Reprinted from:

COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING **37**, 171–195 (1987)

Reconstructing the Third Dimension: Interactions between Color, Texture, Motion, Binocular Disparity, and Shape

PATRICK CAVANAGH

Département de Psychologie, Université de Montréal, Montréal, Québec, Canada H3C 3J7

Received March 25, 1986

The effectiveness of depth cues such as occlusion and shading was examined in images defined by color, texture, binocular disparity or motion. Line drawings represented in any of these modalities were able to signal shape and occlusion showing that contour occlusions are analyzed at a high level, following the reintegration of the separate representations of visual attributes such as color and motion. Subjective contours, on the other hand, could be seen only if the figures were defined by luminance differences. Figures whose depth depended on the interpretation of shadows also required luminance differences: shadow regions had to be darker than the surrounding, non-shadow regions. Shadows areas filled with colors or textures that could not occur in natural scenes were perceived as shadows as readily as real shadows. Even when shadow and non-shadow regions had different depths or had textures that moved in different directions, the depth from shading was still seen as long as there was an appropriate brightness difference. These findings indicate a variety of mechanisms analyze cues to 3-dimensional structure. Occlusion cues in line drawings appear to be analyzed by a general purpose mechanism having access to all pathways of the visual system. Subjective contours and shadows appear to depend on special purpose processes accessing only the luminance pathway. Finally, although natural constraints have proved useful in solving many visual problems, they did not play a significant role in the interpretation of the depth cues examined here.

1. INTRODUCTION

In general, three 2D views of a scene are necessary to specify its 3-dimensional organization and, for best performance, the three views should be from orthogonal directions. The human visual system, however, appears to be quite skilled at recovering 3-dimensional structure from just two eyes that have only slightly different viewpoints, and can even do quite well with a single eye. The visual system must be using some additional, non-retinal sources of information to achieve this performance. In particular, knowledge concerning the constraints of natural images, for example, the properties of light and the interactions between contours of overlapping objects, may be critical to the recovery of depth in the scene [2, 7, 11, 15, 20, 24, 32, 45, 46]. The range of cues entering into the interpretation of the image therefore includes both information sent from the two eyes and inferences that can be drawn from this information based on knowledge of real-world objects. These cues include binocular disparity, vergence, accommodation, relative motion, occlusion, shading, and the structure and size of identified, familiar objects.

How is this variety of cues integrated to produce a coherent impression of the 3-dimensional structure of the scene? As one attempt to understand how are these cues processed, we will look at situations in which cues conflict to see if natural constraints play a role in resolving the conflicts. When there is a conflict between two cues, when binocular disparity suggests one relative depth and occlusion cues another, for example, there are necessarily two possible interpretations of the image. Generally one interpretation violates some property of natural images—a constraint. One example that we will see is an interpretation that requires us to be able to see through an opaque surface. This interpretation violates a constraint concerning opacity while an alternative interpretation does not. The question is whether there is a preference for the interpretation that does not violate any constraints.

We will also look at how these cues are conveyed by the various pathways of the visual system (Fig.1). It is not known how many pathways there are nor exactly what information is processed by each, but it is clear that there are many separate representations of the visual field in the cortex [1, 43]. Zeki [49] has described several regions in the prestriate cortex that may be performing separate analyses of color, motion, binocular disparity, and orientation. Treisman [40, 41] has also proposed that there are several independent feature maps involved in pre-attentive vision. These independent representations appear to be a major aspect of the large-scale organization of image processing in the visual system and may provide a significant functional advantage. Barrow and Tenenbaum [3], in particular, have described how a set of intrinsic images, representing object properties such as reflectance, color, and motion, would be extremely useful in image understanding if they could be derived at an early level.

In order to study the roles of these different representations, or pathways, I will present images that activate only one pathway at a time (Fig. 1). There is not a complete overlap between the properties of the various representations proposed by Barrow and Tenenbaum, Zeki, and Treisman nor is there a complete overlap between any of those and the five prototypical pathways I have placed in Fig. 2. The representations that I have selected are meant to be reasonable candidates, supported by some physiological and behavioral data but open to modification.

Each of these representations, or pathways, performs some specialized function such as the extraction of object color independently of illumination [50], but each is also capable of signaling shape. Julesz [25] has shown this for stimuli defined only by binocular disparity and others have shown it for stimuli defined by relative motion [9] and by equiluminous colors [18]. We will examine whether the shape that can be signaled by each pathway can support depth cues such as occlusion and shadows.

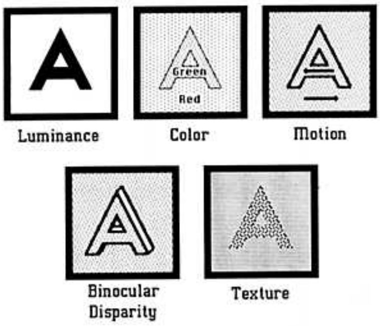

FIG. 1. Individual visual pathways can be isolated perceptually by presenting stimuli that are defined by variations along a single dimension.

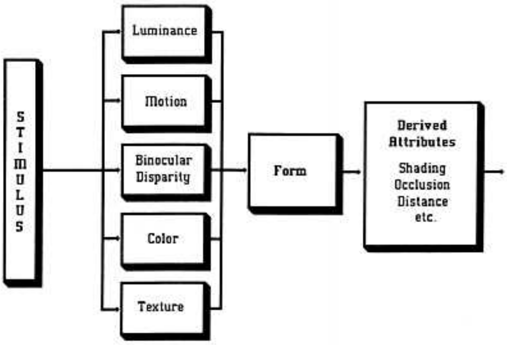

FIG. 2. The role of multiple representations in the visual system. Stimulus information is broken down into several separate representations, each specializing in a particular aspect of the image. Following these specialized analyses, stimulus shapes are determined so as to be maximally consistent with the borders and surfaces determined by the individual analyses. Finally, based on this shape information, depth and surface inferences are derived from possible occlusion and shading interactions in the image. The organization and representations shown here are meant to be reasonable candidates, supported by some physiological and behavioral data but open to modification.

We will look at three particular depth cues: contour occlusions in line drawings, surface occlusions in subjective contour figures and shadows, or illumination occlusions, in images of faces. In each case we will examine two questions. First, what pathways are involved in conveying these cues to higher order analyses and, second, do natural constraints influence the resolution of conflicts?

2. LINE DRAWINGS

A line drawing represents a 3-dimensional scene by the 2-dimensional projection of its discontinuities. Figure 3 shows an example from the caves of Lascaux, one of the oldest drawings on record. This early rendering of a French rhinoceros shows that even the very first drawings made use of contour representations; notice in particular the occlusion of one of the hind legs by the other. Several authors [2, 3, 29] have suggested that something resembling an outline or contour drawing is the underlying representational scheme used by the visual system. If this is the case, this may account for the ease with which we are able to interpret line drawings, an ability which appears to be universal and innate [26]. For example, children can generally identify line drawing by 22 to 26 weeks of age [48] and in one particular case, a child raised without exposure to pictures of any kind until age two could name line drawing of objects at the first presentation [22].

The questions that will be raised concern (1) the cues in the line drawings that allow the interpretation of depth, (2) the pathways in the visual system capable of representing those cues, and (3) the integration of information across pathways in the case of conflicts, for example, when the depth indicated by the line drawing and that by binocular disparity are in conflict.

First of all, there is a wide variety of cues to 3-dimensional structure in line drawings [5, 6, 26]. High-level clues include the known structure of an identified object. Many forms such as faces or automobiles, for example, can be identified on the basis of their 2-dimensional outlines. The stored information concerning the

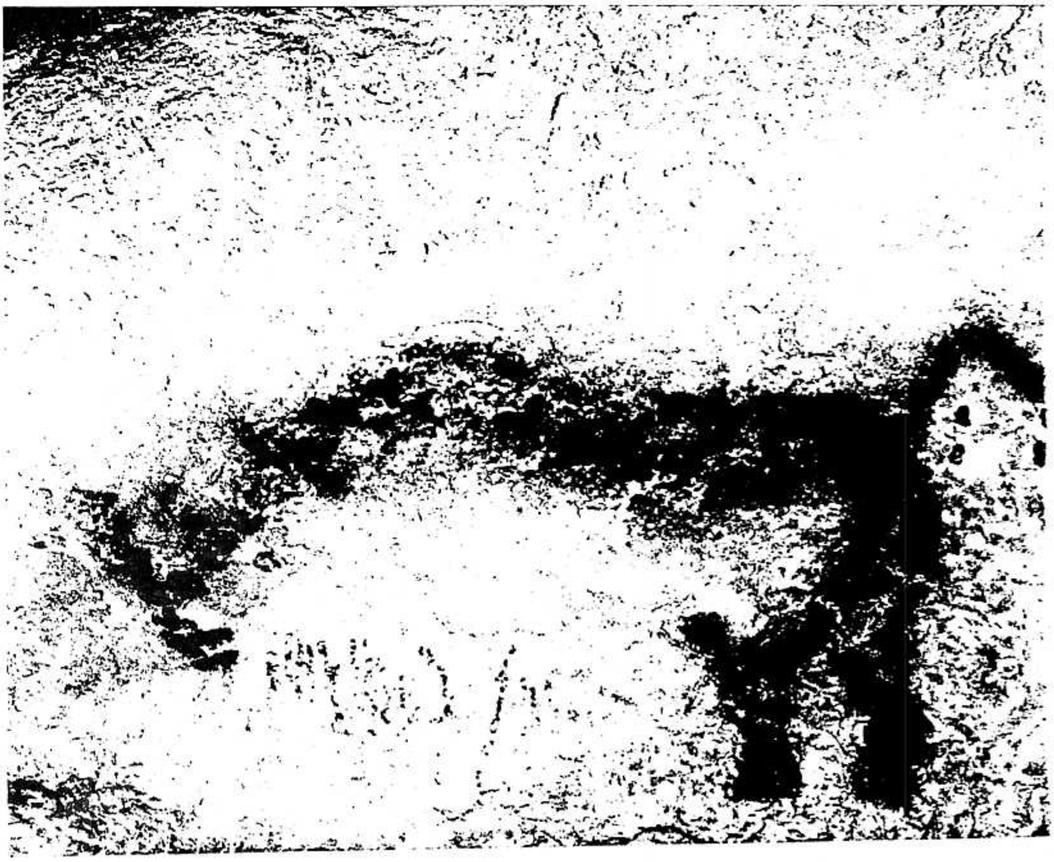

FIG. 3. An early line drawing (15,000 to 20,000 B.C.) from the caves of Lascaux, France.

surface relief of the form then completes their interpretation. Other information available once objects are identified are their size, and therefore their distance relative to other objects, and certain physical properties. For example, once a region is identified as water, specifically, nonturbulent water, its surface orientation can be inferred to be horizontal.

Low-level cues function independently of the content of the drawing and intersection cues, in particular, will be examined here. The intersections of object contours (Fig. 4) have been used to extract of 3-dimensional structure for geometric solids [11, 19, 24, 28, 45], as well as curved [42] and entirely smooth objects [27]. The intersection cues appear to be interpreted locally as can be seen when an impossible figure is examined (Fig. 5). The occlusion implied by the ***T*** intersections is not consistent with the rest of the image and yet the implied depth is still seen in the vicinity of the intersection. The figures that will be examined here will use only the ***T*** intersection, indicating occlusion and the ***X***, or matched ***T***, intersection indicating the crossing of a thin object, a wire or bar, over another contour.

The test figures are a Necker cube and a pair of overlapping, opaque sheets (Fig. 6). In the first, the ***X*** intersections indicate that the areas between the lines are

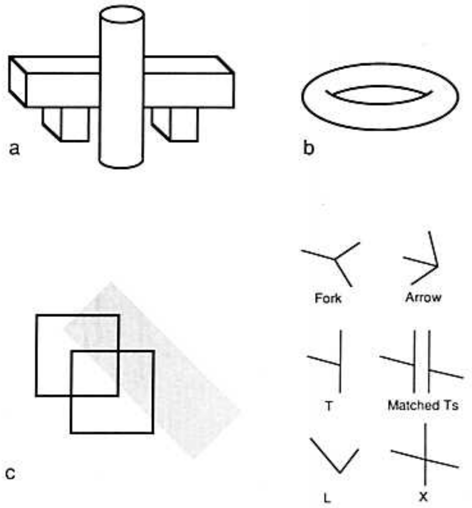

FIG. 4. (a) The intersections of the contours of geometric solids can be classified into a small number of types, some of which are shown in the bottom right-hand panel. Each intersection imposes particular constraints on the organization of the adjacent surfaces. In many cases, simultaneously solving the constraints of all the intersections of the image recovers the 3-dimensional structure of the objects [11, 19, 24, 28, 45]. (b) The ***T*** intersection is a cue to occlusion for smooth surfaced objects as well [27]. (c) The ***X*** intersection between overlying thin objects (bars or wire frames) or between overlying bars and the contours of underlying solid objects indicates that at least one of the surfaces bounded by the contours involved in the ***X*** intersections may be transparent.

transparent. We can see through these areas to the back plane. In the second, ***T*** intersections indicate that the front surface is opaque and the figure appears as two similar opaque sheets staggered one behind the other. The 3-dimensional interpretation of the figures appears to be spontaneous and natural even though binocular disparity indicates that the image is presented on a flat surface. Undoubtably it is this binocular disparity cue that leads us to interpret these as pictures and not as actual 3-dimensional objects.

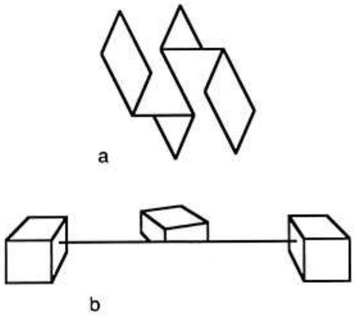

FIG. 5. The relative depth cues from contour intersections act locally, even if they are inconsistent with other relative depth cues elsewhere in the image. Occlusion is seen in both (a) the impossible folded sheets and (b) the hidden box (examples from Kennedy [26]).

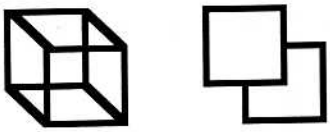

FIG. 6. The two test stimuli used in the experiments. On the left, the *X* intersections in the Necker cube give the impression of a wire frame structure. The areas bounded by the wire frame seem to be transparent; a uniform background surface can be seen through the frame as well as around it. On the right, the two *T* intersections suggest that the area bounded by the complete square is opaque and hides the corner of a similar square that is lying behind it. Perhaps by similarity to the front square, the back square also gives the impression of being opaque even though there is no evidence to support this interpretation.

The figures were presented defined by one attribute at a time: brightness, color, texture, binocular disparity, and finally motion. The observer's task was to determine whether the 3-dimensional interpretations were still available when these figures were represented by a single attribute or whether the figures appeared flat or organized in some other fashion.

2.1. Constraints

In some modes of representation, the figures present a conflict between depth cues. For example, in Fig. 7, if the contours of the two occluding sheets are presented as a random dot stereogram with the lines standing out from the page (the edges would not be visible in the stereogram) it seems unlikely that the figure could be still interpreted as opaque sheets because the central areas of the sheets are now clearly transparent, violating the constraint that a real surface cannot be both opaque and transparent. If the image is interpreted simply as 2-loop wire frame then the opacity constraint would be respected. However, if it is still seen as two occluding sheets then the opacity constraint is ignored by the visual system.

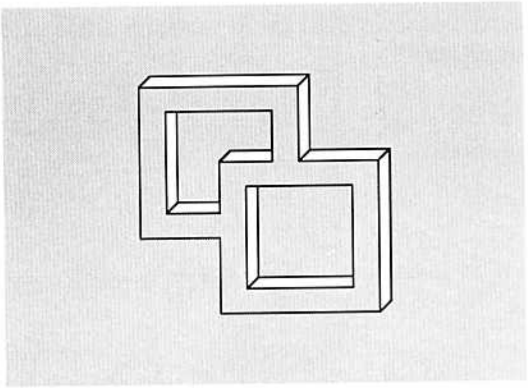

FIG. 7. When the overlapping squares of Fig. 6 are presented as a random-dot stereogram, the lines appear to be raised from the background as shown here, although the edges of the raised areas would not be visible. Since a background surface can be seen through the area bounded by the raised "lines," this area must be transparent, not opaque. It cannot be hiding the missing corner of the incomplete square. The shape information—*T* intersections indicating occlusion by an opaque surface—is therefore in conflict with the binocular disparity information that suggests that the central area is transparent.

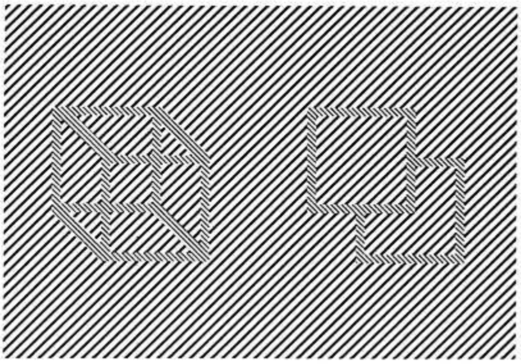

FIG. 8. The two test figures are defined by texture differences alone.

2.2. *Color*

Although the figures are more difficult to see at equiluminance, the observers report that they retained their 3-dimensional structure. The color pathway is thus able to convey intersection information.

2.3. *Texture*

The figures are still visible when the lines differ from the field only in terms of texture (Fig. 8). The figures are most visible with coarse texture on a uniform background, but also visible with uniform lines on a coarse texture field, or with two different textures.

2.4. *Binocular Disparity*

Binocular disparity provides a direct cue to the depth of surfaces that may conflict with the depth suggested by intersections. Certainly some binocular disparity information must be discounted to interpret drawings at all. Binocular disparity cues in a standard drawing indicate that the image is flat and yet it is interpreted as a 3-dimensional object.

The line drawings presented with the lines in front as a stereogram retain the shape of the Necker cube as the depth does not conflict with the transparent interpretation of the spaces between the lines (Fig. 9). On the other hand, the occlusion figure with the lines raised in front of the field (Fig. 10) at first appears to lose its original interpretation as the depth of the field now contradicts the opaque interpretation necessary to see one sheet in front of the other. It might be argued that the ***T*** intersection cannot be interpreted from a disparity representation but, more likely, the intersection information is simply overruled by the depth of the field. However, several observers reported that they still saw this as an occlusion figure with two sheets one in front of the other, as if it were drawn on a transparency, like an etched-glass French door. They were not disturbed by the fact that they could see through what should have been an opaque surface. Other observers could not see this organization even when it was pointed out to them.

It appears that in some cases the opacity constraint can be violated and therefore that there is not a strict precedence of binocular disparity over intersection cues.

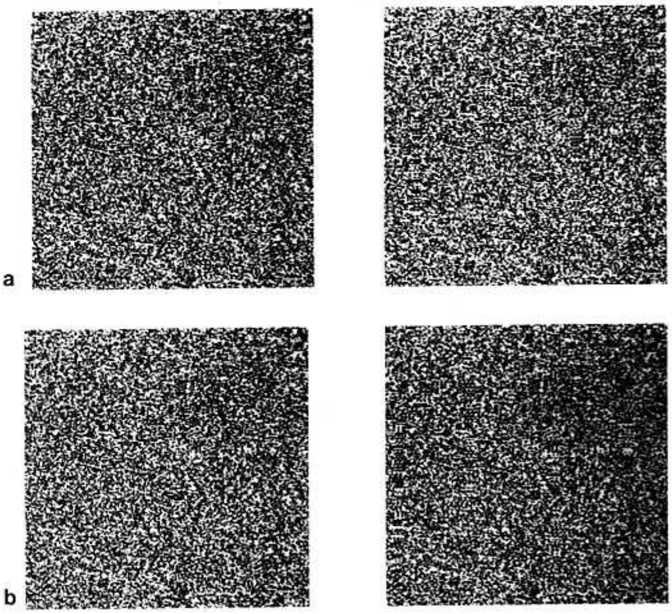

FIG. 9. Necker cube figure. The left- and right-hand images of each pair must be fused to produce a stereo image. If they are fused with uncrossed disparity (left image to the left eye and right image to the right eye) then the "lines" in (a) will appear in front of the background and those in (b) will appear to be cut into the background. If they are fused with crossed disparity (left image to right eye and right image to left eye) then the lines will be behind in (a) and in front in (b).

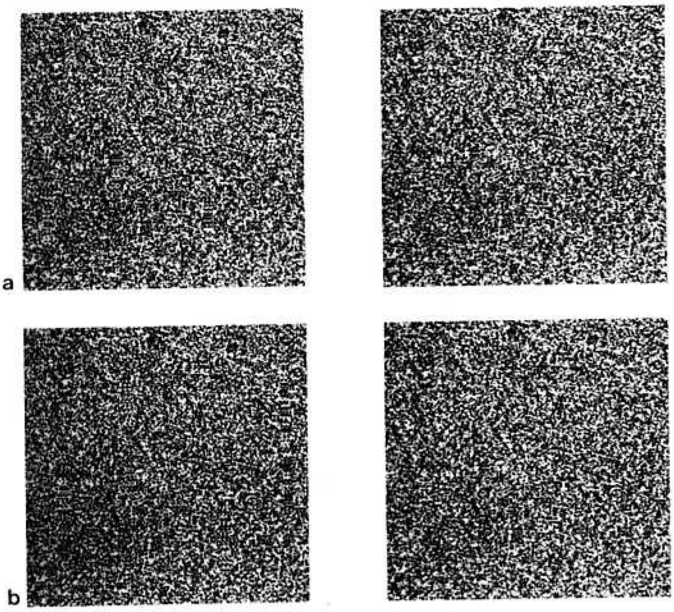

FIG. 10. The overlapping squares figure. The left- and right-hand images of each pair must be fused to produce a stereo image. If they are fused with uncrossed disparity then the "lines" in (a) will appear in front and those in (b) will appear behind. If they are fused with crossed disparity then the lines will be behind in (a) and in front in (b).

The priority of the two cues varies from observer to observer. Dosher, Sperling, and Wurst [12] reported a similar individual variation in the weightings that observers gave to two conflicting depth cues used in their experiment: binocular disparity and luminance-proximity cues.

When the image lines of the Necker cube are presented behind the field (Fig. 9), most observers see geometric islands separated by gaps through which a back plane is visible. Others, however, see the Necker cube in the back plane even though the islands are aligned in front of what should be transparent spaces between the lines. For some observers, then, the improbable coincidence of the alignment between the edges of the Necker cube lines and those of the floating islands does not impede the 3-dimensional interpretation of the Necker cube. For others, this interpretation cannot be seen even when it is pointed out to them.

For the occlusion stimulus (Fig. 10), most see these two areas at the same depth plane, one a square and one a rotated *L*-shape. Others report that the bottom, right-hand surface is seen as a square with one corner hidden behind the other square even though there is an intervening gap through which a back field is visible, contradicting the hypothesis of an opaque surface.

2.5. *Motion*

Like binocular disparity, relative motion provides direct information concerning the depth relations of objects [15, 20]. The stimuli had equivalent random texture within the lines of the figures and in the background field. The texture in the lines was then made to move while the texture in the background remained stationary or vice versa. Because of the accretion and deletion at their edges, the static areas always gave the impression of being in front of the moving areas. Only the textures moved, the areas of moving texture defining the lines of the figures stayed in place. The observations with motion-defined stimuli were analogous to those for stereograms. In instances where there were conflicts between the depth cues from relative motion and those from intersections, some observers gave precedence to the relative motion and others to the intersections.

2.6. *Discussion*

The interpretation of line drawings, and, in particular, of the intersection cues to relative depth, appears to depend on a general purpose analysis that has access to information coming from all visual pathways. Binocular disparity and relative motion cues could override the depth implied by intersections for some observers. For others, the depth implied by intersections took precedence over that implied by binocular disparity and relative motion. For these observers, the precedence of intersections implied that opacity and coincidence constraints could be violated. The difficulty of seeing the intersection interpretations (several observers could not see them even after considerable effort) is perhaps related to the constraint violation.

3. SUBJECTIVE CONTOURS

When an occluding surface is not visibly different from parts of the background that it occludes, there is no physical border at the edge of the surface. But we often do see a contour at these locations, some involving a brightness difference (Fig. 11a) but others simply an edge (Fig. 11b). The cue to these surface occlusions is the

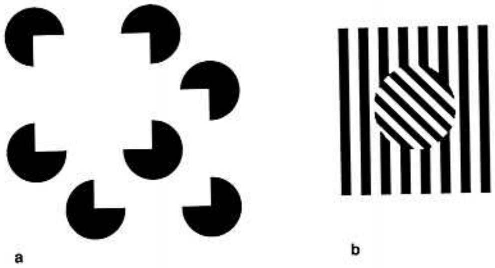

FIG. 11. The two subjective contour test figures. (a) Two subjective squares. The top left square covers the corners of four black disks and partially overlaps the bottom right square. The bottom square covers the corners of the three visible disks and, presumably, the corner of the missing fourth disk hidden underneath the top square. This figure produces an impression that the squares are slightly brighter than the background and that the square sheets and round disks are staggered in depth on top of each other. (b) A subjective disk. A round, striped disk appears to lie slightly above a striped background. Although the disk seems to have a very clearly defined border around it, it does not appear to differ in average brightness from the background.

shape of the occluded object which appears to be a regular shape with an abrupt irregularity: several incomplete disks in the left-hand figure and aligned line terminations in the right-hand figure. Several authors have attributed the subjective contours to a solution of these irregularities, explaining the overall figure as an occlusion between two or more simple shapes rather than by several irregular shapes [17, 35]. Since these occlusion cues are signaled by shape, we might think that they would behave in the same manner as the line intersection cues just examined. In fact, Stevens [38] described a patient with visual agnosia who had a deficit specific to the interpretation of occlusion. He could see neither monocular occlusion cues in line drawings nor subjective contours in Kanizsa figures (such as, Fig. 11a). Stevens therefore felt that both these types of cues were processed by a common module responsible for analyzing occlusion. However, other evidence shows that this cannot be the case. In particular, occlusion cues in line drawings can be interpreted in the absence of a luminance difference between the figure and the field (for example, Fig. 9). Gregory [18], Brussell, Stober, and Bodinger [10], and Prazdny [34], on the other hand, have shown that a luminance difference *is* required for subjective contours to be seen. When the incomplete disks of a Kanisza figure are a different color but the same luminance as the background, no subjective contours are seen [10, 18]; similarly, disks defined by flicker, relative motion, or binocular disparity do not produce subjective contours [34]. This dependence on luminance is unexpected if we think of subjective contours as resulting from a cognitive or perceptual hypothesis based on shape, since shape can be signaled by any pathway, not only the luminance pathway.

The subjective figures tested previously for the role of luminance all involved subjective brightness effects such as that seen in Fig. 11a. In order to examine whether the failure to obtain subjective contours in the absence of a luminance difference is related to the brightness aspect of the figure, a subjective contour (Fig. 11b) that does not involve a brightness effect was tested as well. Figures defined by texture were tested in addition to those defined by color [10, 18], relative motion and binocular disparity [34].

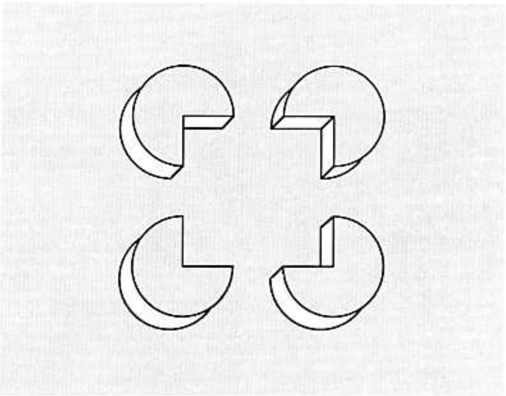

FIG. 12. If the occluded disks of Fig. 11a are presented as a random dot stereogram, the $\frac{3}{4}$ pie shapes will stand out from the background, as represented here. It is clear that there is no central, opaque square hiding the corners of the disks. The corners are, in fact, missing and the central area is transparent, not opaque. The shape information—incomplete disks suggesting occlusion by an opaque surface—is therefore in conflict with the binocular disparity information that suggests that the central area is transparent.

As was the case with the line drawings, there can be conflicts between the depth implied by the occlusion cues in the figures and that implied by binocular disparity or relative motion when these attributes are used to define the figures. These conflicts again test the influence of the opacity constraint. Fig. 12 depicts the one of the subjective sheets of Fig. 11a when represented by binocular disparity or relative motion. Even though the shape information is identical to that for a luminance-defined figure, the added depth information implies that the four disks *are* actually incomplete and not occluded. The central square area, which should be opaque in order to occlude the disks at its corners, is transparent. If observers can maintain the subjective contour interpretation in this situation, they must be violating the opacity constraint.

The observers task was to determine whether the subjective contours and the related relative depths of the surfaces involved were visible in the different viewing conditions. For example, in Fig. 11a, the observers reported whether the two subjective sheets and the disks appeared to be staggered in depth, or whether the figures seemed to be made up of independent shapes, pac-men, all at the same depth plane.

When the figures are defined by attributes other than luminance (e.g., by equiluminous colors or textures) they are more difficult to see, as if they were at a low contrast level and blurred. These factors on their own, however, did not interfere with the perception of subjective figures. Observers were shown the test figures defined by luminance at low contrast (light grey disks on a slightly darker field) or blurred (low pass filtered). They reported that the subjective contours in the low contrast figures became visible as soon as the elements of the figure reached visibility and that good subjective contours could be seen even for very blurred figures. Therefore, a loss of subjective contours on figures defined in pathways other than luminance should not be attributed simply to the lower effective contrast and resolution of the pathway.

3.1. Color

The stimuli of Figs. 11a and b were presented to observers by replacing the black areas with green and the white areas with red. The relative luminances of the red and green were then varied over a range including the equiluminant point, the point at which red and green have the same luminance. At this point, the figures produced no image in the luminance pathway, only in the chromatic pathway. Observers reported that there was a range of relative luminances between the green disks and red field for which no subjective contours were evident in the overlapping sheets figure. Within this range, the figure appeared as a disconnected set of pac-men, all at the same depth plane. The same was true for the overlying disk figure. The central disk which appeared to sit in front of the background when there was a luminance difference clearly fell back into the plane of the field when the red and green were at equiluminance. Luminance therefore appears to be necessary for subjective contours even with figures that do not involve subjective brightness effects. The shapes signaled by the chromatic pathway are not sufficient to trigger the perception of subjective figures.

3.2. Texture

The stimuli were produced by replacing the black areas of Figs. 11a and b with a random speckle of 35% luminance contrast and the white areas with a uniform gray. The relative luminance of the textured and uniform areas were then varied over a

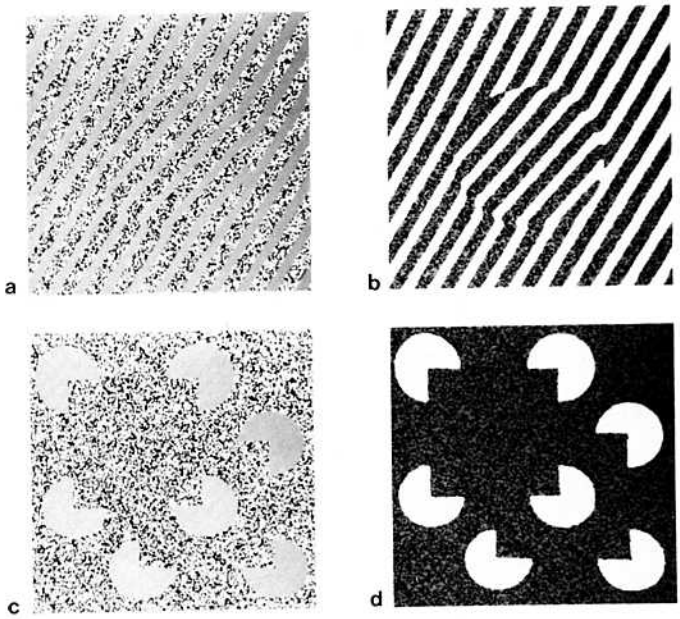

FIG. 13. The subjective disk figure is represented by a texture difference with (a) the background at the same mean luminance as the stripes and (b) the background darker that the stripes. The overlapping squares figure is represented by a texture difference with (c) the background at the same mean luminance as the incomplete disks and (d) the background darker that the incomplete disks.

range that included the equiluminance point. The observations were the same for the texture-defined figures as for the color-defined figures.

Subjective contours were visible when there was a luminance difference (Figs. 13b and d) but not when there was no luminance difference (Figs. 13a and c). Shapes signaled by texture differences alone were not capable of triggering the perception of subjective contours. Note that the texture-defined stimulus, Fig. 13a in particular, is quite visible. It is not the case that the subjective contours disappear because the stimulus is too difficult to see.

3.3. Binocular Disparity

When the figures were presented as stereograms without luminance differences (Figs. 14a and 15a), again no subjective contours were seen. For example, observers

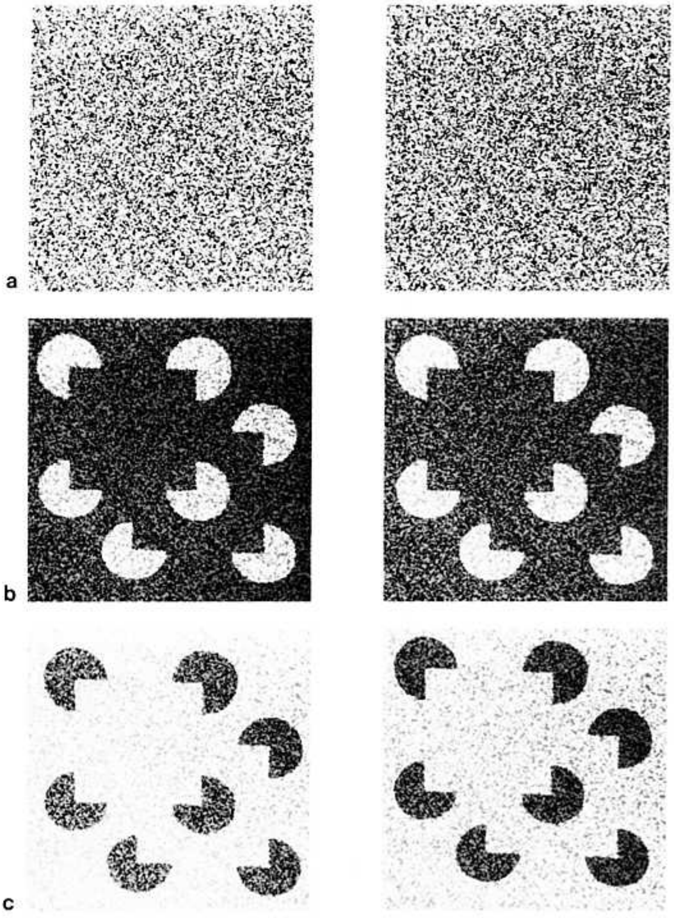

FIG. 14. The overlapping squares figure is represented by a random dot stereogram. The left and right images of each pair must be fused **with uncrossed disparity** to produce the stereo image (the depth edge for the random dots and that for the brightnesses border are not correctly aligned if viewed with crossed disparity). The incomplete disks appear in front of the field in (a) and (b) and behind in (c). The brightnesses of the incomplete disks and the field are equal in (a) while the rearmost plane is darker in (b) and (c).

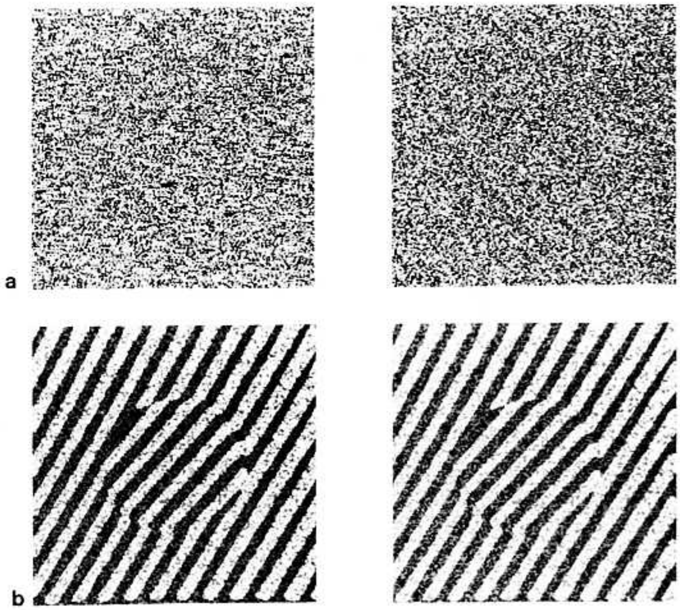

FIG. 15. The subjective disk figure is represented by a random dot stereogram. The left and right images of each pair must be fused **with uncrossed disparity** to produce the stereo image (the depth edge for the random dots and that for the brightness borders are not correctly aligned if viewed with crossed disparity). The stripes appear in front of the field. The brightnesses of the stripes and the field are equal in (a) while the stripes are lighter than the field in (b).

reported that Fig. 14a appeared as a set of unrelated pac-men, although it was difficult to see all the elements of the figures clearly.

When the overlapping sheets image is presented with a brightness difference (Figs. 14b and c), a conflict arises between the depth indicated by binocular disparity and that suggested by the occlusion cues, the incomplete disks. As was the case with conflicting cues in the line drawings, the interpretation varied substantially from observer to observer. When the figure was viewed with the disks in front (Fig. 14b, note that Figs. 14 and 15 *must* be viewed with uncrossed disparity, for example, through a stereo viewer to be seen correctly) some observers saw it as unrelated pac-men floating above a back plane, while others reported seeing an occluding surface that was both opaque, hiding the corners of the disks, and transparent, as if it were a piece of plexiglass. Apparently, it is not necessary for a surface to actually be opaque for it to be interpreted as an occluding surface. When the figure was viewed with the disks in the back plane (Fig. 14c), many observers reported clear subjective contours with a staggering of the sheets even though there was a hole through one sheet where the central disk is seen as a cutout. Others saw no subjective figures here, reporting rather that the front plane seemed to be a thin sheet all at the same depth plane on which there were several pac-men cutouts through which a back plane could be seen.

Note again that the subjective figures are visible monocularly and that any difference in the organization seen when viewed binocularly must arise from the influence of the binocular disparity information. All observers said that they were able to see the depth from binocular disparity in the image but it is possible that some of the individual variation in the interpretations was due to variations in the strength of the observers' stereo vision.

When the overlying disk was presented with a luminance difference (Fig. 15b), the subjective disk of light stripes appeared to arch slightly in front of the field. The depth difference between the light and dark stripes gave the figure the appearance of an egg slicer with rows of stripes in front between which a back plane could be seen. The back plane appeared uniform with no subjective effects. Note that the subjective figure is visible in each eye's image alone and that the depth produced by the binocular disparity does not conflict with that suggested by the line terminators of the light stripes of the disk. The subjective effect in the dark stripes does appear to be suppressed however.

3.4. Motion

The observations with the motion figures were identical to those for the stereo figures. No subjective contours were visible without a luminance difference. When luminance was added, the interpretations varied from individual to individual in the same way that they did for the stereo figures.

3.5. Discussion

As reported previously [10, 18, 34], luminance is essential for the visibility of subjective contours, and this is true as well in figures that do not involve subjective brightness effects (Figs. 13e and 15a). The other representations examined, color, texture, stereo and motion, were all capable of signaling the shapes of the figures but did not produce impressions of subjective contours. The critical aspect of the figure that creates the subjective contour cannot be shape alone, therefore. If subjective surfaces are generated cognitively in order to simplify the organization of incomplete image shapes [17, 35], these surfaces ought to be generated no matter what attributes define the incomplete shapes. Since the subjective effects are seen only for figures defined by luminance, purely cognitive hypotheses for subjective contours must be ruled out.

What processes specific to the luminance pathway lead to the perception of subjective contours? Von der Heydt, Peterhans, and Baumgartner [44] described cells in area 18 that responded to subjective contours. Since luminance is essential for subjective contours to be seen, we could add that these cells should respond only to inducing figures defined by luminance. Area 18 would appear to be a remarkably early level in the visual system to be responsible for an effect that has been labeled cognitive. It is unlikely, however, that the entire occluding surface is inferred at this low level. What may be signaled is simply the contour, or potential contour, in addition to the partial shapes that are physically present in the stimulus. All of this information would be considered in the overall interpretation of the figure at a higher level. Thus it may not be the subjective figure that requires luminance but the contours; these would be signaled at a low level, in the luminance pathway only, and then contribute to an interpretation of an occluding surface.

The processes at this higher level where surfaces are inferred do not seem to respect the opacity constraint in determining the interpretation of the figure. The resolution of the conflict involving opacity varied a great deal from observer to observer, however.

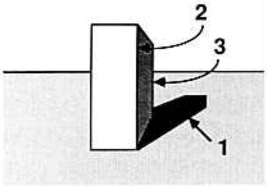

FIG. 16. Three types of shadow border are seen here: (1) a same-surface border, (2) a joined-surface border, and (3) an occluded-surface border.

4. SHADOWS

Shading has been shown to provide very powerful cues for the recovery of 3-dimensional structure [7, 8, 23, 45, 46]. Yvon Leclerc and I have examined which visual pathways convey shadow information and 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 only a change in illumination and not a change in material—we have labeled this a same-surface border and examples are shown in Figs. 16 and 17. A joined-surface border and an occluded-surface border are much less constrained as there may be changes in material and reflectance across the border as well as changes in illumination.

To study the effect of constraints, we concentrated on the same-surface border. Such a border can result from an object casting a shadow on another surface as in the lower shadow of Fig. 17 or a gradual change in surface orientation as in the shading on the sphere in Fig. 17. In either case it is the same material on both sides of the shadow border. Since there is a single material surface being arbitrarily divided by a change of illumination, we do not expect any property to change across the border except, of course, brightness. Specifically, there are five constraints:

(1) Brightness should decrease in the shadow region.

(2) The nature and the contrast of the surface texture, if any, should be the same on both sides of the shadow border.

(3) Only certain color changes can occur across the shadow border.

(4) and (5) No change in either motion or depth is expected at the shadow border. Not only is such a change aligned with the shadow border unlikely but also it generates a conflicting depth cue.

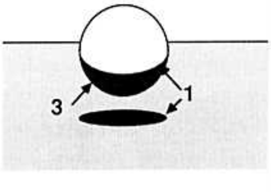

FIG. 17. A same-surface border may arise from a cast shadow as on the lower surface, or from an attached shadow as in the shading edge around the middle of the sphere, both of these borders are labeled with a **1**. The bottom of the sphere presents an occluded-surface shadow border, labeled **3**.

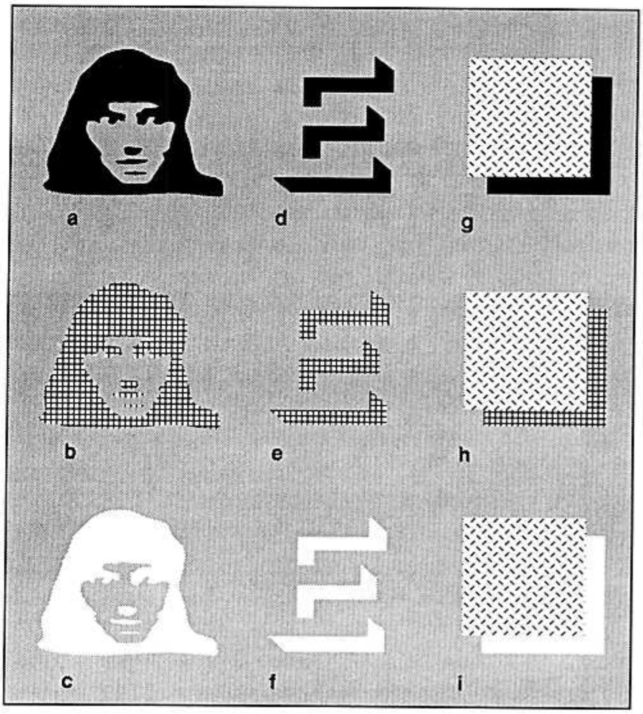

FIG. 18. A binary (i.e., two region) image of a woman's face with three different contrasts between the two regions: (a) the first has positive contrast, appropriate for a shadow, (b) the second has no contrast difference, only a texture difference, and (c) the third has negative contrast, inappropriate for a shadow. A similar sequence, (d), (e), and (f), for a binary image of a 3-dimensional block letter *E*, although here, both directions of contrast are appropriate for producing shadows. The same sequence for a 3-level image of a rectangle covering its cast shadow where only the contrast of the shadow is varied and is (g) appropriate, (h) null, or (i) inappropriate for a shadow region.

We evaluated these constraints on several images. In particular, we were looking for images that changed appearance significantly when the shadows were correctly interpreted. Fig. 18 shows three of the images we examined: a woman's face, a shadowed block letter, and a rectangle casting a shadow.

Figs. 18a, b, and c shows three versions of high contrast image of a face with the brightness difference between shadow and non-shadow regions varying from top to bottom. Note that the shadow is interpreted and the surface of the face seen in relief only if the shadow region is darker than the non-shadow region. Most of the shadows, those cast by the nose, 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.

The other images in Fig. 18 show two of the several other images we have studied. We limited the cues in the images to shadows rather than graded shading for technical reasons. We wanted to substitute attributes such as motion and texture for brightness and since it is difficult to do this in a graded fashion, binary, or 2-level, images were most appropriate. In addition, we needed a stimulus that changed significantly if the shadow regions were not interpreted as shadows. The face and the shadowed letter of Fig. 18 fill this requirement. The shadowed letter has the property of being interpreted as a shadow figure for either positive (Fig. 18d) or negative contrast (Fig. 18f); the illuminated and shadowed surfaces are simply interchanged. At null contrast (Fig. 18e), the letter appears as three unrelated bars. We did not use these letter stimuli however, because they also involve significant subjective contours along the incomplete edges of the block form and we did not want to confound the effects for shadows with those for subjective contours. The rectangle figure with the cast shadow (Figs. 18g, h, and i) has the drawback that even when the back area is not seen as a shadow (Figs. 18h and i), it is still seen behind the front square because of the occlusion cues. There is no dramatic change in the organization of the percept as the back area changes from being seen as a shadow to being seen simply as a partially covered rectangle.

We initially attempted Gilchrist, Delman, and Jacobsen's [16] method for measuring the illumination hypothesis underlying the perception of a shadow. We introduced a test patch halfway into the cast shadow of the rectangle figure and asked observers to set the brightness of the patch inside the shadow area so that it appeared to match that outside the shadow. The shadow area was not called a shadow but merely the outside strip. Their response was that their setting depended on whether they saw the strip as a shadow or not. They were quite happy with two very different adjustments for the identical stimulus depending on their assumption.

Because of these difficulties, we chose a purely subjective measure of the shadow interpretation. The face image was altered by adding color, texture, depth, or motion differences to the two areas of the image. Observers were then asked whether or not the stimulus was seen as a face with noticeable depth in the features. If, for a particular viewing condition, it was difficult to see the face, we concluded that some aspect of the viewing condition was interfering with the perception of the shadow regions.

Generally, when the shadows were not correctly interpreted, the different regions of the figure gave the impression of being segregated surfaces, islands of color or texture. For example, the light/dark borders of the face stimulus, when incorrectly interpreted (Figs. 18b and c), 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 did not appear to stop at the light/dark border and the impression was that of a face (Fig. 18a) that had a distinct surface relief around the nose and under the eyebrows and a particular expression. Thus, to be correctly seen, the borders had to appear as a change in illumination on an unbroken surface and not as a change of material.

4.1. 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 Fig. 19, two differently colored suns. In this case, however, the color that is in the shadow must *always* also fall in its surround. There are therefore constraints on the saturation of the colors in

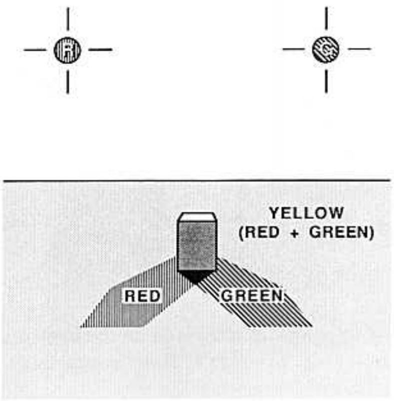

FIG. 19. Colored shadows are produced by two differently colored light sources. Whatever color falls in a shadow must also fall in its surround.

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.

The light and dark areas of the original images were replaced with red and green. Although color alone did not support the perception of shadows, the presence of impossible shadow colors did not interfere with the perception of the shadows as long as there was an appropriate brightness difference. The natural constraints concerning colored shadows are violated for a green shadow and red surround and so we conclude that the color constraint is ignored. Rubin and Richards [36] and others [14, 37] have demonstrated that the constraints on color changes across a shadow border can be used to discriminate changes in material from changes in illumination. Our results imply that this potential source of information is not a factor in human perception.

4.2. Texture

A shadow falling on a textured surface should reduce the brightness of the texture but not change its contrast or nature. A higher contrast texture in the shadow region than in the surrounding region cannot occur naturally in a image.

When there was no difference between the mean luminances of the two areas the shadows could not be correctly perceived (Fig. 20a). The texture difference alone could not support the perception of shadows even though the features could be identified. However, if the shadow regions were darker, then the face could be seen (Fig. 20b).

In summary, shadows could be seen even though the texture in the shadow region violated the physical constraint against a change of texture or contrast across a shadow boundary.

4.3. Binocular Disparity

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 tested the effect of displaying the shadow and non-shadow areas at different depth planes on the interpretation of the shadow

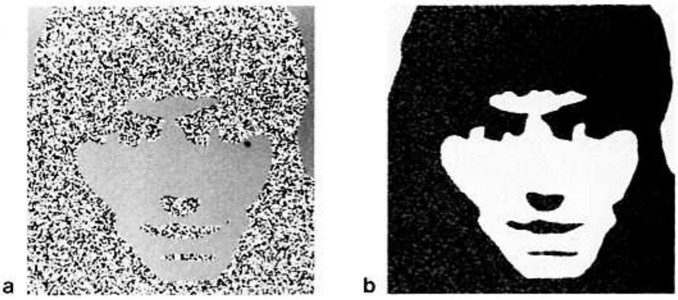

FIG. 20. (a) Depth due to shading cannot be perceived on the basis of a texture difference alone, but can be perceived (b) if the shadow areas are darker, even though there is texture in the shadows, in violation of the texture constraint.

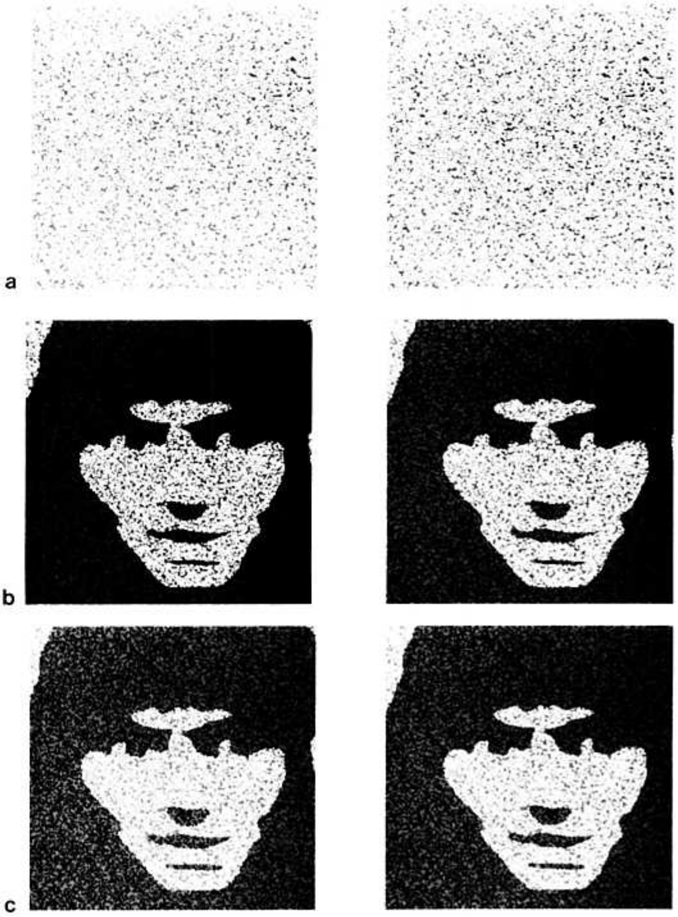

FIG. 21. The figure of the woman's face is represented by a random dot stereogram. The left and right images of each pair must be fused **with uncrossed disparity** to produce the stereo image (the depth edge for the random dots and that for the brightness borders are not correctly aligned if viewed with crossed disparity). The originally dark areas—e.g., hair and shadows—appear behind the field in (a) and (b) and in front in (c). The brightnesses of the originally dark and the originally light areas are equal in (a) but the originally dark areas are appropriately darker in (b) and (c).

regions. (For simplicity, we refer to the originally dark areas as shadow regions and the originally light areas as non-shadow regions, even though the dark areas contain both shadows and areas of low reflectance.)

When the shadow and non-shadow regions had the same brightness (Fig. 21a), the depth alone did not support shadows. The different areas of the face were visible but the overall organization was difficult to see. With an appropriate brightness difference (Figs. 21b and c, note that these figure *must* be viewed with uncrossed disparity), however, the shadows and the face were visible even though the depth change was clearly seen at the brightness edge. When the dark areas of the figure are seen in front (Fig. 21b), it is possible to see through the surrounding areas to a back, lighter plane. The face is visible monocularly but does not change in appearance when viewed binocularly even though the differences in the depth planes of the dark and light areas are very noticeable. It is highly unlikely that a surface such as the cheek will have a depth change exactly aligned everywhere with the shadow border. The face was seen nevertheless. Observers reported that with the light region in front (Fig. 21c), the dark region appeared to be transparent. With the dark region in front, the light region appeared to be seen through a transparency.

4.4. 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 the shape of the 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 a depth change between the two areas and often gives rise to compelling figure/ground organizations [9]. There are therefore two reasons for relative motion between shadow and non-shadow regions to interfere with the perception of 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 gives cues to the organization of the surfaces that are inconsistent with the shadow interpretation of a single surface.

When the shadow and non-shadow regions had the same brightness in our test figure, the motion alone did not support the perception of a shadow even though the shape of the figure could be clearly seen. When an appropriate brightness difference was added in, however, the shadows and the face were correctly interpreted. The change in motion at the shadow border violated 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, whether they both moved in the same direction at different speeds or in opposite directions. 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 and so we conclude that motion cues are ignored when interpreting shadows.

4.5. Discussion

Many authors have proposed that visual information is broken down into multiple representations [1, 43, 49] and that there is cooperation between these representations to determine a consistent interpretation of form [3, 29, 40, 41]. This apparently does not occur in the analysis of shadows as only luminance appeared to influence the interpretation of shadow regions. All other cues, whether from color, texture, motion, or depth, were ignored. Some attempt appeared to be made to deal with contrary evidence, for example, by hypotheses of transparency in the case of inconsistent depth, but this was after the fact. The regions involved had already been signaled as being shadows. Other studies have shown that shading can override binocular disparity. Yellott and Kaiwa [47] and Georgeson [13] 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.

Natural constraints have been useful for solving many visual problems [33]. Disparity constraints have been used in computing stereo images [30, 31] and smoothness constraints to resolve ambiguous contour motion [21] and orientation fields [51]. In the case of shadows, however, constraints other than brightness appear to play no role. The brightness constraint appears to be verified particularly at the edges of shadow regions.

The analysis of shadows therefore appears to be limited to the luminance pathway. This might have arisen as a consequence of the earlier development of the processing of shadows in the visual system. When color and binocular depth analyses evolved later on, they may have simply been added on without interacting with the already established shadow interpretation processes. This argument does not hold for motion cues, however, which most probably evolved as early, if not earlier, than the analysis of shadows. Although the evidence suggests that the interpretation of shadows is a low-level process specific to the luminance pathway, it may be that it is simply the edges of regions that are signaled as appropriate for shadows at this low level. The interpretation of the shadow and non-shadow regions and their relative depth might then occur at a higher level.

Information concerning motion, depth, color and texture may not be used in interpreting shadows, but what about luminance information? Observers are skilled in discounting illumination changes in shadow areas to perceive surface reflectances [7, 16]. On the other hand, they are not very sensitive to the shape cues in shading gradients. Barrow and Tenenbaum [4], 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 [39] reported that observers made errors of up to 50% in estimating surface curvature based on shading. It would appear that the visual system is not making sophisticated computations of light sources and reflectance normals [14, 32, 37, 46], but rather has some very loose rules for what is and what is not a shadow.

The shadows we have studied here have all been presented as binary images with only light and dark areas and so our conclusions must be limited to shadow images of this type for the moment. Although the principal observations have been made with face stimuli, the results are not limited to faces. We find similar observations with other stimuli such as pictures of shadowed hats and shoes. We are also extending our study to more natural shadows but we have not yet observed any qualitative differences.

5. CONCLUSION

The findings here imply that a variety of mechanisms analyze depth cues. Occlusion cues in line drawings were conveyed by all pathways (Fig. 22). When

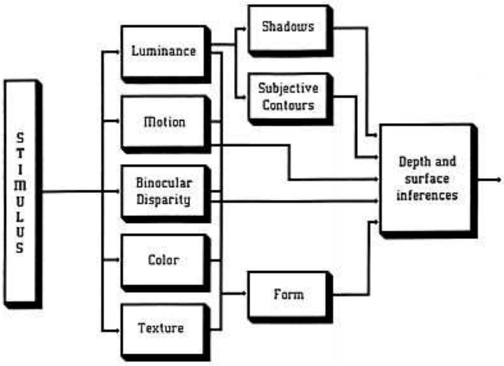

FIG. 22. All pathways were capable of conveying contour intersection cues in line drawings. Subjective contour and shadow cues were signaled only by the luminance pathway. Depth information from binocular disparity and relative motion participated directly in inferences concerning depth and surfaces as well as indirectly through shapes defined by binocular disparity and relative motion.

there were conflicts between cues, the final interpretation could be weighted towards intersection cues or towards binocular disparity and relative motion cues depending on the observer and on the observer's perceptual hypotheses. Occlusion cues in subjective figures were conveyed only by the luminance channel, however. Cue conflicts for subjective figures led to varying interpretations similar to those seen for line drawings. Shadow information was also restricted to the luminance pathway but shadow cues always overruled other cues in the case of conflicts. This diversity of characteristics for the three relative depth cues examined here appears to argue against one general-purpose process analyzing all aspects of depth in the image. The alternative of a multitude of special-purpose processes would entail a great deal of redundancy, however: the analyses concerning form and 3-dimensional organization necessary to interpret each cue in isolation would have to be duplicated in many if not all of the modules.

A less extreme alternative is that all depth cues may be analyzed by a common set of processes which access different types of image data depending on the cue in question. For example, although subjective contour and shadow cues appear to be dependent on low-level processes specific to the luminance pathway, it may be only the signaling of appropriate borders that occurs at this low level. Von der Heydt *et al.* [44], for example, have demonstrated that cells in area 18 respond to subjective contours. Interpretation of the surfaces that could give rise to the signaled contours might occur at a higher level as would the production of the depth and brightness impressions that accompany subjective contours. Similarly for shadows, contours having a consistent and appropriate contrast all along their edge might be signaled at an early stage in the luminance pathway, but a subsequent, general-purpose stage would determine, as one of its tasks, whether the area bounded by that contour is a shadow or not and produce the related depth and brightness properties.

Finally, although natural constraints have proved to be useful in solving many visual problems [21, 30, 33], they do not seem to play a significant role in the interpretation of the depth cues examined here. The perceptual hypotheses of the observers appeared to play a far more important role in determining the interpretations of the figures than did natural constraints. Given the variability of the interpretations seen by our observers, including surfaces that were simultaneously opaque and transparent, there does not seem to be a constraint that a strong perceptual hypothesis cannot overrule. As artists discovered long ago, vision is not limited only to those interpretations of an image that respect the physics of naturally occurring scenes.

ACKNOWLEDGMENT

This research was supported by Grant A8606 to PC from the Natural Sciences of Engineering Research Council of Canada.

REFERENCES

1. J. M. Allman, J. F. Baker, W. T. Newsome, and S. E. Petersen, Visual topography and function: Cortical visual areas in the Owl monkey, in *Cortical Sensory Organization*, Vol. 2: Multiple Visual Areas (C. N. Woolsey, Ed.), pp. 171–186, Humana Press, Clifton, N.J., 1981.
2. D. H. Ballard and C. M. Brown, *Computer Vision*, Prentice-Hall, Englewood Cliffs, N.J., 1982.
3. H. G. Barrow and J. M. Tenenbaum, Recovering intrinsic scene characteristics from images, in *Computer Vision Systems* (A. Hanson and E. Riseman, Eds.), pp. 3–26, Academic Press, New York, 1978.
4. H. G. Barrow and J. M. Tenenbaum, Computational vision, *Proc. IEEE* **69**, 1981, 572–595.
5. H.G. Barrow and J. M. Tenenbaum, Interpreting line drawings as three-dimensional surfaces, *Artif. Intell.* **17**, 1981, 75–116.
6. T. O. Binford, Inferring surfaces from images, *Artif. Intell.* **17**, 1981, 205–244.
7. J. Beck, *Surface Color Perception*, Cornell Univ. Press, Ithaca, 1972.
8. K. Berbaum, D. Tharp, and K. Mroczek, Depth perception of surfaces in pictures: Looking for conventions of depiction in Pandora's box, *Perception* **2**, 1983, 5–20.
9. O. Braddick, A short-range process in apparent motion, *Vision Res.* **14**, 1974, 519–528.
10. E. M. Brussell, S. R. Stober, and D. M. Bodinger, Sensory information and subjective contour, *Amer. J. Psychol.* **90**, 1977, 145–156.
11. M. B. Clowes, On seeing things, *Artif. Intell.* **2**, 1971, 79–116.
12. B. A. Dosher, G. Sperling, and S. A. Wurst, Trade offs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure, *Vision Res.*, **26**, 1986, 973–990.
13. M. A. Georgeson, Random-dot stereograms of real objects: Observations on stereo faces and moulds, *Perception* **8**, 1979, 585–588.
14. R. Gershon, A. D. Jepson, and J. K. Tsotsos, *The Effects of Ambient Illumination on the Structure of Shadows in Chromatic Images*, Technical Report RBCV-TR-86-9, Dept. of Computer Science, University of Toronto, 1986.
15. J. J. Gibson, *The Perception of the Visual World*, Houghton Mifflin, Boston, 1950.
16. A. Gilchrist, S. Delman, and A. Jacobsen, The classification and integration of edges as critical to the perception of reflectance and illumination, *Percept. Pyschophys.* **33**, 1983, 425–436.
17. R. L. Gregory, Cognitive contours, *Nature* **238**, 1972, 51–52.
18. R. L. Gregory, Vision with isoluminant colour contrast: 1. A projection technique and observations, *Perception* **6**, 1977, 113–119.
19. A. Guzman, Decomposition of a visual scene into 3-dimensional bodies, in *Automatic Interpretation and Classification of Images* (A. Grasseli, Ed.), Academic Press, New York, 1969.
20. H. von Helmholtz, *Helmholtz's Treatise on Physiological Optics* (J. P. Southall, Ed.), Optical Society of America, New York.
21. E. C. Hildreth, *The Measurement of Visual Motion*, MIT Press, Cambridge, Mass., 1983.
22. J. E. Hochberg and V. Brooks, Pictorial recognition as an unlearned ability, *Amer. J. Psychol.* **75**, 1962, 624–628.
23. B. K. P. Horn and K. Ikeuchi, Numerical shape from shading and occluding boundaries, *Artif. Intell.* **15**, 1981, 141–184.
24. D. A. Huffman, Impossible objects as nonsense sentences, in *Machine Intelligence* Vol. 8 (B. Meltzer and D. Mitchie, Eds.), Edinburgh Univ. Press, Edinburgh, 1971.
25. B. Julesz, *Foundations of Cyclopean Perceptions* Univ. of Chicago Press, Chicago, 1971.

26. J. M. Kennedy, A *Psychology of Picture Perception: Information and Images*, Jossey–Bass, San Francisco, 1974.
27. J. J. Koenderink and D. van Doorn, The shape of smooth objects and the way contours end, *Perception* **11**, 1982, 129–137.
28. A. K. Mackworth, Interpreting pictures of polyhedral scenes, *Artif. Intell.* **4**, 1973, 121–137.
29. D. Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, Freeman, San Francisco, 1982.
30. D. Marr and T. Poggio, A cooperative computation of stereo disparity, *Science* **194**, 1976, 283–287.
31. J. E. W. Mayhew and J. P. Frisby, Psychophysical and computational studies towards a theory of human stereopsis, *Artif. Intell.* **17**, 1981, 349–387.
32. A. P. Pentland, Finding the illuminant direction. *J. Opt. Soc. Amer.* **72**, 1982, 448–455.
33. T. Poggio, V. Torre, and C. Koch, Computational vision and regularization theory, *Nature* **317**, 1985, 314–319.
34. K. Prazdny, On the nature of inducing forms generating perceptions of illusory contours, *Percept. Psychophys.* **37**, 1985, 237–242.
35. I. Rock and R. Anson, Illusory contours as the solution to a problem, *Perception* **8**, 1979, 655–681.
36. J. M. Rubin and W. A. Richards, Color vision and image intensities: When are changes material? *Biol. Cybern.* **45**, 1982, 215–216.
37. S. A. Shafer, *Using Color to Separate Reflection Components*, Technical Report TR-136, Dept. of Computer Science, University of Rochester, 1984.
38. K. A. Stevens, *Occlusion Clues and Subjective Contours*, MIT A.I. Memo 363, 1976.
39. J. T. Todd and E. Mingolla, Perception of surface curvature and direction of illumination from patterns of shading, *J. Exp. Psychol.: Human Percept. Perform.* **9**, 1983, 583–595.
40. A. Treisman, Focused attention in the perception and retrieval of multidimensional stimuli, *Percept. Psychophys.* **22**, 1977, 1–11.
41. A. Treisman and G. Gelade, A feature-integration theory of attention, *Cognit. Psychol.* **12**, 1980, 97–136.
42. K. J. Turner, Computer perception of curved objects using a television camera, in *Proceedings, 1st AISB Conf.*, Sussex, 1974.
43. D. C. van Essen, J. H. R. Maunsell, and J. L. Bixby, Organization of the extrastriate visual areas in the Macaque monkey, in *Cortical Sensory Organization*, Vol 2: *Multiple Visual Areas* (C. N. Woolsey, Ed.), pp. 157–170, Humana Press, Clifton, N. J., 1981.
44. R. von der Heydt, E. Peterhans, and G. Baumgartner, Illusory contours and cortical neuron responses. *Science* **224**, 1984, 1260–1262.
45. D. Waltz, Generating semantic descriptions from drawings of scenes with shadows, in *The Psychology of Computer Vision* (P. H. Winston, Ed.), McGraw–Hill, New York, 1975.
46. R. J. Woodham, Analyzing images of curved surfaces, *Artif. Intell.* **17**, 1981, 117–140.
47. J. I. Yellott, Jr. and J. L. Kaiwa, Depth inversion despite stereopsis: The appearance of random-dot stereograms on surfaces seen in reverse perspective, *Perception* **8**, 1979, 135–142.
48. A. Yonas, W. T. Cleaves, and L. Pettersen, Development of sensitivity to pictorial depth, *Science* **200**, 1978, 77–79.
49. S. M. Zeki, Functional specialization in the visual cortex of the rhesus monkey, *Nature* **274**, 1978, 423–428.
50. S. M. Zeki, The representation of colours in the cerebral cortex, *Nature* **284**, 1980, 412–418.
51. S. W. Zucker, Early orientation selection: Tangent fields and the dimensionality of their support, *Comput. Vision Graphics Image Process.* **32**, 1985, 74–93.