds. This violates every reasonable assumption we could make about a surface such as a cheek. We conclude that motion cues had surprisingly little effect on the interpretation of shadows in our stimuli. Nevertheless, these cues did seem to invoke a secondary organization of transparent surfaces. The presence of the texture, on the other hand, clearly degraded the perception of the shadow figures.

Experiment 3: Binocular Disparity

As with motion, a change of depth along a cast shadow border is an unlikely occurrence. Certainly, binocular cues to depth should predominate in indicating the surface organization of our figures. We therefore repeated the motion experiment but now used stereoscopic random texture to display the shadow and nonshadow areas at different depth planes.

Method

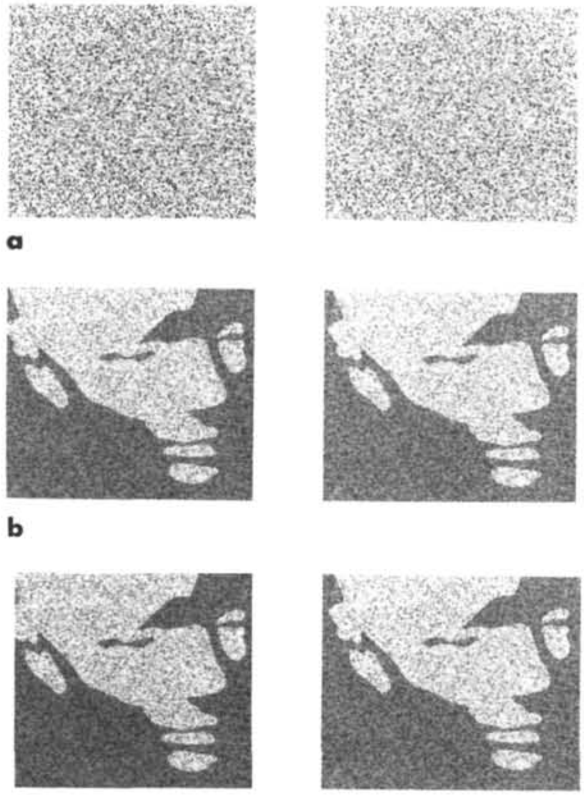
The stimulus presentation was identical to that of Experiment 2 except that the textures were presented to the left and right eyes (using red/green anaglyphs) with different disparities for the shadow and nonshadow regions (Figure 9). When the figure contrast was greater than zero, the shadow and nonshadow areas defined by luminance had the same disparities as the textures that filled them. Conditions were presented with the nonshadow areas in front, with the shadow areas in front, or with both on the same plane. The disparity was +0.09° to place the shadow areas in front, −0.09° to place the nonshadow area in front by the same amount, and 0.0° for both areas to appear on the same plane. The disparity corresponded to a highly visible depth difference of 10 cm at the viewing distance of 1.93 m and image size of 27 cm. The texture contrast was set to 0%, 25%, 50%, or 75% contrast. The mean luminance of the lighter, nonshadow areas was fixed at 40 cd/m^2, and the observer adjusted the mean luminance of the texture in the shadow region until the 3-D shape from shadows just disappeared. The figure contrast at the transition point where the 3-D shape of the shadow figures just disappeared is reported as the Michelson contrast between the mean luminance of the shadow and nonshadow regions.

Four readings were taken in each condition, and 2 observers, PC and LM, were used.

Results

The results were quite consistent across the three stimuli, and Figure 10 shows the average values. These data are very similar to those for motion-defined stimuli. When the texture contrast was 0%, the stimulus was identical to the corresponding stimulus in Experiment 2 and to the monochromatic stimulus (except for color) in Experiment 1. As expected, then, the threshold for perception of the shadow figures was again 3% to 5%. When texture contrast was increased and depth from binocular disparity became visible, the different areas of the figures could be seen when both regions had equal luminance (0% figure contrast), but the overall organization could not be seen (Figure 9a shows the mannequin's face). With an appropriate brightness difference added in, however, the shadows and the figure were visible even though the depth change was clearly seen at the brightness edge (Figure 9b). Again, much more figure contrast was necessary to see the 3-D shape from shadows in these conditions than when there was no texture and no disparity, but this increase in threshold figure contrast was clearly attributable to the presence of the texture and not the binocular disparity. The figure contrasts at which the depth from shadows just became visible were not appreciably different with or without the presence of binocular disparity.

Observers again reported a secondary organization of the figure into two planes. The stimulus was seen on one plane



a

b

c

Figure 9. Panel a: A stereo pair of the woman's face with no luminance difference between the originally light and originally dark areas. Panel b: Stereo pair with an appropriate brightness difference and with the dark areas in front of the light areas. Panel c: Same as Panel b but with the dark areas in back. (The images *must* be viewed with *crossed* disparity for the disparities of the luminance and the texture to correspond.)

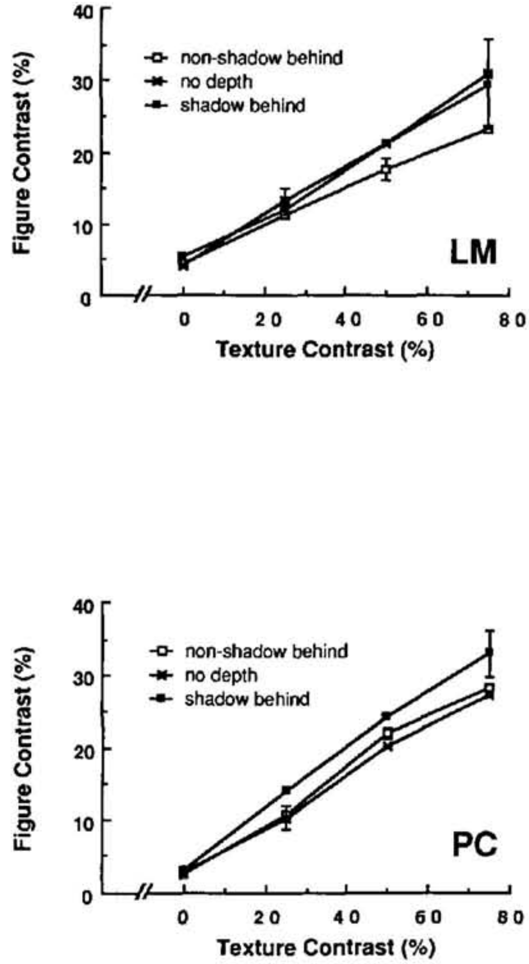


Figure 10. Figure contrast at which depth due to shadows just disappeared when the dark (shadow) areas were seen in front, when the light (nonshadow) areas were seen in front, and when all areas were seen on the same plane, as a function of texture contrast. (Observers LM and PC. Vertical bars show typical standard errors [±1 *SE*].)

as an integrated figure with surface relief, but parts of the figure also appeared to be transparent, and a second depth plane was visible through them. With the dark region in front (Figure 9b), the light region also appeared to be in front but transparent so that a rear depth plane could be seen through it. With the light region in front (Figure 9c), the dark region appeared to be transparent.

In summary, the depth information from binocular disparity did not interfere with the shadow interpretation, but the shape signaled by binocular disparity could not support the perception of shadows. It appears that the interpretation of the shadows ignored binocular disparity and relied solely on luminance information. The binocular disparity was not completely ignored, however, because a secondary transparency organization emerged in the figures, as it had in the figures involving relative motion. The figure appeared on one depth plane with transparent areas through which a rear depth plane was visible. The figure on one depth plane then gave rise to a 3-D interpretation of a face or a cup in the same way as a flat drawing generates a 3-D impression of the object it represents.

As in the previous experiment, the presence of texture degraded the perception of the shadow figures.

Experiment 4: Texture

If a dark, patterned area in a rug is surrounded by a lighter area with a different pattern, the lightness difference is more likely due to a pigment (reflectance) change in the rug than to a shadow that happens to be aligned everywhere with the pattern change. When a shadow falls on a textured surface, such as a rug, the brightness of the texture should be reduced in the shadow area, but neither its contrast nor the texture pattern itself should be affected. This texture constraint can be verified in two different ways: as an area constraint or as a border constraint. First, texture parameters such as contrast, orientation, and average element size for the shadow area should be the same as those for the nonshadow area. In Figure 11, for example, the contrast between the lightest and darkest texture elements must be the same in the nonshadow (C_{NS}) and shadow (C_S) regions. If it is not, this is evidence that there is a material change between the two points and that the border is therefore not a shadow border.

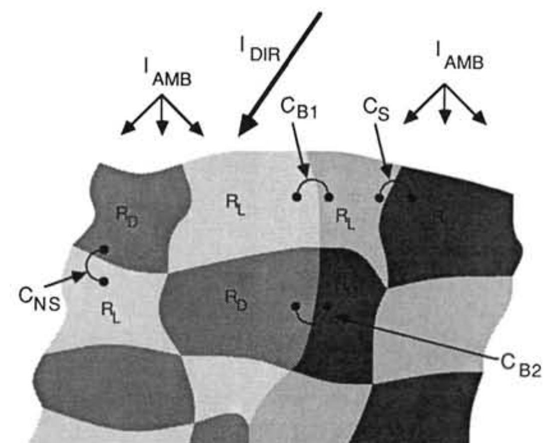


Figure 11. A shadow border crossing a textured surface. (The reflectances of the lightest and darkest texture elements, R_L and R_D, should be the same on both sides of the shadow border. The direct illumination, I_{DIR}, falls only on the nonshadow region, but the ambient illumination, I_{AMB}, if there is any, falls in both shadow and nonshadow areas. The contrast between the lightest and darkest texture elements in the shadow area, $C_S = (R_L - R_D)/(R_L + R_D)$, must be the same as that in the nonshadow area, $C_{NS} = (R_L - R_D)/(R_L + R_D)$. The contrast between adjacent points across the shadow border should be the same at all points, $C_{B1} = C_{B2} = I_{DIR}/(I_{DIR} + 2\ I_{AMB})$, with the exception of chance alignments of the texture element borders and the shadow border such as at C_{B3}, for example.)

A difference in texture contrast between shadow and nonshadow regions does not invariably indicate a change of material, however. It can occur if there are two light sources and if the texture results from the spatial variation of the secondary light source, not from variation in the reflectance of the surface. A textured light such as reflections off a rippled water surface can produce a high-contrast speckle in a shadow area not illuminated by the primary source but a lower contrast speckle in the nonshadow areas that also receive light from the primary source. The reverse, higher contrast texture in the nonshadow region, can occur in at least two situations with a single light source: First, when the shadow is in total darkness, the texture there will effectively have zero contrast; and second, if the sole light source itself is textured, the directly lit areas will have a texture while the shadow areas will have none. Given these possible, though infrequent, violations of the texture area constraint in natural images, violations of the constraint may have only a moderate effect on the acceptability of shadow figures. We examined the area constraint in figures having higher contrast texture in the shadow region than in the surrounding region (Figure 12b), an arrangement that is impossible for shadows in natural scenes having a single illuminant.

The second approach to verifying the texture constraint involves the contrast across the shadow border. A shadow border falling across a textured surface will divide the texture elements through which it runs into lighter and darker parts. Because a texture element has the same reflectance on both sides of the border, the contrast across the border is simply the contrast between the light falling on the nonshadow side (direct plus ambient illumination) and that falling in the shadow (ambient illumination only). The (Michelson) contrast between adjacent points across the border should therefore be the same all along its length ($C_{B1} = C_{B2}$, in Figure 11), even though the texture elements themselves are varying in reflectance along the border. The relation specified by this border contrast constraint is generally, but not always, true. The exceptions occur when the border between two texture elements happens to be aligned with the shadow border (e.g., at C_{B3} in Figure 11), an event that should be relatively infrequent. A weaker version of the border constraint is that the luminance difference between adjacent points across the shadow border should always be in the same direction, darker on the shadow side. The only exception to this border polarity constraint is again the chance alignments of texture element borders and shadow border.

A stimulus with a real material change exactly aligned with the shadow border can have different reflectances in the shadow and nonshadow areas (i.e, the lightest and darkest reflectances, R_L and R_D, are different in the two areas). In this case, both the texture area constraint ($C_S \neq C_{NS}$, Figure 11) and the border contrast constraint ($C_{B1} \neq C_{B2}$) will be violated, although the weaker border polarity constraint ($C_{B1} \cdot C_{B2} \geq 0$) may be satisfied. The converse is not true, however; a stimulus that violates the border contrast constraint may not violate the area constraint, and this possibility allows us to examine the roles of these different constraints independently.

If we take the checkerboard textures of Figure 12 and exchange the light and dark checks in the shadow region only, we keep the area properties such as texture contrast within the regions and figure contrast between the regions the same (spatial frequency content changes very little) and change only the relations between adjacent points across the shadow border (Figures 12c, 12d). Rather than always being darker on the shadow side by the same ratio, there are now parts of the border with high contrast between adjacent points across the shadow border and other parts with low contrast between adjacent points, a violation of the border contrast constraint that should signal a material change. In the extreme case, when the light checks in the shadow region are lighter than the dark checks in the nonshadow region, exchanging the light and dark checks in the shadow area produces a contrast reversal at some points along the shadow border (Figures 12c, 12d).

These particular conditions formed part of a general evaluation of the effect of texture in the shadow and nonshadow

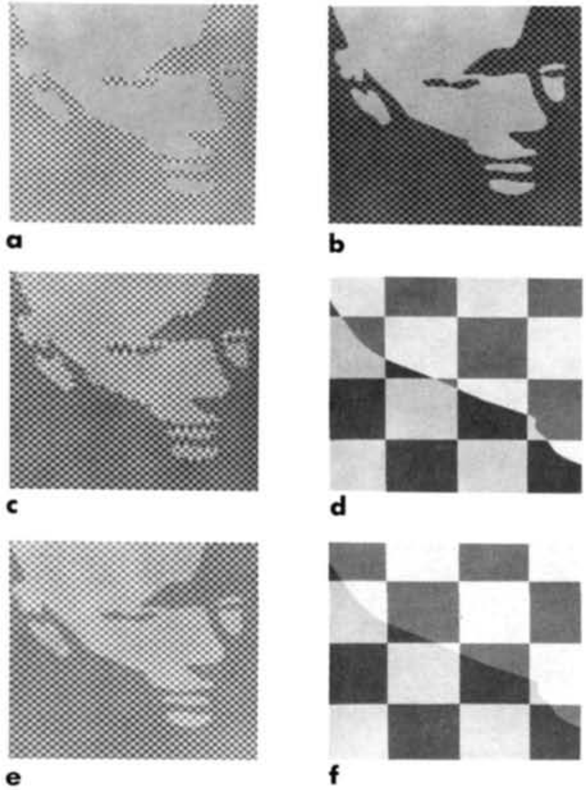


Figure 12. Panel a: Depth due to shading cannot be perceived on the basis of a texture difference alone (shadow texture contrast 60%, nonshadow texture contrast 0%, figure contrast 0%). Panel b: Depth due to shading can be perceived if the shadow areas are darker, even though there is texture in the shadows violating the texture constraint (texture contrasts as before, figure contrast now 50%). Panel c: A figure with the texture out of phase at the shadow border (border detail shown in Panel d) is more difficult to recognize (requires more figure contrast) than the otherwise identical figure (shown in Panel e) with the texture in phase at the shadow border (border detail in Panel f). (The texture contrasts in Panels c through f are 60% [shadow texture contrast −60% in Panels c and d] and the figure contrast 33%.)

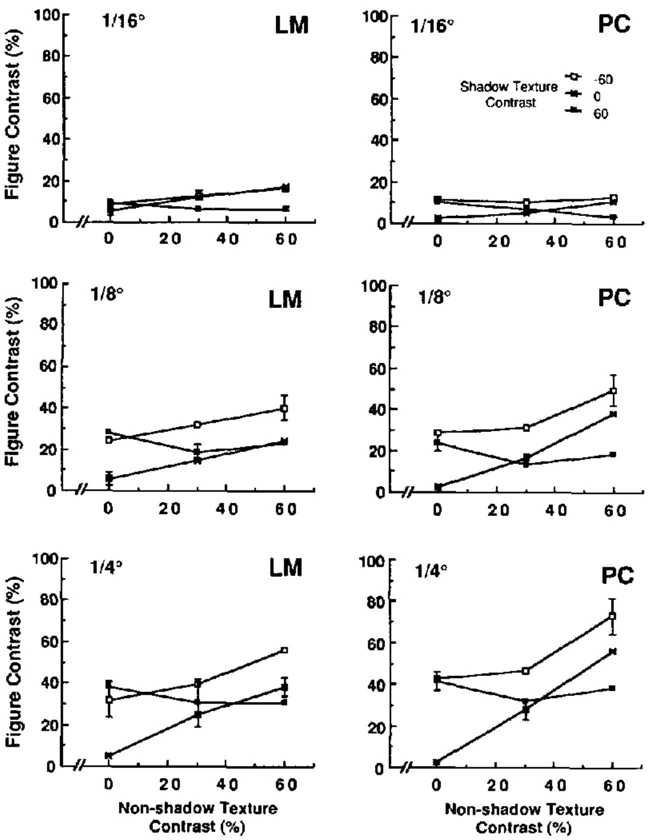


Figure 13. Figure contrast at which depth due to shadows just disappeared as a function of the dark (shadow) texture contrast, the light (nonshadow) texture contrast, and the texture element size. (The results for the three shadow figures have been averaged. Left-hand panel = Observer LM; right-hand panel = Observer PC. Vertical bars show typical standard errors [±1 *SE*].)

regions where we examined the influence of texture constraints and the ability of shape defined by texture to support the perception of shadows. In addition to the woman's face, we also used the mannequin's face and the coffee cup as stimuli. Because the effect of texture must depend on its visibility, we also presented the textures at three different element sizes.

Method

The dark or light areas of the three stimuli were replaced with textures whose mean luminance and contrast could be controlled independently. The textures that filled the shadow and nonshadow areas were made up of checkerboards of light and dark squares (Figure 12) of three possible sizes: 1/16°, 1/8°, and 1/4°. The mean luminance of the nonshadow areas was fixed at 40 cd/m^2, and the observer adjusted the luminance of the shadow area until the shape from shadows just disappeared. Three contrast levels of the texture in the light, or nonshadow, areas were used: 0%, 30%, and 60%. These contrasts refer to the Michelson contrast between the light and dark checks of the checkerboard. Similarly, there were three contrast levels of the dark, or shadow, areas: −60%, 0%, and +60%. The light and dark checks were in phase with the surrounding nonshadow texture in the +60% condition (Figures 12e, 12f) and out of phase (Figures 12c, 12d) in the −60% condition. Each observer made four readings in each of the 81 conditions.

Two observers, PC and LM, made settings.

Results

The results for the three stimuli (the two faces and the coffee cup) were very similar, and the data presented in Figure 13 are averaged over the three stimuli. The most noticeable result is first, as in Experiments 1 and 2, the presence of texture increases the figure contrast threshold, and second, finer textures produce less interference.

We would expect finer textures to interfere less because the drop-off in visual system response to high spatial frequencies lowers their effective contrast. Also, the checks in the larger texture may be as large as some individual stimulus features producing a feature-specific masking.

It cannot be simply the presence of texture contrast that interferes with the shadow perception, however. Increasing the nonshadow, or light, area texture contrast increases the threshold figure contrast when the dark texture contrast is 0% or −60%, but surprisingly, the same increase in light area texture actually improves performance (the threshold decreases) when the dark texture is at 60% contrast (Figure 13). We cannot attribute the effect of texture solely to masking, therefore, but must look more closely at the interactions between the textures in the two areas to understand this result.

When there was no texture in either shadow or nonshadow areas (0% light texture contrast and 0% dark texture contrast), 3%–6% figure contrast was sufficient for the shape from shadows to be perceived. When the texture contrast was different in the shadow and nonshadow regions and there was no difference between the mean luminances, the various areas of the stimuli, defined only by texture differences, could be clearly seen, but the shadow areas were not perceived as shadows. (Figure 12a is an example.) The texture difference alone could not support the perception of shadows. However, if the shadow regions were made sufficiently dark, then the shadow figures could be recognized (Figures 12b, 12e).

Violation of the texture area constraint (and necessarily the border contrast constraint) did not appear to interfere with the perception of the shadows. (Figure 12b is an example.) When there was higher contrast texture in the shadow than in the nonshadow region (60% dark texture contrast and 0% or 30% light texture contrast), the threshold figure contrast was the same as, or only slightly higher than, the condition (60% contrast in both light and dark textures) that did not violate these constraints.

On the other hand, the border polarity constraint appeared to play a significant role. In particular, when the textures in the two regions had the same contrast but were out of phase (60% and −60%, light and dark contrasts, respectively, Figures 12c, 12d), the required figure contrast was about twice as high as when they were in phase (60% and 60%, Figures 12e, 12f). Even though the threshold had increased, there was still a figure contrast at which the 3-D organization of the stimuli was visible. The constraint that was influencing the results could not be that the contrast between adjacent points across the shadow border should remain the same all along the border because this would still be violated at higher figure contrasts. The border constraint may therefore apply only to the extreme cases where the polarity of contrast actually *reverses* at different points along the shadow border. If this were the case, the observer would then increase the luminance difference between the shadow and nonshadow regions until there were no polarity reversals along the border.

Was this the basic factor that determined all the settings that the observers made? If we look at the figure contrasts at which the reversal of polarity is just eliminated *in the stimulus* (Figure 14) for all the combinations of in- and out-of-phase contrasts, the pattern is quite similar to our observed data. This possibility must be qualified to take into account the visibility of the texture at the border. If the contrast reversals occur for a texture so fine that it is invisible, then they can have no effect. We conducted a multiple regression analysis of the data of Figure 13 to determine how much of the data variance could be predicted by this polarity constraint. The predictor for the polarity constraint was simply the figure contrast values at which the contrast reversals are just eliminated in the stimulus (as shown in Figure 14).

We also included the dark texture contrast and the light texture contrast as predictors because there are additional reasons to assume that textures that we used might interfere with the 3-D interpretation. Strong cues to depth are provided by changes in texture element size with distance and by perspective cues in regularly organized textures on 3-D surfaces. Because the textures we used were spatially uniform, without gradients corresponding to the surface orientations, they provided information that the textured surfaces were flat, whereas shadow information supported a different 3-D organization. The presence of uniform textures throughout the background and into the figure areas also may have supported the grouping of connected shadow or nonshadow regions (for example, the shadow area inside the cup and the dark background, Figure 4) even though there are implicit object contours running through these areas when the shadows are correctly interpreted. In addition to these inappropriate pictorial cues in our textures, the texture could also be simply masking the visibility of the mean luminance difference between the shadow and nonshadow regions. Similar figure masking has been reported for face perception (Tieger & Ganz, 1979).

The results of the multiple regression (Figure 15) prediction of the figure contrast, C_F, are shown in terms of the regression coefficients of Equation 2:

$$C_F = \alpha + \beta_S C_S + \beta_{NS} C_{NS} + \beta_B C_B , \tag{2}$$

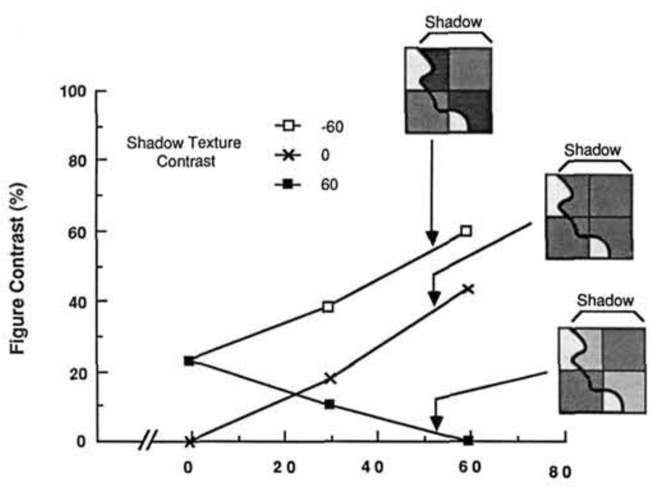


Figure 14. Stimulus figure contrast at which contrast reversals at the shadow border are just eliminated as a function of light and dark area texture contrasts. (Insets show examples of shadow borders for these conditions. The condition of no reversal occurs when some adjacent points across the border may have the same brightness but when no point at the border on the dark, shadow side is brighter than its adjacent point on the light, nonshadow side. At figure contrasts less than the values plotted for each texture contrast condition, contrast polarity reversals occur along the border for those texture contrasts. At higher figure contrasts, border contrast is in the same direction all along the border.)

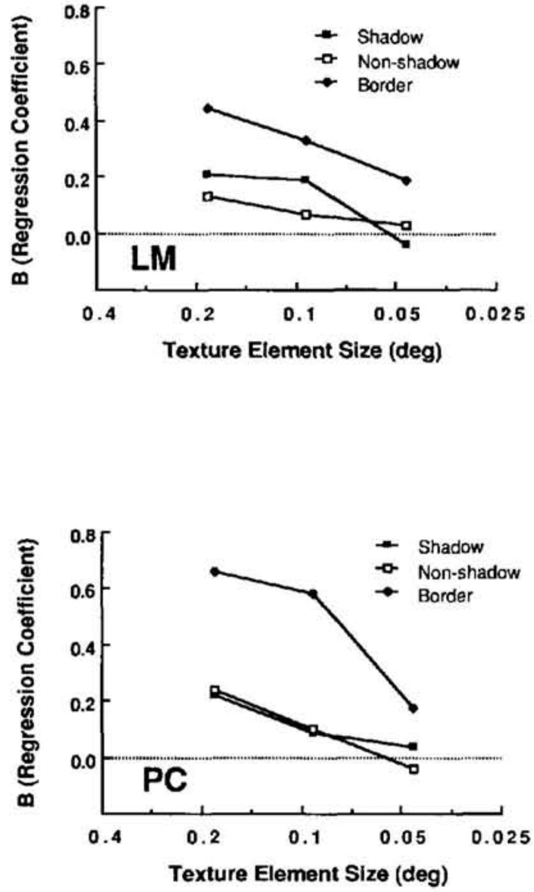


Figure 15. Regression coefficients of three factors in predicting the data of Figure 13 as a function of texture element size for Observers LM and PC. (The factors used were the light area and dark area texture contrasts and the figure contrasts at which polarity reversals are eliminated [as shown in Figure 14].)

where α is a constant; β_S, β_{NS}, and β_B, are the regression coefficients for the three predictors; C_S, the texture contrast in the shadow area; C_{NS}, the texture contrast in the nonshadow area; and C_B, the figure contrast at which there are no polarity reversals at the border (Figure 14), respectively, at each of the texture sizes. The percent of the variance explained was uniformly high, an average of 93% for PC and 96% for LM. Evidently, the border polarity constraint was the most important predictor of the data while the texture contrasts made significant (at large and intermediate texture sizes) but smaller contributions. If the border polarity constraint were an absolute condition for shadow visibility, its regression coefficient would have been 1.0 in all conditions—the observers' figure contrast settings would be identical to those of Figure 14. However, it is quite a bit less than 1.0 and gets smaller with decreasing texture size. This indicates that the visibility of the contrast reversals is an important but not absolute determinant of shadow acceptability: There may be some minimum tolerated visibility of reversals. This tolerance for some reversals may be necessary to allow for the occasional reversals in natural images due to chance alignments of texture and shadow edges. In addition to the large contribution from the contrast polarity factor, there was also a smaller contribution from the shadow and nonshadow textures at the two larger texture sizes. Because the effect of the texture contrast was the same in both shadow and nonshadow regions, this likely represents masking of the figure contrast by the texture contrast although the inappropriate pictorial cues in the texture may also contribute to the interference.

In summary, the acceptability of a shadow region depends on consistent contrast polarity along the shadow border. The shadow region must be darker, and there can be no more than a minimum of visible contrast reversals along the border. Consequently, stimuli defined by texture alone, with equal luminance in shadow and nonshadow regions, cannot support the perception of shadows because there are necessarily numerous contrast reversals along the border. Other aspects of natural constraints concerning texture appear to be ignored. As long as the polarity constraint was satisfied, shape from shadows could be seen when there were different textures in the shadow and nonshadow regions, when the textures were similar but had different contrasts, and when the contrast across the shadow border changed from point to point.

The polarity constraint may also account for the effect of texture contrast seen in Experiments 2 and 3. The greater the texture contrast in those experiments, the higher the figure contrast necessary to see the 3-D shape from shadows. A simple masking explanation would make the same prediction, however, so there is no direct way to disentangle the relative effects of the two factors from the data of Experiments 2 and 3. The multiple regression of Equation 2 does separate the two factors and leads us to propose that the border polarity constraint is the more important factor.

Experiment 5: Texture Size

Is the influence of contrast reversals a function only of the size of the texture elements or of both texture size and image size? We examined the figure contrast necessary to perceive the 3-D organization of shadow figures for four different sizes of texture element and three different sizes of the stimulus by using the most difficult condition of the previous experiment. The textures in the shadow and nonshadow areas were both at 60% contrast but out of phase at the shadow border (Figures 12c, 12d).

Method

The stimulus presentation was identical to that of Experiment 4, except that only the woman's face was used as a stimulus, and it was presented at three different sizes: 2°, 3.3°, and 8° horizontal width, with a 1.09 vertical/horizontal aspect ratio. The textures that filled the shadow and nonshadow areas were made up of checkerboards of light and dark squares (Figure 13) of four possible sizes: 1/32°, 1/16°, 1/8°, and 1/4°. The mean luminance of the nonshadow areas was fixed at 40 cd/m^2, and the observer adjusted the luminance of the shadow area until the surface relief of the stimulus just disappeared. The contrast level of the texture in the light, or nonshadow, areas was +60%, and in the dark, or shadow, areas it was −60%. The light and dark checks were out of phase across the shadow border as shown in Figure 12d. Two observers, PC and LM, made four settings for each of the 12 conditions.

Results

The amount of figure contrast necessary to perceive the 3-D organization (Figure 16) decreased with texture element size as it did in Experiment 4. There was a very consistent effect of stimulus size as well. The smaller stimuli required more figure contrast for the same texture element size. This implies that the contrast reversal along the border interferes with the encoding of the border features and that the importance of the reversals depends on their scale *relative* to the border features. It does not appear to be the case that the presence of visible reversals is sufficient to veto a shadow border because contrast reversals that blocked the perception of shape from shadows for one image size did not do so for a larger image size.

The texture may be masking the image features as well, and if so, the texture size producing the most effective masking may depend on the image size. Although this is a possible contributing factor, Experiment 4 showed that the contrast polarity was a more important factor than was simple figure masking. Another possible explanation is that the bright texture elements of the shadow area that abut the border are being grouped into the brighter nonshadow area at the border. This would disrupt the border features more when the texture elements are large relative to the border features. This grouping effect does appear to occur when the brightness of the light elements in the shadow area equals that of the abutting dark elements in the nonshadow area. (This happens only at 60% figure contrast for the conditions of this experiment.) These adjacent elements then group together to produce a serrated edge at the border. There is no noticeable grouping effect at other contrast levels, however, and the 60% figure contrast that produces the serrated border does not appear to be a significant barrier to visibility. Several figures were visible at lower contrasts and several required higher contrasts.

Our conclusion, therefore, is that the assertion of border features is a cooperative process that is influenced by operators

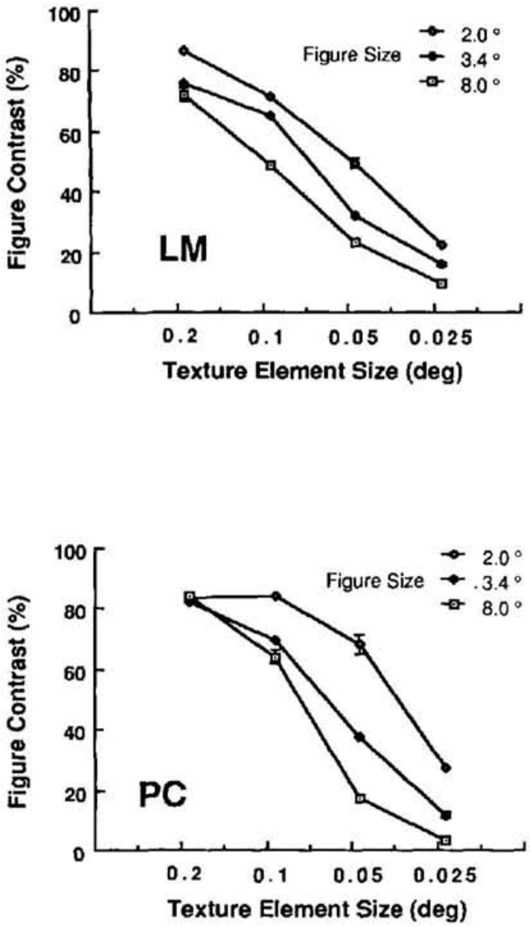


Figure 16. Figure contrast at which depth due to shadows just disappeared as a function of the texture element size and the image size. (Observers LM and PC. Vertical bars show typical standard errors [±1 *SE*].)

signaling border contrast over a range of scales that is particular to the scale of the border feature. Large features can be asserted even in the presence of small scale inconsistencies in border contrast polarity, whereas small features require consistency down to smaller sizes. Cooperation across scales has been proposed for other edge operators (Marr & Hildreth, 1980), but it runs into difficulties when trying to localize curved edges. Operators at different scales will place the edge of the identical, curved stimulus at different radii (Asada & Brady, 1985, Witkin, 1983), and the operators cannot cooperate very effectively if, by nature, they must conflict along many parts of an edge. Our data in this experiment, however, indicate that perhaps only a limited range of scales is involved in the cooperative process, avoiding the problems raised by Witkin. The next experiment examines this possibility.

Experiment 6: Spatial Frequency

In studying the effect of texture on shadows, we considered the shadow border and its contrast polarity element by element. The visual system encodes the image not only point by point but also at a range of scales for each point. In particular, at the cortical level, each local region is represented by oriented receptive fields covering a large range of sizes and all orientations. The importance of the contrast reversals of small texture elements at the shadow border in Experiment 4 indicates that the shadow border is not detected simply by large receptive fields oriented along the shadow border. They would average over small-scale contrast reversals and see only the mean luminance difference between the two areas, a difference that would support the shape from shadows. (Squinting at the images in Figures 12b and 12c produces the same effect.) The data therefore suggest some cooperative process that evaluates the border across a range of scales and rejects it as a potential shadow border if there is inappropriate contrast at one of the scales. The effective range of scales is most likely a function of the image size. In particular, the rate of curvature of important border features puts a lower limit on the spatial frequency of a filter than can follow the curved border.

To evaluate the contribution of different scales, as well as their interactions, we presented filtered shadow images. First, we asked observers to judge whether the images were correctly perceived when low-pass or high-pass filtered (Figures 17a, 17b). Note that the high-pass images have no difference in mean luminance between the shadow and nonshadow region but do have the appropriate brightness difference at the border. Second, the observers judged bandpass and notch-filtered images (Figures 17c, 17d). Observers could see the 3-D organization in all of these versions as long as sufficient informa-

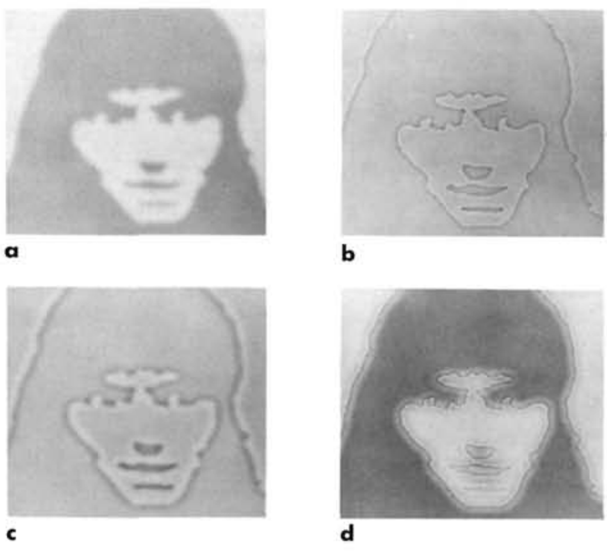


Figure 17. Filtered versions of the woman's face: low-pass filtered (Panel a), high-pass filtered (Panel b), bandpass filtered (Panel c), and notched filtered (Panel d). (The low- and high-pass filters had a half-amplitude spatial frequency of 2.0 cpd [16.0 cycles per image]. In the high-pass image of the face, the Craik-Cornsweet illusion induces the perception of a luminance difference between the central regions of shadow and nonshadow areas when, in fact, they have identical luminances. The bandpass and notch filters both had a half-amplitude bandwidth of 2.0 octaves and center spatial frequency of 2.0 cpd [16.0 cycles per image].)

tion was present from within a band of spatial frequencies from about 1.5 to 6.0 cycles per degree (cpd).

We devised a nulling technique to evaluate the strength of contribution of each of these filtered components to the perception of the shadow image. The image was filtered and presented at maximum positive contrast (100%). The observer could then add a negative, unfiltered version of the same image to this (Figure 18). The observer adjusted the contrast of the negative image until it just nulled the perception of the shadow figure. The ratio of the contrast of the negative and positive images was taken as the relative strength of the filtered component in contributing to the visibility of the shadow image. The more negative the contrast required to null the shape-from-shadows organization of the figure, the stronger the contribution of the positive components to shape from shadows. Some specific examples demonstrate that this technique produces a scale of relative strength that can have negative or positive values and that increases monotonically with the contribution of the filtered image to the shape-from-shadows process.

1. If the positive test image is unfiltered, containing all the spectral components, it must have the highest possible relative strength. The strength that we should measure can be easily determined because the negative and positive versions are identical except for a contrast reversal. Increasing the contrast of the negative component of the combined image is therefore equivalent to decreasing the contrast of a positive image. To null the perception of the shadow figure, the negative contrast would have to be raised to about 96% of the positive version's contrast, leaving a remaining contrast of 4%, the threshold measured for untextured images in the previous experiments. The relative strength of an unfiltered image (all-pass filtered) should therefore be about 0.96.

2. If the positive test image is filtered and contributes only a bare minimum of information necessary to support the perception of the 3-D structure, then any amount of the negative, unfiltered image will eliminate the 3-D organization. On the other hand, if the observer reduces the contrast of the negative image below 0%, then he or she is adding in a *positive* image, and this will improve visibility. The threshold setting in this case would therefore be 0.0.

3. If the filtered image did not support the perception of shadows at all, the observer would have to reduce the strength of the negative, unfiltered version until it reversed contrast (becoming a positive, unfiltered image) before any 3-D organization would be visible. In this case the relative strength would be negative, and it would measure the component's *interference* with the shadow perception.

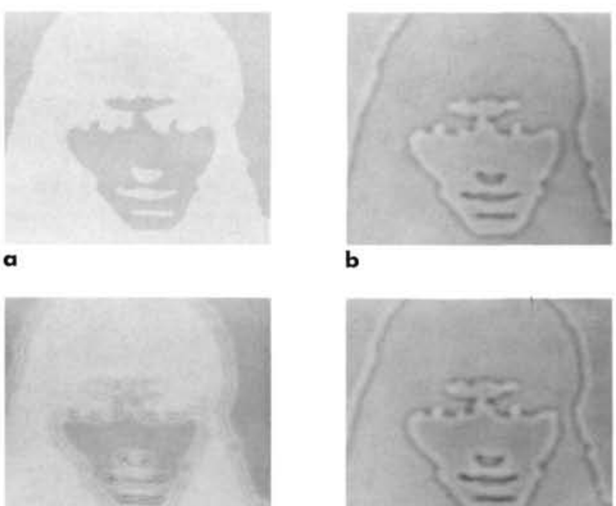


Figure 18. Nulling technique for measuring the relative contribution of filtered components of an image to the perception of depth due to shadows. A negative, unfiltered version of the image (Panel a) is added to the filtered, positive version (Panel b) to produce the combined image (Panel d). (The observer adjusts the contrast of the negative component until the depth due to shadows is eliminated (Panel c) and then finds a setting at the transition between these two organizations of the image.)

Figure 19 shows the frequency spectra of the filters used to produce two of the combined images: one a low-pass image and the other a notch-filtered image, both combined with a negative, unfiltered image having 20% contrast. Figure 19 shows that for the low- and high-pass combined images, the half-amplitude frequency may not be the best characterization of the filter. It may be more important to consider a second factor: the "crossover frequencies" of the resulting spatial frequency spectrum. The crossover frequencies are the points at which the spectrum changes from positive to negative: The positive portions favor the retrieval of shape from shadows, whereas the negative portions interfere with it. For the bandpass and notch images, the two crossover points extend outward from the center frequency of the filter fairly symmetrically as the contrast of the unfiltered negative component is varied. The center frequency of the filter, therefore, corresponds reasonably well with the center of the resulting band of positive or negative information. For the low- and high-pass images, however, the crossover frequency shifts as a function both of the filter half-amplitude frequency and the contrast of the negative unfiltered component. Varying the contrast of the negative component affects the crossover frequency even though the filter half-amplitude frequency remains constant. Consequently, for the low- and high-pass images, we will analyze the results both in terms of relative strength and crossover frequency.

Method

The image of the woman's face (Figure 4a) was convolved with 2-D gaussian kernels to produce low-pass filtered images. Bandpass images were generated by differencing two low-pass images having different gaussian space constants. The spatial frequency at half-amplitude for the low-pass images was 0.51, 1.00, 1.41, 2.00, 2.83, and 8.07 cpd. The corresponding high-pass images having the same half-amplitude spatial frequencies were produced by subtracting the low-pass images from the original images. The center spatial frequencies of the bandpass images were 1.0, 2.0, 4.0, and 8.0 cpd, and all had a full bandwidth at half-amplitude of 2.0 octaves. The corresponding notch-filtered images having the same center frequencies and bandwidths were produced by subtracting the bandpass images

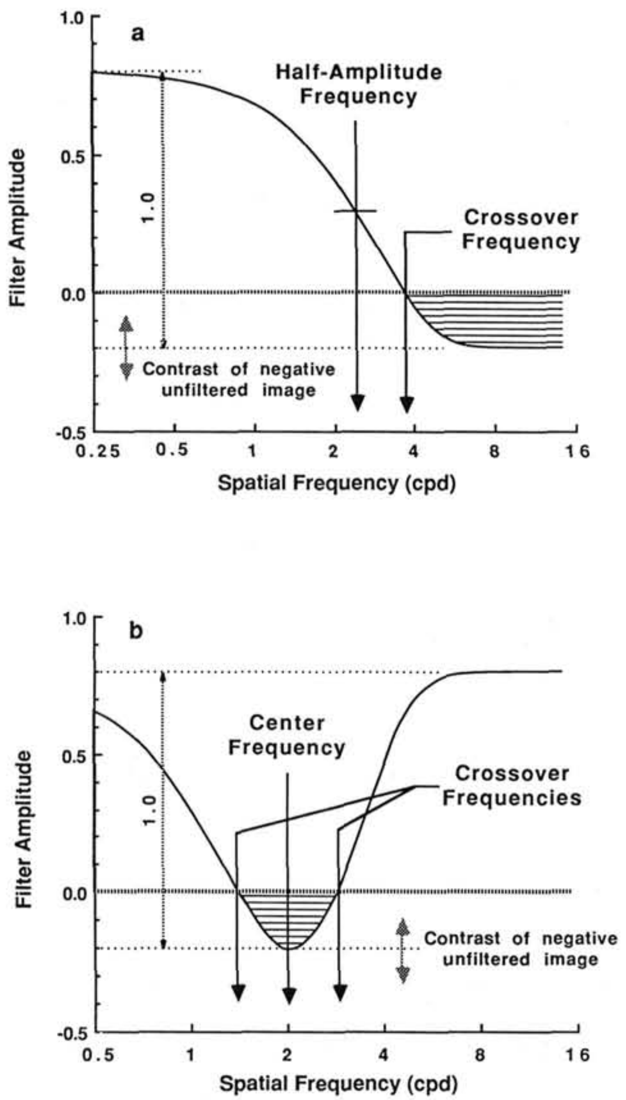


Figure 19. Panel a: The spatial frequency spectrum of the positive, low-pass filter at 100% amplitude combined with the negative, all-pass filter at 20% amplitude. (The frequency spectrum of the combined image is given by the product of the spectrum of the original image and the spectrum of this combined filter. Frequency is shown in cycles per degree [cpd]. The relative strength measure for the low-pass image is given by amplitude of the negative, all-pass filter [i.e., the contrast of the negative, unfiltered image] for which three-dimensional shape from shadows just disappears in the combined image. The crossover frequency shown on the graph is the point at which the spectrum changes from positive to negative. The crossover frequency moves to lower frequencies as the contrast of the negative, unfiltered image is increased even though the half-amplitude frequency remains fixed.) Panel b: The spatial frequency spectrum of the positive, notch filter at 100% amplitude combined with the negative, all-pass filter at 20% amplitude. (These components combine to produce a center range of negative contrast flanked by upper and lower ranges of positive contrast that are separated by about one octave. The crossover frequencies indicated on the graph are the points at which the spectrum changes from positive to negative. The crossover frequencies move fairly symmetrically away from the filter center frequency as the contrast of the negative unfiltered image is increased.)

from the original images after equalizing their amplitudes at center frequency.

A negative version of the unfiltered, original image was then added to the filtered image. The observer adjusted the contrast of the negative image until the surface relief due to shadows was just eliminated. The contrast of the negative image at this null point was then taken as the relative strength of the filtered components. The mean luminance of the display was approximately 40 cd/m^2.

Four readings were taken for each combination by 2 observers, PC and LM.

Results

The relative strengths for the low- and high-pass images are shown in Figure 20a as a function of the half-amplitude frequency of the filter. Low-pass images produced an impression of surface relief as long as their half-amplitude point fell above about 1.0 cpd (Figure 20a). Below this point, the relative strength fell below zero, and the images were too blurred to identify image features. This low frequency cut-off is most likely a function of the image size itself: Larger images, equally blurred, would regain visibility of features and surface relief.

High-pass images (Figure 20a) similarly produced impressions of 3-D organization as long as their half-amplitude frequency fell below about 4.0 cpd for Observer LM and 8 cpd for Observer PC. An appropriate contrast along the edge alone is therefore sufficient to support shape from shadows. There is no difference between the interior luminance of the shadow and nonshadow areas in high-pass filtered images such as Figure 17b. The illusory perception of a difference in the brightness between the two areas is an example of the Cornsweet-O'Brien illusion—brightness induction due to luminance edge transients. Above 4.0 cpd for Observer LM and 8 cpd for Observer PC, the relative strength of the information in the high-pass filtered image fell below zero.

The data for low- and high-pass images therefore show that images containing only spatial frequencies below about 1.5 cpd or only above about 6.0 cpd did not support 3-D shape from shadows. On the other hand, Figure 20a shows that the relative strength of the low-pass images continued to increase as components above 4.0 cpd were added, and that of the high-pass images also increased when components below 1.5 cpd were added. Do these results imply that frequencies outside the 1.5–6.0 cpd range contributed to the shadow interpretation process? An analysis of the crossover frequencies of the spectra of these images shows (Figure 21), in fact, that this was not the case.

Increasing relative strength means that the observers were able to raise the contrast of the negative, unfiltered image to even higher levels before losing the 3-D organization of the combined image, and one effect of increasing the contrast of the negative component is to shift the crossover frequency (Figure 19). When the half-amplitude frequency of the low-pass filter was at or below about 1.5 cpd, the measured relative strength was at or below zero, indicating that there was no negative image information at any place in the spectrum. When the half-amplitude frequency was increased above 1.5 cpd, observers were able to add in the negative image and still recognize the 3-D structure. The highest spatial frequencies

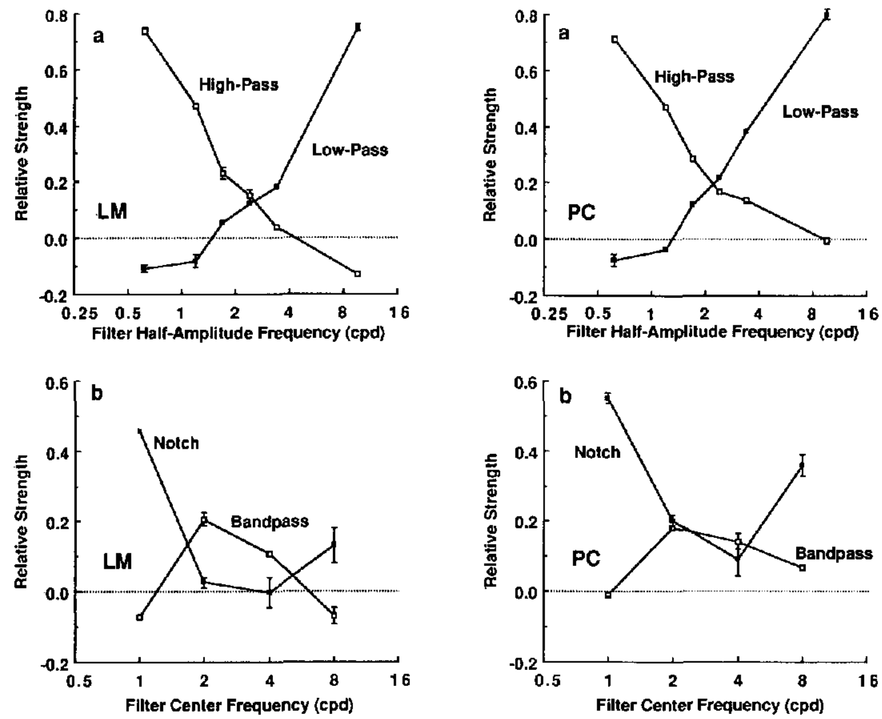


Figure 20. Panel a: Relative strength of high- and low-pass components of the image of the woman's face as a function of the half-amplitude frequency of the filter. Panel b: Relative strength of notch and bandpass components of the image of the woman's face as a function of the center frequency of the filter. (Observers LM and PC. Vertical bars show standard errors [±1 *SE*] wherever they extended beyond the data symbols. cpd = cycles per degree.)

are the first to become negative, and as the contrast of the negative, unfiltered image is increased, the crossover frequency moves to progressively lower frequencies. The results showed that the observers increased the contrast of negative image until the crossover point moved down to the upper limit of the critical region, 4.0–6.0 cpd (Figure 21, upper curves). Similarly, as the half-amplitude frequency of the high-pass filter was decreased below about 6 cpd, observers increased the contrast of negative image until the crossover point moved up to the lower limit of the critical region, 1.0–1.5 cpd (Figure 21, lower curves). These results suggest that the 3-D organization of the figure was visible as long as there were no negative image components within the 1.5–6.0 cpd range.

The results for the notch and bandpass images (Figure 20b) also show that the central range of frequencies from 1.5 to 6.0 cpd is more important in supporting the perception of the shadow image. A bandpass of two octaves for the filtered image appears to have been sufficient for the visibility of the 3-D organization. This result is consistent with that of Hayes, Morrone, and Burr (1986), who showed that a 1.5-octave bandwidth filter was sufficient for recognition of faces presented, in their case, as gray-scale images.

The results for notch-filtered images differ qualitatively between the two observers. For LM, the relative strength of the 2.0 and 4.0 cpd notch-filtered images was essentially zero, indicating that no negative contrast could be tolerated within the critical range. For PC, the notch-filtered image at 2.0 cpd could tolerate up to 20% contrast of the negative image and that at 4.0 cpd up to almost 10%. Negative components are present within the critical range for both images. If Observer PC has a slightly broader critical range, it may be that he receives sufficient positive contrast within the range to allow some negative contrast. The fact that Observer LM could tolerate no negative contrast within the critical band suggests that the total positive signal within the band must outweigh the negative by a substantial margin and that given the gradual filter functions that we used, that margin could not be reached for LM if any negative signal was within the critical range.

Our results, therefore, suggest a very simple model of the shadow border decision. A shadow border must have a consistent polarity of luminance contrast both from point to point along its length, as shown in the previous experiment, and across scales at each point (Figure 22). Because the operators must signal the polarity of the contrast at the border, they may be oriented operators such as those identified by Hubel and Wiesel (1968), although arrangements for signaling edge polarity are possible with nonoriented operators (Marr, 1982). We assume that operators signaling contrast polarity are uniformly distributed (logarithmically) across scales and are sampled within a contributing range of about two octaves. Operators indicating positive contrast must outweigh those signaling negative contrast. If a border meets these criteria, it is potentially, though not necessarily, a shadow border; if it

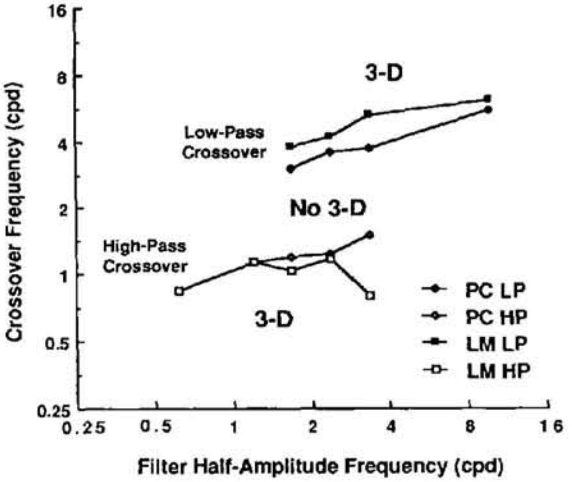


Figure 21. Crossover frequencies as a function of the half-amplitude frequency of the filter for low-pass (LP) and high-pass (HP) images. (Observers LM and PC. The crossover frequency corresponds to the point where the filter spectrum crosses from positive to negative values for the relative strengths shown in Figure 20. The three-dimensional [3-D] organization of the test stimulus was visible when the crossover frequencies were above the upper lines for the low-pass test or below the lower lines for the high-pass test. Because there is a crossover frequency in the combined filter spectrum only when the spectrum is negative over some range and because this occurs only when the relative strength setting is greater than zero, several filter spectra do not have crossover frequencies. Specifically, relative strength settings were at or below zero for low-pass images with half-amplitude frequencies below 2.0 cycles per degree [cpd] and for high-pass images with half-amplitude frequencies above 4.0 cpd; as a result, there are no crossover frequencies plotted for these filtered images.)

fails to meet these criteria, it is not acceptable as a shadow border.

The relatively narrow range of spatial frequencies that contribute to the shape-from-shadows process may ease the problem of locating curved edges. As mentioned previously, operators at different scales place a curved edge at different positions (Witkin, 1983). Because the results here show that a frequency range of only about two octaves contributes in asserting shadow borders in the figure that we tested, no scale-specific correction for curved borders may be necessary. In fact, although the range of two octaves could include several operators at different scales within that range, it may include as few as one. Unfortunately, our bandpass and notch images have too broad a bandwidth (also two octaves) to accurately assess the number of operators within the critical range. We assume that there is more than one.

The location of the critical frequency range must be strongly influenced by the scale of border detail necessary for the interpretation of the shadow image. Tightly curved contours will be completely missed by large, low-spatial-frequency operators. Although the curvature of essential contours determines the lower bound of the spatial frequency of useful operators, there is no equivalent upper limit imposed by the image structure. The fact that we measured an upper limit that is only two octaves above the lower limit indicates that processing considerations may be setting the upper limit.

If we change the image size, the change in the scale of border detail should therefore produce a change in the contributing range we measure for the bandpass image. We informally verified this by displaying a bandpass image (center frequency 4.0 cpd) that produced a recognizable shadow figure at the experimental viewing distance (1.93 m) and then by observing the image over a wide range of viewing distances. The figure retained its shadow interpretation from 0.2 m to 16.0 m (our maximum possible viewing distance). This represents a range of center frequencies extending from 0.4 cpd to 32.0 cpd. For the stimulus at its original size, only the range of center frequencies from about 1.5 to 6.0 cpd produced recognizable 3-D interpretations (Figure 20b); therefore, the central range of frequencies that contribute most effectively appears to be stimulus dependent. Note in particular that the higher spatial frequencies (above 6.0 cpd) that did not contribute to shape-from-shadows for the original image size do contribute for the smaller images, indicating again that the upper limit of the contributing band may be set by processing considerations such as optimizing cooperation across scales. On the other hand, a bandpass image that did not produce depth from shadows (1.0 cpd center frequency) remained ineffective up to about 8.0 m distance but produced a recognizable shadow figure at distances beyond that. The critical range may therefore be a function of both stimulus dependent factors and absolute spatial frequency factors.

Hayes et al. (1986) also report that the important dimension for recognition is stimulus dependent: cycles per object and not cycles per degree. The best performance in their experiment occurred for bandpass images with a center frequency of about 20 cycles per face width. Rolls, Baylis, and Hasselmo

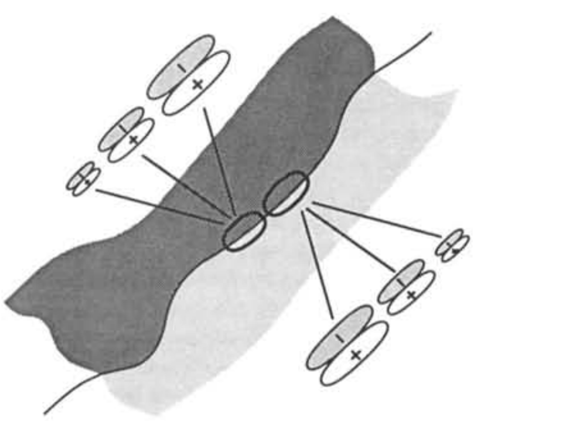


Figure 22. The border operator for signaling consistent contrast polarity at a border. (Oriented operators across a narrow range of scales must signal the same contrast polarity at a point on the border, and the polarity must remain the same from point to point along the border. There is a certain tolerance for inconsistent polarity in order to deal with chance reversals resulting from occasional alignments of texture element borders and the shadow border. The range of scales contributing to the assertion of the border depends on the scale of the border features being asserted.)

(1987) report that neurons in monkey superior temporal sulcus that are selective for face stimuli respond optimally to bandpass images with 8–16 cycles per face width. The optimal bandpass range for stimuli in our experiment was about 18 cycles per face width. (The woman's face covered about 6° of the 8° stimulus width, and the optimum center frequency for the bandpass images fell somewhere between 2.0 and 4.0 cpd).

General Discussion

Pathways

Many authors have proposed that visual information is broken down into multiple representations (Allman, Baker, Newsome, and Petersen, 1981; van Essen et al., 1981; Zeki, 1978) and that there is cooperation between these representations to determine a consistent interpretation of form (Barrow & Tenenbaum, 1978; Marr, 1982; Treisman, 1977; Treisman & Gelade, 1980). This apparently does not occur in the analysis of shape from shadows, because only luminance appeared to support the perception of our shadow figures. Shape descriptions provided by other attributes—whether color, texture, motion, or depth—were unsuccessful. Apparently, shape alone is not sufficient to establish a shadow region: The region must also be darker than its surround.

It could be argued that the loss of the shadow figures for the representations without luminance was due simply to the lower resolution and lower effective contrast of these representations. This is not the case, however. Our data showed that the shadow figures became recognizable in untextured images at about 3%–5% figure (luminance) contrast. At this point, the component shapes of the image were only just visible themselves. These same shapes when defined by color, texture, binocular disparity, or movement did not generate recognizable shadow figures even though the 2-D shapes of the components were clearly visible. The experiment on spatial frequency showed that for the luminance-defined image of the woman's face, a two-octave bandwidth and a center frequency as low as 1.5 cpd were sufficient for the perception of the shadow figure. All of the stimulus representations that we used are capable of producing resolvable images within this band. The spatial resolution for color extends to 11 or 12 cpd (Mullen, 1985), and for stereo-defined and motion-defined images up to 3.0 cpd (Nakayama & Tyler, 1981; Tyler, 1974). There are no comparable data for texture-defined images, but our impression was that they support a resolution similar to that for color-defined images. We conclude that these representations have the contrast and resolution capabilities necessary for supporting shape from shadows in our stimuli but did not do so because of the absence of the appropriate luminance contrast that is necessary for the perception of shadows. As an aside, it seems likely that the difficulty in recognizing negatives of faces (Galper, 1970; Galper & Hochberg, 1971; Hayes et al., 1986; Laughery, Alexander, & Lane 1971; Luria & Strauss, 1978; Phillips, 1972; also see Figures 4j, 4k, and 20a in this article) is due to the inappropriate luminance contrast for the shadows and shading in the face. The reverse contrast will produce inappropriate depth interpretations that disrupt the surface structure of the face.

It might seem self-evident that shadows would require luminance information to be properly interpreted: A real shadow is always darker than the adjacent nonshadow region. On the other hand, real rainbows are always colored, and yet they can be identified in black and white images simply by their shape. Why, then, is depth from shadows not perceived in images defined only by color, for example, when all the essential shape information is present? It may be that shadow analysis is part of the specialized luminance analysis just as seeing colors is part of the specialized color analysis. However, it seems unlikely that an early level of the visual system such as the luminance pathway could be solely responsible for the depth and surface inferences involved in interpreting shadows. Higher level analyses must be generating these inferences, and for reasons that are not clear, perhaps simplicity or evolutionary precedence, they appear to access only luminance information. By ignoring shape information in other pathways, the visual system gives up opportunities to reject areas as shadows because of impossible colors or inappropriate depths, motions, or textures. This is what our data showed as observers saw depth in shadow images having appropriate luminance patterns even when they violated the color, depth, motion, and texture contraints of natural shadows.

Constraints

Natural constraints have been useful for solving many visual problems (Poggio, Torre, & Koch, 1985). Disparity constraints have been used in computing stereo images (Marr & Poggio, 1976; Mayhew & Frisby, 1981), smoothness constraints to resolve ambiguous contour motion (Hildreth, 1983) and orientation fields (Zucker, 1985), and intersection constraints to solve for object and shadow surfaces (Waltz, 1975). In the case of shadows, however, constraints other than brightness appeared to play no role. Violations of constraints involving color, texture, binocular disparity, or motion had no effect on the interpretation of the shadow regions. They did influence the overall organization of the figure, however, leading to hypotheses of transparency in the case of inconsistent depth, for example. Other studies have shown that shading can override binocular disparity cues (Bülthoff & Mallot, 1988). Yellott and Kaiwa (1979) and Georgeson (1979) have shown that even with binocular viewing, an inside-out face (a mold of a face) looks right-side-out as long as shading is present. If the mold is presented solely as a random-dot stereogram with no shading, it is seen inside out.

It is possible that shadow cues were able to overrule the other cues to surface organization in our stimuli because of the familiarity or simplicity of the figures that we used. Stimuli that are more difficult to interpret to begin with, such as many of the Leeper (1935) and Mooney (1957) figures that involve both shadowed and fragmented images, are less robust in the presence of contradictory motion and binocular disparity cues. In addition, there may be large individual differences in the ability to derive shadow interpretations from luminance patterns. Even though luminance cues to shadows may not always take precedence over other cues that we have examined, the other cues appear to dominate only when the luminance information on its own is ambiguous.

Border Operators

Clearly, the one constraint that did influence the perception of shadows was that the shadow area must be darker than its surround. In particular, the contrast polarity along the border must be consistent within a range of scales at each point, as well as from point to point along the border (Figure 22).

Our data suggest the following process for identifying acceptable shadow borders. We assume, as indicated by physiological results (Hubel & Wiesel, 1968; Maffei & Fiorentini, 1977), that oriented operators are available to signal the presence of a border at a range of preferred spatial scales. For a given image location to form part of a potential shadow border, the operators signaling one polarity of contrast at that point must exceed those signaling the opposite direction by a substantial margin. The range of scales covered by the operators appears to be fairly narrow, about two octaves, and the center of the range depends on the scale of features in the stimulus. From that point, the border continues in both directions as long as the contrast polarity at subsequent points is consistent with that of the first. Our experiments gave no indication of the spatial extent of consistent polarity that is necessary to support a shadow interpretation, although we propose that it is asserted in a piecewise manner rather than required over the entire shadow border.

Identifying Shadows

So far, we have discovered only one criterion for accepting an area as a shadow: It must be consistently darker than its surround all along its border. If this were the only criterion, then all appropriately dark regions would be interpreted as shadows. However, a quick glance around any scene shows that not everything that is darker than its surround is a shadow. Before discussing other possible reasons for rejecting regions as shadows, we must make clear the distinction between extracting shape from shadows that we have studied in our experiments and judging whether or not a dark area is a shadow.

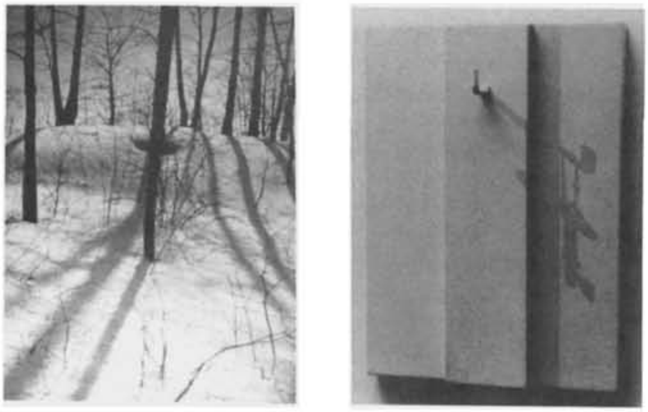
In our figures, observers were asked whether or not they could recognize the shape defined by shadows; they were not asked whether the dark areas looked like shadows. The figures had two possible organizations, including one in which shape from shadows was recognizable. Under some conditions, the shape was clearly recognizable even though the shadow area looked unnatural; for example, the color in the shadow area might look too bright for a shadow (Cavanagh et al., 1987). It is possible that two processes are involved in analyzing shadows, one that uses dark regions to extract shape from shadows and a second that judges the acceptability of a shadow region as a shadow. Even when a region may not be acceptable as a shadow in a scene for reasons of texture or color, it may nevertheless provide the darkness and shape information necessary for the perception of the image as a representation (picture) of a 3-D, shadowed object. That is, the shape-from-shadow process generates a 3-D interpretation, but the surface qualities of the shadow region lead it to be rejected as a scene shadow. The figure is then seen not as a shadowed object in the scene but as a picture of such an object.

In a real scene, there may be many dark areas that are good candidates for shadows. A road surface is a good example because we can often see shadows, wetting, stains, and cracks on the same surface. Changes of pigment (paint, tire skid marks), changes of reflectance (oil or water wetting the surface), and surface cracks all conform to the surface on which they lie, do not look like objects, and can be darker than the surround. They do not look like shadows, however, because they are the wrong color or because their internal texture differs from that of the surround. They may even appear too dark, as if the visual system independently evaluates the ambient light and judges the acceptable level of darkness for a shadow. These discriminations show a sensitivity to factors that appeared to have no influence on shape-from-shadows in our experiments. One important difference between the natural scene and our images is that our stimuli had only two levels in them, and so there was no basis for independently judging what color or darkness the shadows should have had. If shadow regions produce no recognizable shape, then evidence that they are not shadows (incorrect color or darkness or absence of a casting object) may tip their interpretation toward alternatives such as stains or wetting. Finally, it is clear that the ambiguity in interpreting dark areas is never completely resolved because many shadow areas can be taken as material changes on first glance, and conversely, material changes are often misinterpreted as shadows. The last case is, of course, the basis for pictures where the misinterpretation is intentional.

Although the surface qualities of a dark region may influence its acceptability as a cast shadow, the lack of correspondence between the shadow and the casting object does not seem to have much effect. The dark regions in Figure 23a are accepted as shadows without comparing them to the trees in shape or even in number. Sometimes, the object casting the shadow may not be present in the visual field or may not exist at all (Figures 23a, 23b).

To summarize, there may be a hierarchy of interpretations available for a dark area in an image. We suggest that the shadow interpretation processes are strongly model driven so that the interpretations with the highest priority are those supporting familiar object shapes. The dark areas in an image may be self-shadows within a familiar shape (e.g., under the woman's nose in Figure 4a), or they may be regions of low reflectance of a familiar object (e.g., the woman's eyebrows in Figure 4a). If a dark region supports neither interpretation, it may be seen as a cast shadow if its surface qualities do not differ from those of its surround (e.g., the tree shadows in Figure 23a); otherwise, it is seen as an amorphous material change, a stain, or wetting. If a dark region supports a shape due to shadows but does not have the surface qualities required for a cast shadow in the scene, the shape from shadows is still seen, but the image is interpreted as a picture.

This suggests that a cast shadow is one of the last choices the visual system considers when attempting to classify a dark



a b

Figure 23. Panel a: Tree shadows. (It is unlikely that any comparison is made between the trees and their shadows, either in shape or in number.) Panel b: This key shadow looks quite natural even though there is no key. (Photograph of construction "No. 227" by Jiro Takamatsu. Adapted from *Sensation and Perception* [p. 156] by E. B. Goldstein, 1980, Belmont, CA: Wadsworth. Copyright 1980 by Wadsworth. Adapted by permission of the author and the publisher.)

area. The human visual system is not well suited to the solution of the optics problem posed by cast shadows and so apparently avoids it as much as possible. Clearly, this approach can work only if the visual system excels at identifying objects, so that the number of regions that have to be considered as possible cast shadows is small. Computer vision techniques, on the other hand, are very ineffective when it comes to identifying objects because they cannot access the immense number of stored object descriptions that biological systems can. A hierarchical classification scheme like the one we suggest for human vision is therefore ill advised for computer vision (Binford, 1984; Shafer, 1984). However, computer algorithms can be very good at solving optics constraints for shadows, such as permissible shadow colors (Gershon, Jepson, & Tsotsos, 1986; Rubin & Richards, 1982; Shafer, 1984) and depth, motion, and texture contraints. These constraint criteria could compensate to some extent for the less powerful model-driven processes available in computer vision.

Shape From Shadows

How intelligent is the process that determines the 3-D shape of an object from the 2-D shadow shapes? We have already noted that it is not possible to recover the illuminant, the object, and the surface from the observed shadow borders alone. One possibility is that nothing needs to be recovered. Once shadow areas are identified, they are taken as occluded areas of the object surface, and what remains is sufficient to support the perception of the object. Thus shadow areas, if misinterpreted as material changes, can disrupt the analysis of the image, but once they are correctly identified, they do not need to contribute anything further. An analogous situation can be seen in a demonstration from Bregman (1981). The shapes of Figure 24a appear as jumbled letter pieces. Once an overlying surface is made visible (Figure 24b), it is possible to distinguish between object contours and occluding contours, and the shapes become recognizable. Most important, separated bits of familiar patterns, capital letter *B*s, can be grouped across the occluding area. This grouping was less evident in Figure 24a, where we had little basis for assuming an occluding surface. The perception of the shadow figures may follow a similar process. When a border is identified as a potential shadow border, hypothesized object details can be completed within the shadow area, and regions can be patched together across the shadow area. If the border is signaled as a material border that does not conform to the polarity constraints for a shadow, then no hidden region can be hypothesized, and no grouping occurs across the area.

When there are actual object contours visible due to regions of low reflectance, as is true for the woman's face of Figure 4a and *B*s of Figure 24, it may be sufficient simply to ignore the shadows as occluded areas and rely on the subset of visible contours to identify the object. That we can do this indicates that the matching process can identify candidate objects from a subset of contours in the presence of unrelated contours (the shadow contours) even though there are no explicit cues distinguishing object contours from shadow contours.

When there are no areas of low reflectance producing object contours, the problem is more difficult, and we will consider the three different types of shadow contours (extremal, terminator, and cast) separately. Outer contours of the object are visible against the background if one is lit and the other is not, and these outer, or extremal, contours are highly informative. The left-hand, top, and bottom contours of the cup in Figure 25a are extremal contours and are marked with *E* in Figure 25b. These contours follow convexities on the object surface that are normal to the observer's line of gaze.

A second class of shadow contours—terminators—follows convexities on the surface that are everywhere normal to the direction of the illuminant. (A familiar terminator contour is the division between the dark side and the light side of the moon.) In the cup of Figure 25a, the right-hand vertical

a b

Figure 24. Fragmented shapes that appear unconnected in Panel a but emerge as capital letter *B*s partly occluded in Panel b when an occluding surface is made explicit. (Adapted from "Ask the 'What for' question in auditory perception" by A. S. Bregman, 1981. In M. Kubovy and J. Pomerantz [Eds.], *Perceptual Organization* [pp. 106–107]. Hillsdale, NJ: Erlbaum. Adapted by permission. This identification of shapes from partial contours and the grouping of the pieces across the occluding surface is analogous to the processes involved in identifying shadow figures. In a sense, a shadow area occludes the surfaces within it, and once the shadow is identified, the surfaces on either side of it can be linked together.)

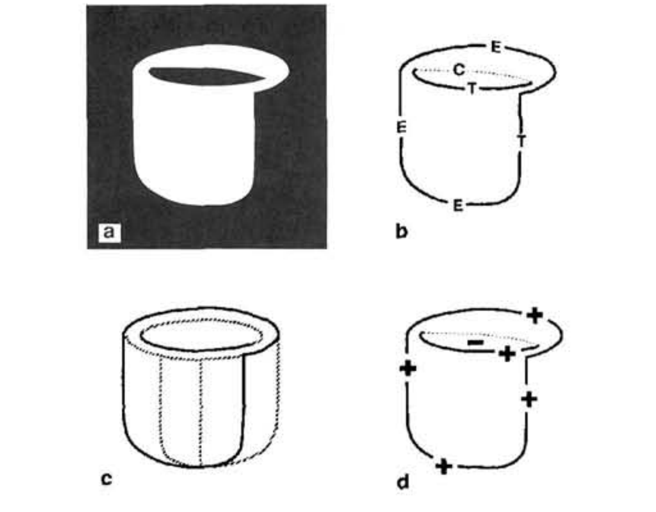


Figure 25. Panel a: A shadow figure of a thick-sided cup or crucible. Panel b: The same figure but retaining only the extremal (e), terminator (t), and cast (c) contours. Panel c: The terminator and extremal contours produced from several different light directions produce the impression of surface contours. Panel d: Extremal and terminator contours occur only at positions of surface convexity. (Enclosed shadow regions must contain a convexity.)

contour and the upper edge of the lower rim of the cup's opening are terminal contours and are marked with a *T* in Figure 25b. In the case of face stimuli such as Figures 4a and 4b, terminal contours fall principally along the ridges of the nose, eyebrows, or lips, tracing important contours of the surface structure. These are the same surface folds, eyebrows, lips, and noses that an artist would represent when making a line drawing. Not all of the terminator contours will follow such ridges of high convexity, however. In objects such as a cup, the surface of the cylindrical body of the cup has equal curvature everywhere.

If the visual system can make a match between some subset of the shadow contours and a familiar shape, it can then assume that the remaining contours are cast shadow contours, for example, the edge of the shadow cast by lip of the cup onto the cup's interior (marked with *C* in Figure 25b). The cast shadow contours are the least informative of the three contour types.

So far we have suggested that shadows contribute to the retrieval of 3-D shape in two ways: (a) A subset of shadow contours may be sufficient for object recognition—3-D shape is then retrieved from the known structure of the object; and (b) once a shadow is identified, the parts of the object's surface separated by the shadow can be linked together as a single surface.

Shadows may also provide two cues to surface curvature without having to solve the correspondence problem: the presence of a convexity at the terminator contour and a concavity within a closed shadow region. These may be used in constructing the object's 3-D organization (Figure 25d).

Does visual system use correspondence information at all? We appear to be fairly insensitive to variations in cast shadows that deviate noticeably from the actual shadow (Figure 23b). On the other hand, Berbaum, Bever, and Chung (1984) showed that cast shadows can resolve surface shading ambiguities by revealing the direction of the light source, and Yonas (1979) showed two instances where cast shadows are considered in determining the casting object's position and shape (Figure 26). The cast shadow of the sphere in Figure 26 appears to determine its position in space even though its picture position has not changed. In his second example, an ellipse in the picture represents first a flat disk and then an egg-shaped object as a function of the shape of its shadow.

Although these cast shadows may help somewhat, observers are not very accurate in using shading cues to judge object shape. Barrow and Tenenbaum (1981), for example, showed that luminance gradients on a cylindrical surface could depart substantially from natural shading without changing the perceived shape of the surface. Todd and Mingolla (1983) reported that observers made errors of up to 50% in estimating surface curvature based on shading. Mingolla and Todd (1986) showed, as well, that adding a cast shadow did not improve the judgments of surface shape. It appears that the visual system is not making very precise computations of light sources and reflectance normals (Gershon et al., 1986; Ikeuchi & Horn, 1981; Pentland, 1982; Shafer, 1984, 1985; Woodham, 1981, 1984) in either the shape-from-shading or the shape-from-shadows situation.

In fact, the contribution of shading to surface understanding may not be mediated by the computation of surface normals (Pentland, 1982; Woodham, 1981) at all but perhaps proceeds on the basis of surface contours in a manner similar to that proposed here for shadows. The shading on a surface may be interpreted as a tangent field, with each local tangent oriented orthogonally to the direction of maximum brightness change. This tangent field could act as a set of surface contours sufficient for the reconstruction of the surface relief (Stevens, 1981, 1986; Zucker, 1985).

We can summarize the contribution of shadow shape to object shape in four possible levels. We believe that the first level occurs in all shadowed images and is sufficient for

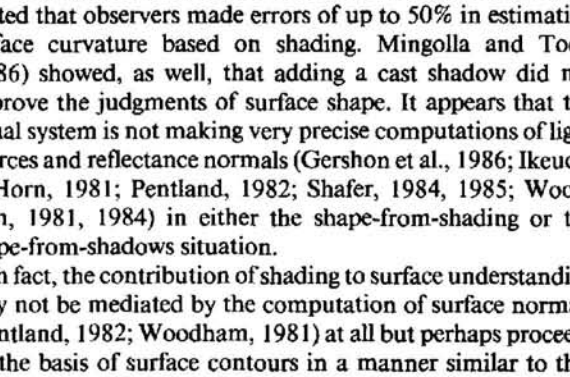


Figure 26. Cast shadows can influence the apparent position of an object and its shape. (Adapted from "Attached and cast shadows" by A. Yonas, 1979. In C. F. Nodine and D. F. Fisher [Eds.], *Perception and Pictorial Representation* [pp. 104 and 108]. New York: Praeger. Copyright 1979 by Praeger Publishers. Adapted and reprinted by permission. It is not clear whether the visual system actually solves the correspondence between the object and its shadow or if it has only a few simple rules for particular cases.)

determining shape from shadows. The subsequent levels may contribute in particular instances.

1. Shadows provide a subset of object contours and a hypothesis of occluded object regions.

2. Terminator and extremal contours signal surface convexities; enclosed shadow regions signal concavities.

3. The impression of the surface relief on which the shadow falls is influenced by the cast shadow shape in special cases.

4. The impression of the position and shape of the object is influenced by its cast shadow position in special cases.

Conclusions

We have found that the visual system verifies only the luminance along the border of a region to determine whethei it is an acceptable shadow region for determining shape from shadows. The operator identifying the shadow border requires consistent polarity across a range of scales at each border point, as well as from point to point along the border. The critical range is fairly narrow and is influenced by the scale of the image features.

Our experiments have examined the low-level criteria that the visual system uses to identify potential shadow areas when retrieving shape from shadows. The evidence that only a single criterion was used points to a very particular, high-level approach to shadow interpretation—one that emphasizes model-based analysis over the analysis of illumination constraints in the scene. Using this observation as a guideline, we have proposed that shadow shape is used in three ways in retrieving 3-D object shape. (a) A subset of the shadow contours is matched against familiar prototypes to identify known shapes and to recover their 3-D organization from stored information. (b) Shadows may also contribute to the analysis of surface curvature in signaling convexity along the terminator and extremal contours and concavity within closed shadow regions. (c) In some instances, cast shadow borders may provide information about surface relief and object shape.

References

Allman, J. M., Baker, J. F., Newsome, W. T., & Petersen, S. E. (1981). Visual topography and function: Cortical visual areas in the owl monkey. In C. N. Woolsey (Ed.), *Cortical sensory organization: Vol 2. Multiple visual areas* (pp. 171–186). Clifton, NJ: Humana Press.

Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research, 25*, 1411–1430.

Asada, H., & Brady, M. (1985). The curvature primal sketch. *IEEE transactions on pattern analysis and machine intelligence*, PAMI-8, 2–14.

Barrow, H. G., & Tenenbaum, J. M. (1978). Recovering intrinsic scene characteristics from images. In A. Hanson & E. Riseman (Eds.), *Computer vision systems* (pp. 3–26). New York: Academic Press.

Barrow, H. G., & Tenenbaum, J. M. (1981). Computational vision. *Proceedings of the IEEE, 69*, 572–595.

Beck, J. (1972). *Surface color perception.* Ithaca, NY: Cornell University Press.

Berbaum, K., Bever, T., & Chung, C. S. (1984). Extending the perception of shape from known to unknown shading. *Perception, 13*, 5–20.

Berbaum, K., Tharp, D., & Mroczek, K. (1983). Depth perception of surfaces in pictures: Looking for conventions of depiction in Pandora's box. *Perception, 12*, 5–20.

Binford, T. O. (1984). Survey of model-based image analysis systems. *International Journal of Robotics Research, 1*, 18–64.

Braddick, O. (1974). A short-range process in apparent motion. *Vision Research, 14*, 519–528.

Bregman, A. S. (1981). Ask the "What for" question in auditory perception. In M. Kubovy & J. Pomerantz (Eds.), *Perceptual organization* (pp. 99–118). Hillsdale, NJ: Erlbaum.

Bülthoff, H. H., & Mallot, H. A. (1988). Integration of depth modules: Stereo and shading. *Journal of the Optical Society of America A, 5*, 1749–1758.

Cavanagh, P. (1985). Depth and surface inferences in line drawings. *Journal of the Optical Society of America A, 2*, P51.

Cavanagh, P. (1987). Reconstructing the third dimension: Interactions between color, texture, motion, binocular disparity and shape. *Computer Vision, Graphics and Image Processing, 37*, 171–195.

Cavanagh, P., Anstis, S. M., & MacLeod, D. I. A. (1987). Equiluminance: Spatial and temporal factors and the contribution of the B-cones. *Journal of the Optical Society of America A, 4*, 1428–1438.

Cavanagh, P., Shioiri, S., & MacLeod, D. I. A. (1987). Is the achromatic form pathway based on luminance or brightness? *Investigative Ophthalmology and Visual Science Supplement, 28*, 362.

De Valois, K. K., & Switkes, E. (1983). Simultaneous masking interactions between chromatic and luminance gratings. *Journal of the Optical Society of America, 73*, 11–18.

Galper, R. E. (1970). Recognition of faces in photographic negative. *Psychonomic Science, 19*, 207–208.

Galper, R. E., & Hochberg, J. (1971). Recognition memory for photographs of faces. *American Journal of Psychology, 84*, 351–354.

Georgeson, M. A. (1979). Random-dot stereograms of real objects: Observations on stereo faces and moulds. *Perception, 8*, 585–588.

Gershon, R., Jepson, A. D., & Tsotsos, J. K. (1986). Ambient illumination and the determination of material changes. *Journal of the Optical Society of America A, 3*, 1700–1708.

Gilchrist, A., Delman, S., & Jacobsen, A. (1983). The classification and integration of edges as critical to the perception of reflectance and illumination. *Perception & Psychophysics, 33*, 425–436.

Goldstein, E. B. (1980). *Sensation and perception.* Belmont, CA: Wadsworth.

Gregory, R. L. (1977). Vision with isoluminant colour contrast: 1. A projection technique and observations. *Perception, 6*, 113–119.

Gregory, R. L. (1979). Stereo vision and isoluminance. *Proceedings of the Royal Society of London B, 204*, 467–476.

Hayes, T., Morrone, M. C., & Burr, D. C. (1986). Recognition of positive and negative bandpass-filtered images. *Perception, 15*, 595–602.

Hildreth, E. C. (1983). *The measurement of visual motion.* Cambridge, MA: MIT Press.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology, 195*, 215–243.

Ikeuchi, K., & Horn, B. K. P. (1981). Numerical shape from shading and occluding boundaries. *Artificial Intelligence, 15*, 141–184.

Julesz, B. (1971). *Foundations of cyclopean perception.* Chicago: University of Chicago Press.

Laughery, K. R., Alexander, J. F., & Lane, A. B. (1971). Recognition of human faces: Effects of target exposure time, target position, and type of photograph. *Journal of Applied Psychology, 55*, 477–485.

Leeper, R. (1935). A study of a neglected portion of the field of learning. *Journal of Genetic Psychology, 46,* 41–75.

Lezak, M. D. (1976). *Neuropsychological assessment.* New York: Oxford University Press.

Luria, S. M., & Strauss, M. S. (1978). Comparison of eye movements over faces in photographic positives and negatives. *Perception, 7,* 349–358.

Maffei, L., & Fiorentini, A. (1977). Spatial frequency rows in the striate visual cortex. *Vision Research, 17,* 257–264.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* San Francisco: Freeman.

Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B, 207,* 187–217.

Marr, D., & Poggio, T. (1976). A cooperative computation of stereo disparity. *Science, 194,* 283–287.

Maunsell, J. H. R., & Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neurosciences, 10,* 363–401.

Mayhew, J. E. W., & Frisby, J. P. (1981). Psychophysical and computational studies towards a theory of human stereopsis. *Artificial Intelligence, 17,* 349–387.

Mingolla, E., & Todd, J. T. (1986). Perception of solid shape from shading. *Biological Cybernetics, 53,* 137–151.

Mooney, C. M. (1957). Age in the development of closure ability in children. *Canadian Journal of Psychology, 11,* 219–226.

Mullen, K. T. (1985). The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. *Journal of Physiology, 359,* 381–400.

Nakayama, K., & Tyler, C. W. (1981). Psychophysical isolation of movement sensitivity by removal of familiar position cues. *Vision Research, 21,* 427–433.

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

Phillips, R. J. (1972). Why are faces hard to recognize in photographic negative? *Perception & Psychophysics, 12,* 425–426.

Poggio, T., Torre, V., & Koch, C. (1985). Computational vision and regularization theory. *Nature, 317,* 314–319.

Richter, J. P. (1970). *The notebooks of Leonardo da Vinci.* New York: Dover.

Rolls, E. T., Baylis, G. C., & Hasselmo, M. E. (1987). The responses of neurons in the cortex in the superior temporal sulcus of the monkey to bandpass spatial frequency filtered faces. *Vision Research, 27,* 311–326.

Rubin, J. M., & Richards, W. A. (1982). Color vision and image intensities: When are changes material? *Biological Cybernetics, 45,* 215–216.

Shafer, S. A. (1984). *Using color to separate reflection components* (Tech. Rep. TR-136). Rochester, NY: Department of Computer Science, University of Rochester.

Shafer, S. A. (1985). *Shadows and silhouettes in computer vision.* Boston: Kluwer Academic Publishers.

Stevens, K. A. (1981). The visual interpretation of surface contours. *Artificial Intelligence, 17,* 47–73.

Stevens, K. A. (1986). Inferring shape from configurations of contours across surfaces. In A. P. Pentland (Ed.), *From pixels to predicates: Recent advances in computer vision* (pp. 93–110). Norwood, NJ: Ablex.

Tieger, T., & Ganz, G. (1979). Recognition of faces in the presence of two-dimensional sinusoidal masks. *Perception & Psychophysics, 26,* 163–167.

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.

Treisman, A. (1977). Focused attention in the perception and retrieval of multidimensional stimuli. *Perception & Psychophysics, 22,* 1–11.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136.

Tyler, C. W. (1974). Depth perception in disparity gratings. *Nature, 251,* 140–142.

Ullman, S. (1979). *The interpretation of visual motion.* Cambridge, MA: MIT Press.

van Essen, D. C. (1985). Functional organization of primate visual cortex. In A. Peters & E. G. Jones (Eds.), *Cerebral cortex: 3* (pp. 259–329). New York: Plenum.

van Essen, D. C., Maunsell, J. H. R., & Bixby, J. L. (1981). Organization of the extrastriate visual areas in the macaque monkey. In C. N. Woolsey (Ed.), *Cortical sensory organization: Vol 2. Multiple visual areas* (pp. 157–170). Clifton, NJ: Humana Press.

Waltz, D. (1975). Generating semantic descriptions from drawings of scenes with shadows. In P. H. Winston (Ed.), *The psychology of computer vision* (pp. 19–29). New York: McGraw-Hill.

Witkin, A. P. (1983). Scale-space filtering. *Proceedings of the 8th International Joint Conference on Artificial Intelligence, 2,* 1019–1022.

Woodham, R. J. (1981). Analyzing images of curved surfaces. *Artificial Intelligence, 17,* 117–140.

Woodham, R. J. (1984). Photometric method for determining shape from shading. In S. Ullman & W. Richards (Eds.), *Image understanding 1984* (pp. 97–125). Norwood, NJ: Ablex Publishing.

Yellott Jr., J. I., & Kaiwa, J. L. (1979). Depth inversion despite stereopsis: The appearance of random-dot stereograms on surfaces seen in reverse perspective. *Perception, 8,* 135–142.

Yonas, A. (1979). Attached and cast shadows. In C. F. Nodine & D. F. Fisher (Eds.), *Perception and pictorial representation* (pp. 100–109). New York: Praeger.

Yonas, A., Cleaves, W. T., & Petersen, L. (1978). Development of sensitivity to pictorial depth. *Science, 200,* 77–78.

Zeki, S. M. (1978). Functional specialization in the visual cortex of the rhesus monkey. *Nature, 274,* 423–428.

Zeki, S. M. (1980). The representation of colours in the cerebral cortex. *Nature, 284,* 412–418.

Zucker, S. W. (1985). Early orientation selection: Tangent fields and the dimensionality of their support. *Computer Vision, Graphics and Image Processing, 32,* 74–93.

Received November 20, 1987
Revision received March 7, 1988
Accepted March 8, 1988 ■