COGNITION

Cognition 67 (1998) 45–71

Recovery of 3D volume from 2-tone images of novel objects

Cassandra Moore*, Patrick Cavanagh

Vision Sciences Laboratory, Harvard University, Cambridge, MA, USA

Abstract

In 2-tone images (e.g., Dallenbach's cow), only two levels of brightness are used to convey image structure – dark object regions and shadows are turned to black and light regions are turned white. Despite a lack of shading, hue and texture information, many 2-tone images of familiar objects and scenes are accurately interpreted, even by naive observers. Objects frequently appear fully volumetric and are distinct from their shadows. If perceptual interpretation of 2-tone images is accomplished via bottom-up processes on the basis of geometrical structure projected to the image (e.g., volumetric parts, contour and junction information) novel objects should appear volumetric as readily as their familiar counterparts. We demonstrate that accurate volumetric representations are rarely extracted from 2-tone images of novel objects, even when these objects are constructed from volumetric primitives such as generalized cones (Marr, D., Nishihara, H.K., 1978. Proceedings of the Royal Society London 200, 269–294; Biederman, I. 1985. Computer Vision, Graphics, and Image Processing 32, 29–73), or from the rearranged components of a familiar object which is itself recognizable as a 2-tone image. Even familiar volumes such as canonical bricks and cylinders require scenes with redundant structure (e.g., rows of cylinders) or explicit lighting (a lamp in the image) for recovery of global volumetric shape. We conclude that 2-tone image perception is not mediated by bottom-up extraction of geometrical features such as junctions or volumetric parts, but may rely on previously stored representations in memory and a model of the illumination of the scene. The success of this top-down strategy implies it is available for general object recognition in natural scenes.

Keywords: Two-tone images; Object recognition; Perceptual organization; Shadows

1. Introduction

Two-tone images are sparse, luminance-based representations which typically use

* Corresponding author. Present address: University of California, Los Angeles, Psychology Department, 128a Franz Hall, Los Angeles, CA 90095-1563, USA.

PII S0010-0277(98)00014-6

only black ink on a white background. Though scant in information, this form of depiction is highly representative of familiar scenes and objects (Fig. 1) and the perception of the 3D structure in these scenes is often immediate and effortless. However, not all such images are so easily interpreted. Several famous examples, such as the Dalmatian dog or Dallenbach's cow, are seen initially (for many of us) as scattered black islands on a flat white background. This interpretation may persist indefinitely but typically it reorganizes suddenly into a volumetric object (either by itself or following helpful hints from someone who already sees the object); thereafter, the image retains the volumetric interpretation, and the former percept is usually lost forever. To the human perceiver, seeing a 2-tone image as a 3D volumetric object can be a relatively effortless task, but computationally, it is a very powerful feat of perceptual organization and interpretation. Two-tone images have no exact counterpart in the natural world, yet many of these images are interpretable to human observers (and in some cases, monkeys as well; Perrett et al., 1984). Despite their non-ecological character, or perhaps because of it, 2-tone images offer numerous insights into the image features, mental representations, and perceptual processes underlying general object recognition.

The perceptual understanding of 2-tone images has typically focused only on *familiar* objects (e.g., Mooney, 1957; Galper, 1970; Phillips, 1972; Hayes, 1988; Cavanagh and Leclerc, 1989). It has been generally assumed that past experience enables interpretation (Rock, 1984), or more specifically, experience with familiar objects allows a partial match between 2-tone image contour and a stored memory representation; this partial match then influences organization of the rest of the image (Cavanagh, 1991). Object recovery from 2-tone images may require top-down guidance and thus be possible only when the image depicts familiar objects.

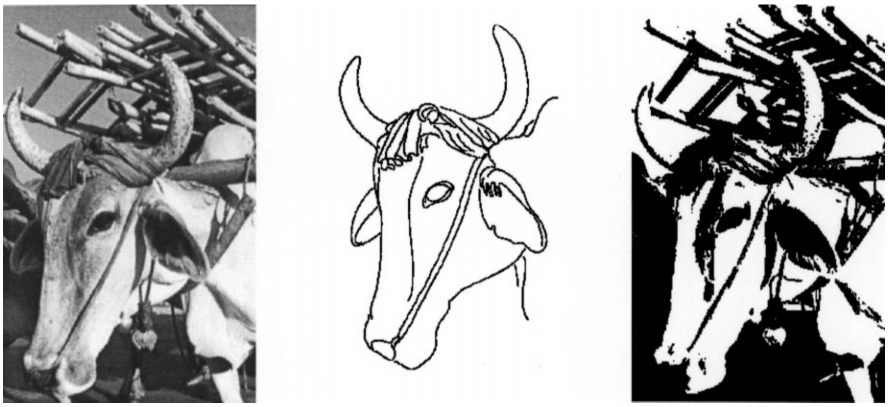

Fig. 1. Grayscale, line drawing and 2-tone image of a bullock. In the grayscale image most material changes (e.g., bridle to face) correspond to high contrast luminance contours; shadow contours tend to be softer. The drawing contains lines only for structural features (e.g., convexities and concavities at the ear and nostril, folds and creases in the rag on horns). Luminance contours in the 2-tone image occur at the occluding edges of objects (bridle strap), cast shadow edges (ear shadow on neck), and at attached shadow boundaries (right horn). These three contour types are physically undifferentiable in 2-tone images.

However, the recovery of *unfamiliar* objects from 2-tone images, and the bottom-up schemes which might enable such recovery, have never been directly tested. Object recovery from 2-tone images might use a bottom-up process in which essential features (e.g., contours, junctions, volumetric primitives) are first extracted from the image, then concatenated to form objects or object parts, which are only then compared to object representations in memory. If this were the case, a memory representation of the depicted object would not be necessary for the accurate recovery of at least the parts and structure of that object from a 2-tone image. However, 2-tone images present a daunting challenge to current models of bottom-up object shape recovery because the images do not appear to have any of the critical features on which these models rely, and yet they are recognizable. Using 2-tone images of novel objects, we test the explanatory power of three part-based approaches to object recovery.

We will demonstrate that accurate volumetric representations are rarely extracted from 2-tone images of novel objects, even when these objects are constructed from volumetric primitives such as generalized cones (Marr and Nishihara, 1978; Biederman, 1985), or from the rearranged components of a familiar object which is itself recognizable as a 2-tone image. This failure of volume recovery is not simply due to a lack of object contour and junction information. We will show that partially occluded line-drawings of the same objects, with exactly the same contour and junction features hidden, are readily seen as 3D. The failure appears to hinge on the interpretation of shadow areas. Shadow regions are typically mistaken for object surfaces, indicating observers do not have appropriate models of the lighting and shadow in the scene. Accordingly, we attempted to facilitate recovery of novel objects by making the nature of the lighting explicit in the 2-tone image. We placed the novel object beside a familiar, recognizable image where both had the same lighting; we presented several versions of the image with different light sources; we provided direct evidence of the location and direction of the lighting. None of these manipulations improved recovery. A prior memory representation is necessary for the recovery of volume from 2-tone images. Even then, only familiar objects of sufficient complexity were recovered in 2-tone images. Despite their high degree of familiarity, simple canonical volumes such as cubes or cylinders presented alone seldom appeared volumetric in 2-tone images. This failure of familiarity was perhaps not surprising given that the 2-tone images of cubes and cylinders offered so few cues and allowed so many alternative interpretations.

2. Characterization of the image and the problem

Two-tone images are created by thresholding a grayscale image; all pixels above a particular luminance level are set to white, and those below are set to black. Modeling (shading), texture and color information, features which commonly facilitate determination of 3D object shape and shadow identification, are absent in 2-tone images. There is no magic number (e.g., the average image luminance) for setting the threshold in an image. The point of maximum interpretability of each 2-tone image depends upon the lighting, the shape of the object surfaces, and the camera viewpoint.

Two-tone image interpretation presents a complex problem because the images conflate information about object structure and illumination. Luminance contour in a 2-tone image is a combination of: informative occluding object contour (Fig. 2a,b), potentially informative interior (e.g., attached shadow, pigment) contour (Fig. 2c), and potentially misleading cast shadow contour (Fig. 2d,e). The interior attached shadow contours, and the cast shadow contours, can vary dramatically with changes in illumination direction. The direction of illumination striking a curved object surface determines the contour along which the incident light is orthogonal to the surface normal (Fig. 2c). At this contour, the pictured surface in a 2-tone image turns from white (lit) to black (unlit), forming an attached shadow contour that does not correspond to the occluding contour of the object. If the direction of illumination is changed relative to the object surface, the location and shape of this contour also changes. In contrast, occluding object contours, and the attached shadow contours corresponding to occluding edges of the object (Fig. 2b), are relatively stable. Once the image has been thresholded, however, object and shadow contours are physically indistinguishable.

Surface segmentation in a 2-tone image is not simply a figure/ground differentiation in which the white region is exclusively ground and the black region is exclusively figure (or vice versa). In order to accurately recover the structure of an object, the perceiver must differentiate the contour types; mistaking shadow contour for object contour would lead to errors in calculating object structure. In addition, parts of the white region must be combined with parts of the black region to accurately recover a volumetric object from the image. Consider the 2-tone image in Fig. 2. The ends of the curved cylinder are white, but much of the rest of the object is black; these regions have to be united to form a volumetric cylinder.

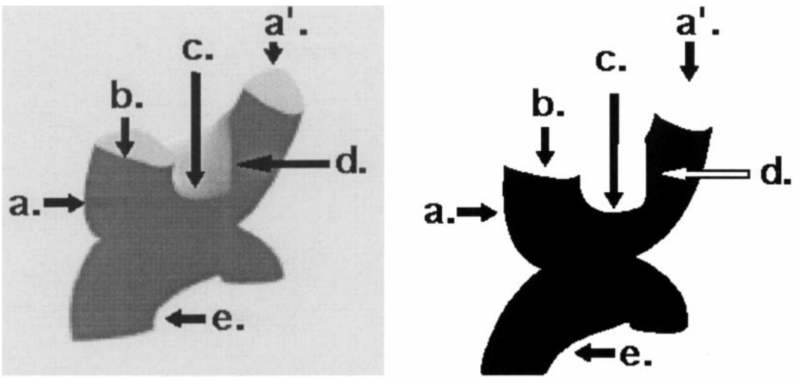

Fig. 2. Types of luminance contour: (a) external object boundary present in both the grayscale and 2-tone images, (a′) external object boundary absent in the 2-tone image, (b) attached shadow at object boundary, (c) attached shadow boundary not corresponding to object boundary, (d) self-shadow boundary, cast on the object by itself, and (e) cast shadow on ground plane. Contours (a,b) are relatively stable with changes in illumination, although (a′) illustrates that only part of the bounding contour of an object is present in a 2-tone image. Contours (c–e) vary dramatically with changes in the direction of illumination.

Although the computational problem of integrating black and white regions is complex, the human perceptual system is capable of deriving sufficient information from 2-tone images for naive observers to recognize faces and other familiar objects and scenes (Street, 1931; Mooney, 1957). Numerous researchers have shown that 2-tone images of faces are readily recognized by most observers (e.g., Mooney, 1957; Galper, 1970; Phillips, 1972; Hayes, 1988; Cavanagh and Leclerc, 1989). Even though 2-tone images contain none of the traditional depth cues, the observer does not simply see a 2D pattern that could be interpreted as a face, but rather sees the face with concave and convex regions, cast and attached shadows, much as full grayscale images are perceived. In 2-tone images of novel objects, which lack the guidance of familiarity cues, it may not be possible to differentiate contours arising from illumination effects (e.g., cast shadows, highlights) and those arising from object structure (e.g., occlusions, changes in material) and this could block the recovery of depicted object. Alternatively, stable contours which are relatively invariant to the effects of illumination (e.g., occluding object contours) might be sufficient for volume recovery via bottom-up methods that do not rely on familiarity.

The research reported in this paper is an attempt to assess the viability of bottom-up approaches for recovery of 3D structure from 2-tone images of novel objects. Two such approaches will be considered: the volumetric primitives approach (Binford, 1971; Marr and Nishihara, 1978; Biederman, 1985; Lowe, 1985) in which simple volumes are extracted from the image prior to derivation of global object structure, and the line labeling approach (Clowes, 1971; Huffman, 1971; Waltz, 1975; Malik, 1987) in which the unique interpretation of the lines constrains the formation of structural features such as corners, junctions and edges, theoretically allowing a single coherent 3D object to emerge.

The hypothesis to be tested is simple: if the perceived volumetric appearance of 2-tone images is attributable to the prior recovery of low-level structure such as corners, edges, volumetric primitives, or familiar parts, then novel single and multiple-part objects should appear volumetric as readily as familiar objects. Empirical support for this hypothesis would indicate that a part-based explanation of 2-tone image perception is sufficient. However, if images of novel objects appeared fragmented and 2D, or the volumes depicted could not be accurately recovered, an account of 2-tone image perception mediated by object familiarity, or top-down knowledge would gain support.

3. Experiment 1: single-part objects

3.1. Stimuli

To address this question we generated a set of 2-tone images of generalized cylinders after the primitives suggested by Marr and Nishihara (1978) and by Biederman (1985). These simple volumes were uniformly gray in color with a matte surface. They rested on an infinitely large white matte ground plane that was not visible in the image, but caught shadows cast by the objects. Illumination of the objects mimicked natural sunlight as nearly as possible. The objects were illuminated from above by a single point light source set at an infinite distance from the object. Placing the point light at an infinite distance insures that light rays are parallel and shadow shape is not distorted. The angle of the light relative to the surface normal of the ground plane was approximately 45°. This scheme produced fairly realistic grayscale images. Two-tone images were formed by setting the threshold in each image to a level at which all regions of the object not receiving direct illumination turned black, and all lit regions turned white (Fig. 2). This threshold level corresponded approximately to the average illumination of the image area encompassed by the smallest square that would fit around the depicted object. A total of 12 objects, each subtending 5° of visual angle, were included in the experiment.

3.2. Methods

The perceived organization of the 2-tone images was tested in 13 naive observers and several observers in our laboratory who were generally familiar with 2-tone images, but had not seen the experimental images. Observers were told they would be viewing a series of objects, some of which might appear flat or 2D, and some which might appear volumetric or 3D. They were further instructed that some objects might appear to have pieces missing, but they were to determine the shape of the object as best they could. The 2-tone images were presented individually on a computer monitor; the display was terminated by the observer's keypress response. The task of the observer was to decide whether a gray star placed on the image was, or was not, on the surface of the depicted object (Fig. 3). The star could appear in four different image areas: (1) the black region of the object surface, (2) the black region of the cast shadow, (3) the white region of the object surface, (4) the white region of the background or ground plane (these two were contiguous and undifferentiated). A correct decision about the location of the star required the observer to differentiate object regions, both black and white, from shadow and background regions. After performing the star task, the observers were asked to provide a qualitative description of the shape of the object in the image as it would appear if the object itself were before them.

To ensure the task could be performed with grayscale images, 10 subjects performed the 'star task' on the grayscale counterparts of the 2-tone images. Every object and every star location produced at least 90% accuracy with a 200 ms

Fig. 3. The observer's task was to decide whether the gray star was, or was not, on the object surface. The star appeared (left to right) on the black object region, black shadow region, white object region, white background region. The object is a cylinder illuminated from the upper left casting a shadow across the ground plane to the right.

exposure. Thus, any observed differences in the 2-tone images should be due to the information lost in converting the image to the 2-tone format.

3.3. Results and discussion

Our own observations, and those of several naive viewers revealed that the perceptual interpretation of 2-tone images of simple, isolated volumes, even common objects like cylinders and rectangular prisms, was rarely volumetric. Examples of the 2-tone objects appear in Fig. 4; readers should first view the images in Fig. 4, then compare their interpretations to the grayscale images in Fig. 15.

The most common impediment to the accurate recovery of the depicted object was misinterpretation of cast shadow areas as part of the object. The 13 naive observers tended to see all black areas as object regions and white areas as background. When the star was on the black object region, *all* observers correctly reported it to be on the object surface in *all* images. When the star was on the black *shadow* region, however, observers correctly reported that it was *not* on the object surface only 15% of the time. In signal detection terms, the observers' hit rate was 100%, whereas their correct rejection rate was only 15%. Integration of the white region into the object was marginally more successful. If the star was on the white background, observers correctly reported it was not on the object (correct rejection) in 81% of the trials; when it was on the white portion of the object, however, only 38% of the trials prompted a correct response (hit).

Signal detection analysis allows us to characterize the observer's performance in terms of the detectability of the object versus the background (d′) and a response bias (beta). Although the detectability score for black regions, d′ = 2.04, was higher than that of the white regions, d′ = 0.57, this difference is probably exaggerated as a result of the 100% hit rate for black areas. More notable is the large difference in

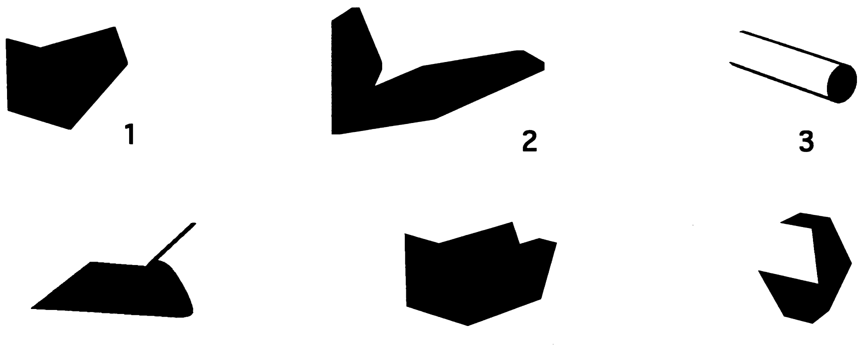

Fig. 4. Many observers fail to see volume in 2-tone images of generalized cones. Some of these single-part novel objects have cast and attached shadows (2, 4, 5), others have only attached shadows (1, 3, 6). The additional information provided by the cast shadows did not appear to be useful. For a volumetric view, see Fig. 15.

bias. The bias to see black regions as object (beta = 12.3) was far greater than the bias to see white regions as background (beta = 0.4). The first result clearly shows that observers were simply labeling black areas as figure (high bias). The second indicates that observers were willing to see white regions as part of the figure (low bias) but were doing so in a way that only weakly corresponded to the original figure areas (low d′).

Observers described the objects as either substance on a surface (e.g., paint), or curved or folded sheets of material (e.g., paper). Occasionally the images were seen as silhouettes of 3D objects whose volume did not extend beyond the dark areas. For example, object 4 in Fig. 4 can be seen as a slug facing right with an extended antenna. Notice that only the black area is seen as the slug, whereas the *actual* object surface extends into the white area.

After providing their responses, a few observers were shown the grayscale versions of the objects and asked to reconsider their interpretation of the 2-tone images. Only with considerable effort could they see volumetric objects which encompass both dark and light regions of the image. Even then, identification of the volumes was frequently incorrect. Interestingly, the 2-tone representation of some objects appears more volumetric after viewing their grayscale versions but in others the perception of volume is effortful even after exposure to a grayscale version.

Even knowing the 2-tone images are pictures of objects with shadows does not necessarily enable the viewer to accurately distinguish surface from shadow regions of the image. In addition, the absence of volumetric junction information (due to thresholding) favors an image interpretation in which the entire black region is seen as a silhouette-like object against a white background. Incorporation of portions of the white area into the hypothesized object is necessary for accurate volume recovery, but this rarely occurred. Removal of cast shadows, so the black parts of the image correspond *only* to non-illuminated object surfaces, does not seem to increase the volumetric appearance of the depicted object, but rather reduces the perceived complexity of the percept (compare objects 1 and 5, or 2 and 6 in Fig. 4).

4. Experiment 2: multiple-part objects

Considering the complexity of objects that *are* recognizable in 2-tone images (e.g., faces) objects with multiple parts might be perceived as volumetric more readily than single-part objects. Shadows cast by some object parts fall on other parts (self-shadows), possibly revealing the shape of the shadow-receiving parts.

4.1. Stimuli and methods

To test this possibility we constructed 12 novel, multiple-part objects containing 3 or 4 of the generalized cones used in the previous 2-tone images (Fig. 5). The multiple-part objects were illuminated from above (the northern hemisphere of a viewing sphere) in each of the 4 quadrants. This lighting scheme caused some images of a particular object to contain many self-shadows, whereas shadows in

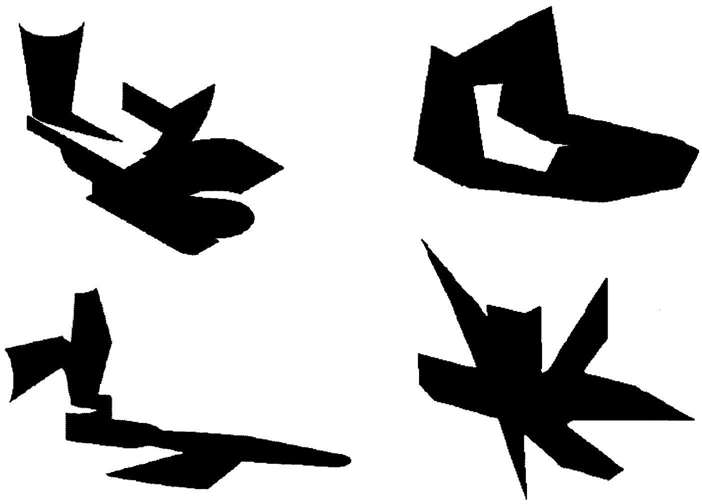

Fig. 5. The 2-tone images of objects constructed from generalized cylinders often appear to be complex paper cut-outs or silhouettes. Readers should compare their impression of volume in these images to grayscale versions of the images in Fig. 16.

other images were cast primarily on the ground plane. The task and instructions to the subjects were the same as those in the single-part object experiment.

4.2. Results and discussion

Once again, 13 naive observers displayed a bias to identify black regions as figure and white regions as ground. Black shadow regions were interpreted as non-object regions only 14% of the time; black object regions were identified as such on 91% of trials. White background regions were correctly identified as non-object regions 76% of the time, whereas white object regions were correctly identified as object surfaces in 36% of trials. Signal detection analysis again showed that observers were biased to see the black region as object (beta = 7.7) and the white region as non-object (beta = 0.43). In both cases responses were only weakly governed by the original figure and ground areas ($d' = 0.26$ for black and 0.35 for white image regions).

The objects with few self-shadows appeared to be complex paper cut-outs or folded or curved sheets of material. Perceived depth in the scene was less prevalent than in the single-part object images, and volumetric interpretations were partial at best. Some image regions were identified as shadow rather than surface regions (e.g., Fig. 5, rightmost part of the right object), but the differentiation was frequently incomplete (where does the object end and the shadow begin?) and the subsequent attempt to describe the shape of the depicted object was usually inaccurate. (Readers should compare their interpretation of the 2-tone images in Fig. 5 with the grayscale versions of the objects in Fig. 16). Self-shadows create luminance contours in the 2-

tone images that do not correspond to the edges of the object. These contours tended to obscure, rather than reveal, the shape of the object.

In general, the objects composed of several volumes did not appear more volumetric in 2-tone images than the single-part objects. That is, the ability of observers to differentiate the object from the background (as indicated by d′) was equally poor in both experiments. The only exception was the moderate value of d′ for the black regions of simple objects. But in this case, *all* black regions were labeled as 'on the object' and this 100% hit rate degrades the validity of the estimate of d′. Clearly, even with the single-part objects, observers were doing a poor job of distinguishing object from background as they always saw black regions as figure whether or not it corresponded to the original object or to its cast shadow.

4.3. General discussion

On the basis of the observations described above, the geometrical structure in 2-tone images of unfamiliar objects appears insufficient to produce the impression of 3D typical of 2-tone images of faces and other familiar objects. Object contour and shadow contour are not readily distinguished; neither simple volumes nor multiple-part novel objects are accurately recovered from 2-tone images. These results suggests that interpretable 2-tone images may require depiction of familiar objects, and that the interpretation may be mediated, top-down, by the perceiver's knowledge of specific objects.

Before accepting this explanation, however, two alternatives must be addressed.

First, the difference between comprehensible and incoherent 2-tone images might still be based upon the complexity of the object depicted, not its familiarity. Our multiple-part objects contained fewer parts, and were less compact than the typical 2-tone face. If only complex objects appear volumetric in 2-tone images, multiple-part novel objects made of 'face parts' should appear volumetric. If objects made of face parts do *not* appear volumetric, whereas the same parts in a face-like structure *do*, a familiarity explanation for 2-tone images would be supported.

Second, the amount of visible object contour in our images may have been too meager to allow recovery of 3D volume. It might be argued that a greater amount of object contour is available in traditional 2-tone images, e.g., in faces, than in our generalized cone objects, thus making recovery of volume in our objects problematic. If this were the case, line-drawings of the objects in which an equivalent amount of contour was deleted or occluded should appear as 2D as their corresponding 2-tone images. However, if partial line-drawings *did* appear volumetric, the misperception of the 2-tone images could not be attributed to lack of object contour or junction information.

We explore these two possibilities in turn in the following sections.

5. Object complexity – a demonstration

In an effort to create a novel object as 'face-like' as possible, we constructed a face (and several other objects, see Fig. 6) composed of several simple volumes, then rearranged those volumes. Both the face and the scrambled face were illuminated with the same lighting parameters, and submitted to a thresholding procedure to produce 2-tone images. As expected, the 2-tone image of the face looked quite volumetric, and was easily identified as a face. Although the 2-tone image of the scrambled face contained more luminance contour than the novel objects used earlier, it appeared either flat or silhouette-like, much as the single- and multiple-part objects.

From these observations we conclude that object complexity cannot explain why 2-tone images of faces and familiar objects are perceived veridically, whereas simple volumes are not. Furthermore, we conclude that recognition of familiar objects in 2-tone images cannot be mediated by bottom-up recovery of generalized

Fig. 6. The volumetric parts of the familiar object on the left were rearranged to create the novel object on the right. The surface of the face is mostly white, the espresso pot mostly black, but in both cases, the perceived volume incorporates both light and dark image regions. Most perceivers are unable to combine white and black regions in the objects on the left, resulting in non-volumetric or inaccurate representations. Grayscale versions of these images appear in Fig. 17.

cones, nor can the schemes that use them (Binford, 1971; Marr and Nishihara, 1978; Biederman, 1985; Lowe, 1985) be directly extended to explain the perception of volume in 2-tone images.

Although the familiar objects appeared appropriately volumetric, neither the objects made of their component parts, nor the isolated components, appeared volumetric. If generalized cones were the basic units of object recognition for 2-tone images, and were identified prior to their structural relations in the object as a whole (as hypothesized for line-drawings or full color images by Biederman, 1985), then we would expect the primitives to appear volumetric in isolation (Fig. 4), and in novel configurations (Fig. 6, right side), as well as in familiar objects (Fig. 6, left side). This was not the case. The perceptual interpretation of the generalized cones was dependent upon the familiarity of the configuration in which it appeared – the part itself was not ‘primary’.

The ineffectiveness of volumetric primitives for 2-tone image interpretation seems to arise from a difficulty in image segmentation. Typically, luminance contour is used to derive the axis of the part (Blum, 1973), or deep concavities in the occluding contour of an object delineate its parts (Hoffman and Richards, 1985). Both part-recovery methods encounter problems in 2-tone images. Deriving part or object axes requires prior, successful discrimination of object and shadow regions of the image, otherwise axes are assigned to shadow regions as well as object regions. Concavities that could indicate appropriate part boundaries in line-drawings or grayscale images are obscured in 2-tone images and spurious concavities are introduced where luminance contours created by object edges meet contours created by shadows (Fig. 5b). The known means of volumetric part recovery do not seem applicable, without substantial modification, to 2-tone images.

6. Experiment 3: adequate object contour

It is possible that the object contour and junction information in 2-tone images of novel objects is either insufficient, or too ambiguous to produce a volumetric interpretation of the image. In line labeling schemes proposed by Waltz and others (Clowes, 1971; Huffman, 1971; Waltz, 1975; Malik, 1987) the constraint imposed by the unique interpretation of luminance edges, and subsequent structural features such as corners and junctions, allows only a single coherent 3D scene to emerge.

Thresholding a grayscale image to create a 2-tone obscures many of these key features for object recovery (Fig. 7b). Segments of occluding contour are missing or obscured by cast shadows; most interior edge and corner segments are eliminated as well. The loss of these contours and their defining labels changes the identity of the remaining visible edges, especially at intersections. There are no ‘arrow’, ‘Y’, or ‘K’ junctions to indicate corners in 2-tone images. All junctions consisting of three or more coterminating lines either become ‘L’ junctions in which only two contours coterminate, or become straight lines in which the contour of the object and that of the shadow are smoothly joined. Interposition information, mediated by T junctions in line-drawings, is lost in a 2-tone image. The alteration of junctions destroys what

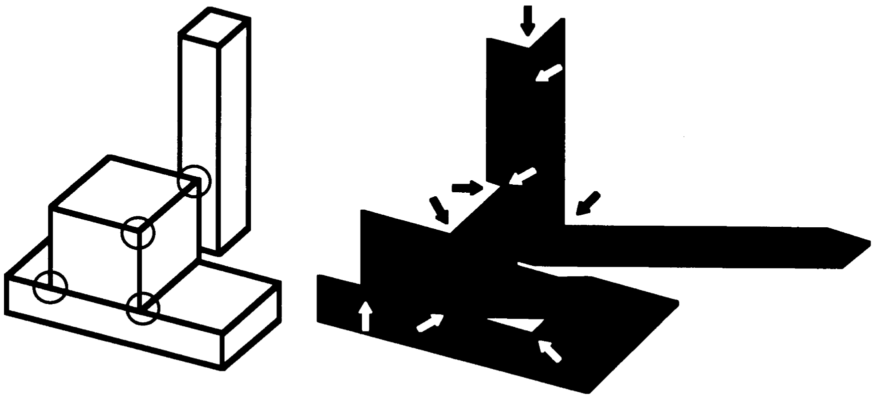

Fig. 7. Blocks world after Waltz (1975), (a) line-drawings with T, K, Y, and arrow junctions circled, (b) in a 2-tone image all 3-way junctions become L junctions or are obscured completely. The white arrows mark examples of obscured lines at junctions. Black arrows mark examples of spurious concavities which would irretrievably derail any attempt to accurately parse the image.

might have been a unique interpretation of the contours, rendering this type of representation intractable for Waltz-type line-labeling algorithms, and possibly interfering with volume recovery in human observers.

6.1. Stimuli

Testing the possibly deleterious effect of junction disruption required the dissociation of object contour and junction information from shadow or illumination information. This was accomplished by creating line-drawings of the single- and multiple-part objects used in the original 2-tone images. Segments of object contour corresponding to the segments obscured by thresholding the 2-tone images were deleted in one condition and occluded in the other. Thus the effects of contour and junction information were dissociated from illumination information.

6.2. Methods

Ten observers were shown five single-part and five multiple-part objects; each object was seen with contour occluded[1] and contour deleted (Figs. 9 and 10). Observer instruction began with the presentation of several examples of deleted and occluded line-drawings (Fig. 8). Observers were asked to form an impression of whether the drawings looked like a volume, a folded or curved sheet, or a wire. Examples of the drawings as volumes (line-drawings of whole objects), as wires

[1]The occluders were randomly shaped black blobs giving the appearance of paint on a transparent surface in front of the object. Many shapes of occluders were created with no observed difference in effectiveness. The properties of the various occluders (e.g., simplicity/complexity, symmetry/asymmetry, curvature/rectilinearity, filled/outline) did not seem to affect the perceived dimensionality of the occluded object, though unfilled occluders slowed the separability of occluder and object.

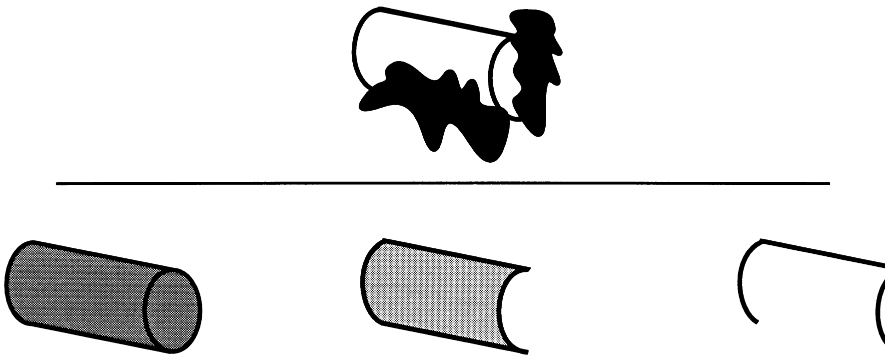

Fig. 8. Observers viewed line-drawings such as the occluded cylinder above, and were asked to state whether the depicted object looked more like a volume, a sheet of paper, or a wire. After making the decision, the observer was shown the three versions beneath the line and asked to indicate which of the three their impression was most like.

(contour deleted versions) and as surfaces (contours were closed by continuing each line end to its nearest neighbor) were then shown for comparison with observers' percepts.

During the test phase, single black line-drawings on white paper were sequentially presented to each observer. Objects were presented in a single random order with the stipulation that occluded and deleted versions of the same object were separated by at least two other objects presentations. Half the subjects received the random order, the other half saw the reverse order. The observer viewed each drawing, then indicated whether the sample drawing appeared more like a volume, a surface, or a wire. After a response was taken, observers were presented with the volumetric, surface, and wire versions of the object to ascertain how nearly their percepts corresponded to the depicted versions.

6.3. Results

When contour was *deleted* from line-drawings of single- and multiple-part novel objects leaving only those contours available in their 2-tone counterparts, the drawings did not appear volumetric (Fig. 9). Novel objects appeared volumetric in only 2% of trials, and as wires or surfaces in 98% of trials. The bare line terminations discouraged completion of the lines across missing sections, and the loss of a contour in Y or arrow junctions caused them to be interpreted as L junctions. Without these completions, the original volumetric shape was not recovered.

However, contour deletion in *familiar* objects, the cylinder and cube, was not nearly as deleterious. The cube and cylinder were interpreted as volumes in 56% of the deleted-contour trials, and as wires or surfaces in 44%. Contour completion in these familiar objects corroborates previous findings (Biederman, 1985; Biederman and Cooper, 1991; but see also Bregman, 1981). Our study of the *novel* objects suggests that completion of contour deleted volumes may require familiarity with the depicted object as well as local cues to contour relatedness.

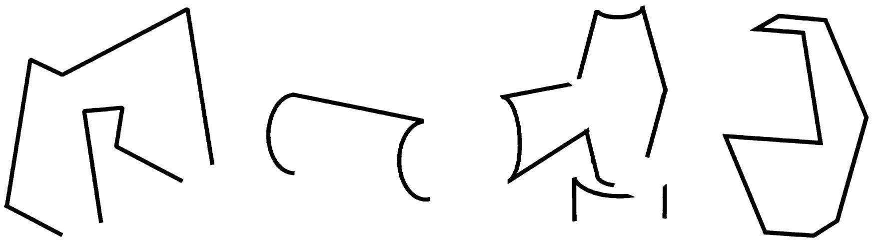

Fig. 9. Line-drawings of single- and multiple-part objects isolate the contribution of contour information from illumination and surface information. Deleted contour segments correspond to obscured contour segments in the 2-tone counterparts of these objects. Like their 2-tone versions, these contour versions fail to appear volumetric.

When occluders covered the same lengths of contour that had been deleted (or obscured in the 2-tone images) 3D shape was accurately recovered in almost all instances. A ‘volume’ response was produced in 90% of all trials, whereas only 10% produced a ‘surface’ or ‘wire’ response. (Recall that *none* of these appeared volumetric in 2-tone images!) The strength of the volumetric appearance of the object seemed to depend upon the amount of contour occluded (Fig. 10) rather than the type of image feature hidden or revealed by the occluder. Single- and multiple-part objects did not differ in frequency of appearing volumetric, however, the simpler, more canonical objects (e.g., the cylinder) were usually perceived with greater accuracy than the more complex or unusual objects. The more unusual objects were seen as volumes, but the accuracy of the hypothesized structure is reduced for some observers. The volumes the observers described were consistent with the partially occluded or deleted drawings, but occasional simplification of the volumes did occur. For example, the leftmost object in Fig. 10 was usually described as having a flat top.

6.4. Discussion

Although one or more lines in every Y, K or arrow junction is occluded, the junctions complete spontaneously and perception of volume is immediate. The

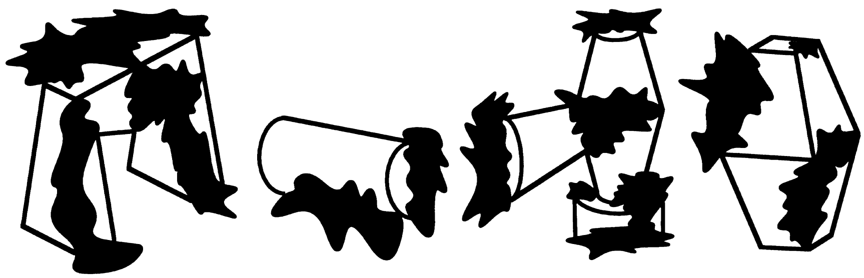

Fig. 10. Line-drawings with occlusion of the segments deleted in Fig. 8 and obscured in the 2-tone images. Unlike the latter two versions, the occluded objects appear volumetric. Notice that occlusions occur at junctions, not in the middle of line segments.

completing junctions in the line-drawings contain the same contour as those that *did not* complete in the 2-tone images. The Y, K, or arrow junctions need not be explicit for the line-drawings to be interpreted as 3D objects; the lines appear to continue behind the occluder completing the implicit junction. These results imply that there *is* sufficient contour and junction information in 2-tone images to form a volumetric interpretation. Given the same contour and junction information appears in the different forms of representation, why would a volumetric interpretation succeed with the occluded line-drawings, but not with the 2-tone images?

This result exemplifies the ‘generic’ or ‘non-accidental’ nature of the world model used by the human visual system (e.g., Lowe and Binford, 1981; Witkin and Tenenbaum, 1983; Lowe, 1985; Koenderink, 1990; Freeman, 1994; Albert and Hoffman, 1995). Rather than assuming the image (either retinal or pictorial) resulted from unusual viewing conditions, the visual system assumes the image contains a generic or non-accidental, and therefore representative, view of objects in the world. Non-accidental interpretations are made on the basis of a few image features (e.g., T junctions, parallelism, symmetry) that are reliable indicators of the structure and location of objects in the 3D world. Their recovery may occur in the early stages of visual processing, suggesting the possibility of bottom-up image reconstruction.

In the case of the occluded objects, T junctions are formed at the intersection of the occluder and the bounding edge of the object, providing a simple (possibly bottom-up) mechanism for separating the occluder and the object (Fig. 10). The contour of the object appears to terminate extrinsically at the junction with the occluder (Nakayama et al., 1987). Even though the placement of the occluder itself is highly accidental, the resultant T junctions separate the occluder from the line drawing, implying that if the viewpoint of the observer were to change, more object contour might be revealed. Thus the presence of the T junctions provides an impetus for separating the object from the occluder, and the presence of the occluder provides support for postulating continuation of the object contour behind the occluder (Bregman, 1990; Fig. 11).

In contrast, the deleted-contour drawings provide no evidence that the object lines might continue into the solid white background, or that further information could be revealed by a change of viewpoint. Observers frequently see the lines of the object as

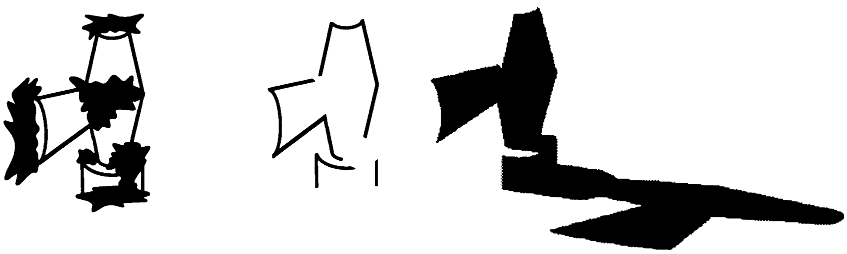

Fig. 11. The presence of the occluder suggests the object contour could continue on to form a volumetric object; the line terminates extrinsically at the junction with the occluder. In contrast, deleted line-drawings and 2-tone images provide no cues for continuation of their contours or surfaces into the white region. Consequently, no volumetric interpretation is formed.

intrinsically terminated (Nakayama et al., 1987) at the point of deletion. Similarly, the black regions in the 2-tone images are usually seen as surfaces which also appear to be intrinsically bounded. The image provides no obvious reason for postulating continuation of the surface beyond the luminance contour. The black regions occasionally appear as 3D silhouettes in which the entire black region (including the shadow) constitutes the object projecting the silhouette, but again the hypothesized object does not appear to encompass any of the white region. Since there are no differences between luminance contours caused by shadows and those caused by object edges, both are interpreted as surface edges, and both contribute to the perceived shape of the surface.

7. Illumination hypotheses

Although the available object contour is identical in the contour-occluded line-drawings and 2-tone images, the percept is very different. The compelling volume in the occluded line-drawings implies that ample contour is present, but contour alone is insufficient to enable a volumetric interpretation of novel objects in 2-tone images. What seems to be missing is a model of light and shadow in the scene. Since line-drawings do not contain illumination information, a lighting model is unnecessary, but it may be required to understand 2-tone images. If an object surface in a 2-tone image is to continue beyond a black region, across a luminance boundary, into a white region, there must be an external (non-object based) explanation of the black/white boundary. In the 2-tone images, shadows, caused by the particular illumination in the scene are the explanation.

The observer must generate a representation which includes a model of light and shadow for these borders to be attributed to illumination effects and not to object structure. If the object in the 2-tone image is a familiar one, a partial contour match may trigger, or even compel a lighting and shadow hypothesis. For example, some contours of the image may resemble contours of a face, but the whole image could only be a face if strong directional lighting and deep shadows were also present. If our conjecture is correct, then global familiarity is one means of triggering an appropriate lighting hypothesis. However, other means might be available; in the following sections we examine some alternative possibilities. Perhaps, if given the illumination direction, observers could complete surfaces in 2-tone images of novel objects just as they completed line-drawings behind the occluders.

7.1. Illumination information from recognizable objects

Previous research has shown that illumination direction is instrumental in determining the 3D shape of ambiguous *grayscale* objects that could be interpreted as bumps or dents (Yonas et al., 1979; Berbaum et al., 1983a; Berbaum et al., 1983b). Yonas et al. found that illumination direction determined perceived shape for adult observers. Berbaum et al. used the presence of a familiar object as implicit evidence of a particular illumination direction and found that observers incorporated this

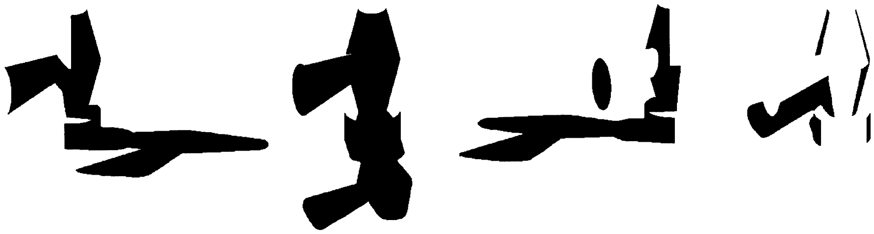

Fig. 12. A multiple-part object illuminated from each of 4 quadrants of the northern hemisphere; the angle of the light is approximately 45°. The object contour *could* be computed from the information present, but it is unlikely to alter the how the observer perceives the object in the image.

information into their interpretations of the ambiguous objects. The illumination direction suggested by the unambiguous shading of the familiar object was extended to the ambiguous object, determining whether it will appear to be a bump or a dent.

A simple test of this effect in our 2-tone images is seen in Fig. 6 which contains both a familiar and an unfamiliar object. Unlike the results of Berbaum's grayscale studies the presence of a recognizable object beside a novel object did not provide a sufficiently strong cue to enable volume recovery in 2-tone images.

7.2. Illumination information from multiple views

We next considered whether multiple views of the same object, each illuminated from a different direction, could induce a volumetric interpretation. Illumination of the same object from different quadrants[2] of the northern hemisphere (Fig. 12) produced differently shaped shadows, but the external object contours remained fairly consistent from one lighting condition to another. In each image, some previously exposed segments, and some new segments of the object contour were exposed. By viewing four images of an object, each lit from a different quadrant, it is theoretically possible to extract and integrate consistent (object) contours while discounting the changing (shadow) contours, thus discriminating the object from its shadow. Previous studies have shown that observers can integrate partially deleted contours across multiple views of line-drawings well enough to enable object naming (e.g., Snodgrass and Feenan, 1990; Biederman and Cooper, 1991; Srinivas, 1993).

Observers who are knowledgeable about the creation of 2-tone images may be able to consciously extract and integrate the available contour for constructing an accurate object representation. However, these observers, the authors included, are unlikely to 'see' a volumetric interpretation of a novel object, such as occurred in the occluded line-drawings. The 'moment of insight' that accompanies the organization of a difficult 2-tone image of a familiar object (e.g., the famous Dalmatian) never arrives. The process seems more cognitive.

[2]When images of individual objects were shown in isolation, no particular illumination quadrant (in the northern hemisphere) revealed volume or improved shadow/object contour discrimination better than any other. Illumination from either the southern hemisphere or the equator of the viewing sphere produced nearly incomprehensible images.

7.3. Explicit illumination information

In previous images, lighting direction could (theoretically) be derived by locating the shadow in the image. However, deciding which image regions are shadows and which are object surfaces may depend upon knowing the direction and location of the light source. Perhaps the illumination and contour information in the 2-tone images is too interdependent to allow derivation of both simultaneously without the aid of cues like familiarity.

In an attempt to decouple the two, we made the direction, strength, and source of illumination explicit. Two-tone images of novel objects were placed in the context of a grayscale scene depicting a lamp with a very bright spotlight. When the object is truly novel (Fig. 13, left side) explicit information about the light striking the object does little to unite black and white regions or enable perception of volume (for a grayscale view, see Fig. 16). For several observers asked to describe the depicted shape, the explicit lighting depiction creates a volumetric appearance of familiar objects like the cylinder in the image on the right side of Fig. 13.

In previous 2-tone images of novel objects, neither familiar volumes like the cylinder nor more novel volumes like those depicted in Fig. 4 (all are generalized cylinders) appeared volumetric. When illumination was specified, familiar but simple objects such as the cylinder in Fig. 13 appeared reliably volumetric. If explicit lighting could create volume in the cylinder, why was it not also seen in other objects?

Cylinders are extremely common in the world, and it would be unlikely that perceivers would store representations of more complex namable objects (e.g., faces, espresso makers), but not cylinders. However, the match of such a simple shape may fail on some 2-tone images because it is not sufficiently complex to rule out other interpretations, or because more parsimonious interpretations are readily available. For example, adopting the folded sheet interpretation for the objects in

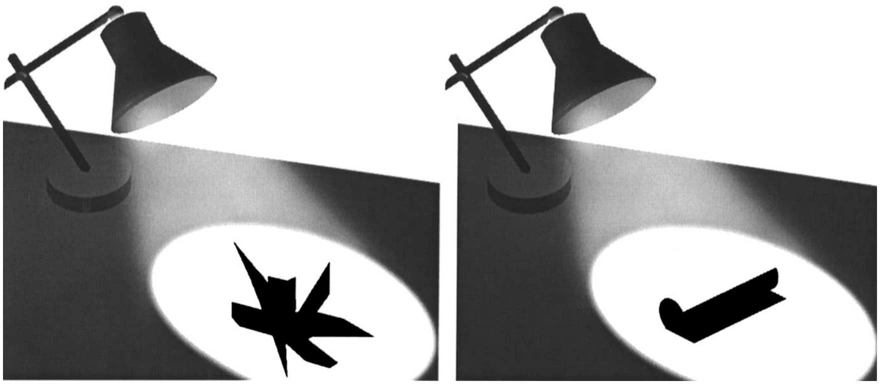

Fig. 13. Even though the illumination source is explicit, as are the lighting direction and strength, the novel object on the left (see grayscale version, Fig. 16) appears to be a folded sheet whereas the more familiar cylinder on the right appears volumetric to most observers.

Fig. 13 would allow the perceiver to ignore the illumination information in the image. Alternatively, a volumetric interpretation requires that the lamp generate a level of brightness uncharacteristic of the normal world. When an explicit lighting model is provided, the additional information may be sufficient to allow a volumetric match of the regular cylinder, but insufficient to enable a volumetric interpretation of an unfamiliar generalized cylinder. The interpretation of the novel generalized cylinder could be either a volumetric object with a complex lighting model, or a folded sheet with a generic lighting model. Both alternatives are compatible with the image information, but the latter may conform better to a generic model of the world, especially when a lighting model is not explicitly specified.

7.4. Illumination information from scenes

All the novel objects in the 2-tone images we have presented thus far appear in isolation. Although many interpretable 2-tone images contain only a single object (e.g., a face), the face may contain many redundant cues consistent with a memory-representation of a face and inconsistent other alternatives. In order to create 2-tone images with a multiplicity of cues to shadow location, we constructed scenes containing multiple novel objects, hoping the redundancy would facilitate differentiation of shadow from object surfaces.

Redundancy of shadow cues was facilitated by the spacing of the objects and the long, thin shadows from the vertical objects spanning the width of the horizontal object (and its shadow), and continuing on the ground plane. Although these shadow conditions may have enhanced the separation of the objects, the perception of a ground plane (Stevens, 1981), and the skewed symmetry relation between the object and shadow (Kennedy, 1974), they produced only a partial volumetric representation of the horizontal object on which the shadows were cast (Fig. 14, far left).

However, when familiar objects such as cylinders were depicted in the scene, they did appear to attain volume (Fig. 14, right). As with the explicit lighting model, the redundant cues in the scenes appeared sufficient to create volumetric interpretations of simple familiar objects, but less so with simple novel objects. The presence of several vertical objects in the scene, creating more redundancy, seems to enhance volume recovery. Some instances of cotermination of cast shadows created a modest

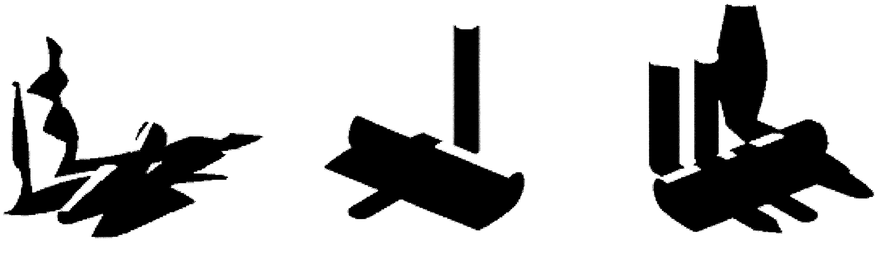

Fig. 14. Tall regular volumes, located far from one another, casting long shadows, produce a volumetric interpretation (center, right). Irregular volumes reduce the redundancy of segmentation cues and images appear less volumetric (left).

illusory contour coincident with the actual bounding contour of the horizontal object (Fig. 14, center), possibly facilitating segmentation of image regions into shadows and object surfaces, and encouraging continuation of object surfaces into the white area. Many of the organizing relations used in Fig. 14 also appear in 2-tone images of single objects (see especially Figs. 2 and 6), but in isolation they do not appear strong enough to differentiate shadow and object regions.

8. Conclusions

The observations reported here were an attempt to discover whether bottom-up, part-based approaches to object perception could be applied to the interpretation of 2-tone images of novel objects. It is well established that 2-tone images of familiar objects can be recognized, even by naive observers. We examined 2-tone images of novel objects to explore *how* objects might be recovered.

First, generalized cones were depicted in 2-tone images, both in isolation and in the context of multiple-part novel objects. If the volumetric primitives were the basic units of object recognition in 2-tone images, we would expect them to appear volumetric in isolation, in novel configurations, and in familiar objects. The parts did not appear reliably volumetric either alone or together. Next we created nonsense objects by rearranging parts that had appeared volumetric in 2-tone images of familiar objects. Only in the context of the familiar objects did the parts appear volumetric. From these results we conclude that recovery of volumetric parts does not precede, and thereby enable, object recognition in 2-tone images.

We next asked whether the amount of visible object contour in our images was sufficient to support a volumetric interpretation. To isolate the contribution of contour from that of illumination, we constructed line-drawings of each object, occluding segments of contour corresponding to segments obscured in the 2-tone images. Although both had identical contour exposed, the 2-tone images did not appear volumetric whereas the line-drawings with occluded segments did. The presence of the occluders enabled a perceptual hypothesis in which the object corners appeared to complete behind the occluders. The compelling appearance of volume in the occluded line-drawings indicated that the lack of volume in the 2-tone images cannot be attributed to insufficiency of local components such as junctions and contour.

Having established that bottom-up recovery of volumetric parts was not generally useful for 2-tone image interpretation, and that the images of novel objects did indeed contain ample object contour for volume recovery, we turned from the structural properties of objects to their illumination. When the objects were truly novel, even explicit lighting depiction did not produce volumetric representations. In a 2-tone image containing both a familiar object and a novel object, the unambiguous lighting of the familiar object failed to transfer to the novel object. The presentation of multiple 2-tone images of the same object, each illuminated from a different angle failed as well; even knowing the contour of the object did not make it look volumetric. Finally, we created a grayscale scene of a glaring spotlight shining on a 2-tone image of an object. The addition of an explicit lighting model enabled volume recovery of simple familiar objects, but neither simple nor complex novel objects. Even though the lighting strength, location and direction was explicit in the image, it did not induce a perception of volume in the novel objects.

We suggest that determination of illumination strength and direction is a necessary aspect of the perceptual understanding of 2-tone images. When a 2-tone image depicts a complex familiar object, a segment of identifiable contour can initiate the process of differentiating shadow from object regions and integrating an object surface that continues beyond a luminance boundary in the image. However, familiarity alone is not a sufficient condition for volume recovery from 2-tone images. Many of the generalized cones we used were simple, highly familiar objects, e.g., cubes and cylinders. In isolation, familiar volumes were interpreted as unusual bent or silhouetted shapes under generic lighting. The 2-tone images of these volumes are so sparse they allow multiple simple interpretations, many of which do not compel a complex lighting model with deep shadows. The possibility that the image depicts some unfamiliar object under generic lighting conditions appears preferable to the interpretation of a familiar volume with unusual lighting. Simple familiar shapes require explicit lighting or regular arrays with highly redundant shadow patterns to constrain possible interpretations and enable appropriate 3D recovery. These findings suggest that strictly bottom-up, or even part-based models are inadequate for explaining perceptual interpretation of 2-tone images.

We favor an approach in which recognition of familiar objects in 2-tone images (and perhaps natural scenes) is essentially top-down, requiring some hypothesis concerning the identity of the object depicted before the assignment of regions to figure and ground is made. This approach is similar to that of Peterson and Gibson (1994) who claim that object recognition precedes the assignments of figure and ground in simple black and white figures. We suggest that the first guess for the identity of the object is mediated by viewpoint-specific memory representations of familiar objects (Cavanagh, 1991; Poggio and Hurlbert, 1994; Tarr and Bulthoff, 1995; Shashua, 1997). This approach uses distinctive, perhaps deformable, templates (Yuille, 1991) at either the whole object level or the local part level. Recent work in computer vision (Belhumeur et al., 1996; Shashua, 1997) also demonstrates that only a few directions of illumination need to be stored in order to recognize an object illuminated from a novel direction. Although it is clear that canonical parts in our 2-tone images could not *initiate* the process of object recovery, evidence from line-drawings, explicit illumination and object redundancy in our studies suggests their presence may be beneficial once an illumination direction has been hypothesized. In natural scenes, which contain multiple segmentation cues such as T junctions, color and texture that could initiate object/shadow segmentation, part templates may play a more active role.

Object recognition in natural scenes may use a top-down match of characteristic views of familiar objects, or of familiar parts of unfamiliar objects, and image information. This template approach might be used to identify familiar parts (e.g., cylinders, bricks, or spheres) of unfamiliar objects (Dickinson et al., 1992) as well as whole namable objects. A hypothesis about a few simple volumes in probable orientations could provide a start for the structural analysis of the overall 3D shape. The end result of this recovery of familiar parts would be a set of volumetric parts and their orientations and locations very much like that provided by other part-based approaches. The difference is that the volumetric parts would be identified from stored 2D views and might be less vulnerable to the missing elements and ambiguous shadow regions which make 2-tone images so difficult. Observers may have mental models of familiar parts which, once triggered by other image cues, create expectations or predictions about where completing lines might be found, thus facilitating object recovery.

Clearly, parts are not interpretable in 2-tone images. Neither generalized cones, nor even familiar, canonical shapes like cylinders and cubes provided any reliable volumetric interpretation either in isolation or in complex groups. The recognition of sufficiently complex familiar objects appears to employ a more direct, holistic process that bypasses an initial part-based analysis.

Although our demonstrations reject part-based analyses in the interpretation of 2-tone images, they do not exclude part-based analyses in the interpretation of natural images with redundant cues to shadow and object edges. The critical message of 2-tone images is that familiar objects *can* be recognized without recourse to any known bottom-up schemes. The presence of this level of analysis for 2-tone images implies that it operates in natural scenes as well, for who would imagine that such a strategy would evolve for a style of image that did not exist before 1895. An holistic analysis for familiar objects would offer speed advantages in dealing with familiar items in the world, but undoubtedly would operate in concert with other more general, perhaps structural approaches to object recognition.

Appendix A. Figs. 15–17.

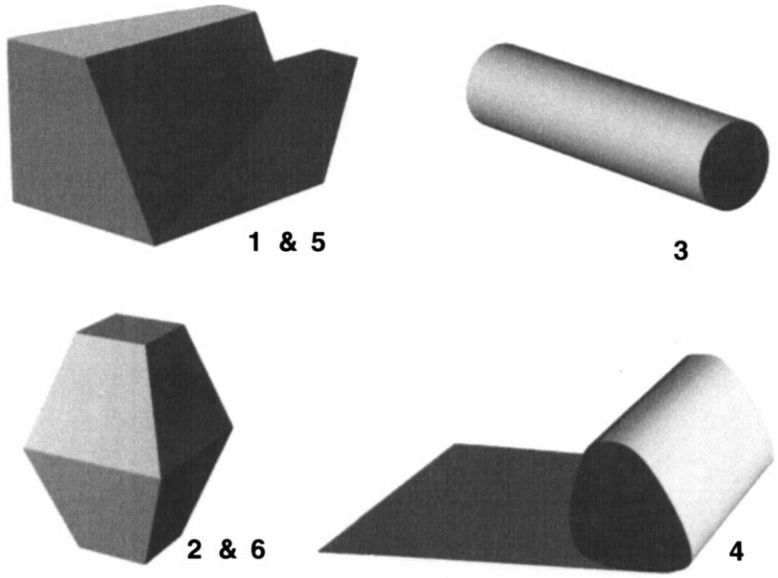

Fig. 15. Grayscale versions of 2-tone images in Fig. 4. Object numbers correspond to 2-tone versions with and without cast shadows.

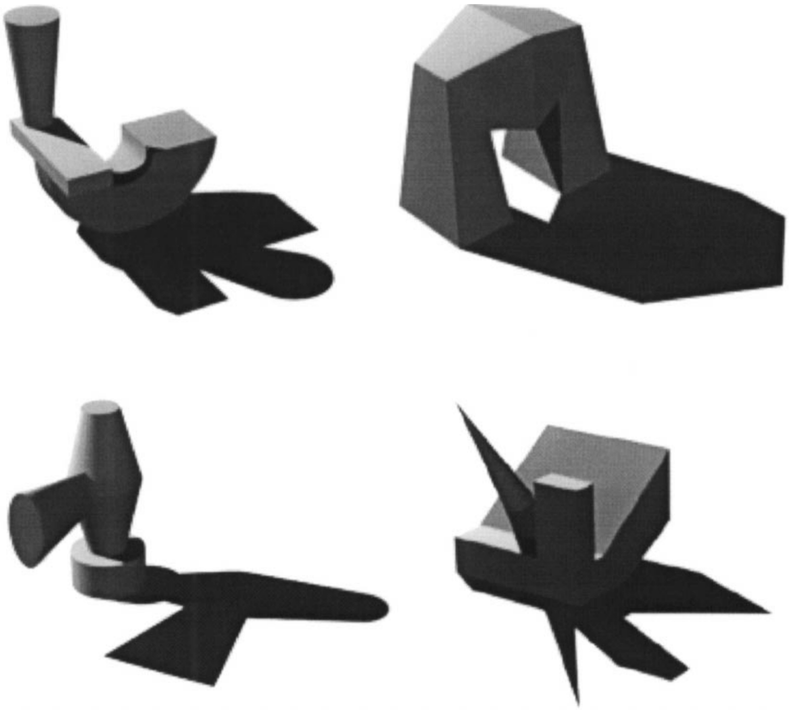

Fig. 16. Grayscale images of 2-tone objects in Fig. 5.

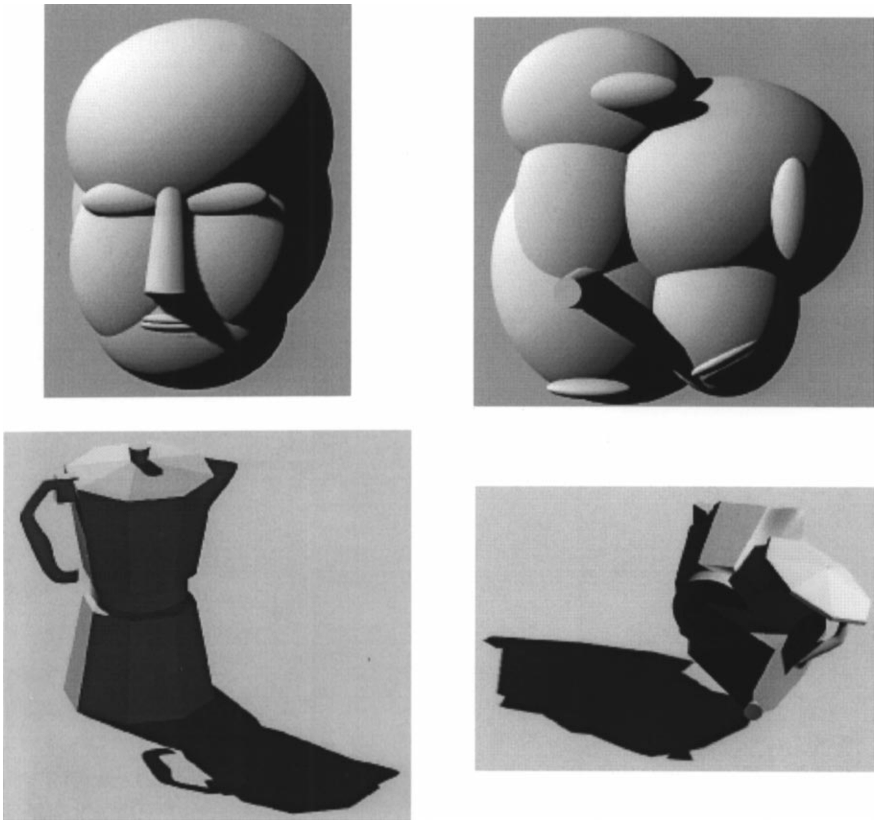

Fig. 17. Grayscale versions of 2-tone images in Fig. 6. Although the objects on the right are constructed from the rearranged parts of the objects on the left, the novel arrangements are rarely seen as volumetric.

References

Albert, M., Hoffman, D., 1995. Genericity in spatial vision. In: Luce R.D. (Ed.), Geometric Representations of Perceptual Phenomena: Papers in honor of Tarow Indow on his 70th birthday. Erlbaum, Mahwah, NJ, pp. 95–112.

Belhumeur, P.N., Yuille, A.L., Epstein, R., 1996. Learning and recognizing objects using illumination subspaces. Proceedings of the International Workshop on Object Representation for Computer Vision April, 1996.

Berbaum, K., Tharp, D., Mroczek, K., 1983a. Depth perception of surfaces in pictures: looking for conventions of depiction in Pandora’s box. Perception 1–2, 5–20.

Berbaum, K., Bever, T., Chung, C.S., 1983b. Light source position in the perception of object shape. Perception 1–2, 411–416.

Biederman, I., 1985. Human image understanding: recent research and a theory. Computer Vision, Graphics, and Image Processing 32, 29–73.

Biederman, I., Cooper, E.E., 1991. Size invariance in visual object priming. Journal of Experimental Psychology: Human Perception and Performance 18, 121–133.

Binford, T., 1971. Visual perception by computer. Proceedings, IEEE Conference on Systems Science and Cybernetics. Miami, FL.

Blum, H., 1973. Biological shape and visual science, Part 1. Journal of Theoretical Biology 38, 205–287.

Bregman, A.S., 1981. Asking the 'What For' question in auditory perception. In: Kubovy, M., Pomerantz, J. (Eds.), Perceptual Organization. Lawrence Erlbaum, Hillsdale, NJ.
Bregman, A.S., 1990. Auditory Scene Analysis. MIT Press, Cambridge, MA.
Cavanagh, P., 1991. What's up in top-down processing? In: Gorea, A. (Ed.), Representations of Vision: Trends and Tacit Assumptions in Vision Research. Cambridge University Press, Cambridge, UK.
Cavanagh, P., Leclerc, Y.G., 1989. Shape from shadows. Journal of Experimental Psychology 15, 3–27.
Clowes, M.B., 1971. On seeing things. Artificial Intelligence 1, 79–116.
Dickinson, S.J., Pentland, A.P., Rosenfeld, A., 1992. From volumes to views: an approach to 3-D object recognition. CVGIP: Image Understanding 55 (2), 130–154.
Freeman, W.T., 1994. The generic viewpoint assumption in a framework for visual perception. Nature 368 (6471), 542–545.
Galper, R.E., 1970. Recognition of faces in photographic negative. Psychonomic Sciences 194, 207–208.
Hayes, A., 1988. Identification of two-tone images; some implications for high- and low-spatial-frequency processes in human vision. Perception 174, 429–436.
Hoffman, D.D., Richards, W.A., 1985. Parts of recognition. Cognition 18, 65–96.
Huffman, D.A., 1971. Impossible objects as nonsense sentences. Machine Intelligence 5, 295–323.
Kennedy, J.M., 1974. A Psychology of Picture Perception: Images and Information. Jossey-Bass, San Francisco, CA.
Koenderink, J.J., 1990. Solid Shape. MIT Press, Cambridge, MA.
Lowe, D.G., 1985. Perceptual Organization and Visual Recognition. Kluwer Academic, Boston, MA.
Lowe, D.G., Binford, T.O., 1981. The interpretation of three-dimensional structure from image curves. Proceedings of IJCAI-7 Vancouver, August, 613–618.
Malik, J., 1987. Interpreting line drawings of curved objects. International Journal of Computer Vision 1, 73–107.
Marr, D., Nishihara, H.K., 1978. Representation and recognition of the spatial organization of three-dimensional shapes. Proceedings of the Royal Society London 200, 269–294.
Mooney, C.M., 1957. Age in the development of closure ability in children. Canadian Journal of Psychology 114, 219–226.
Nakayama, K., Shimojo, S., Silverman, G.H., 1987. Stereoscopic occluding contours: a critical role in pattern recognition of background objects. Investigative Ophthalmology and Visual Science Supplement 28, 365.
Perrett, D.I., Smith, P.A., Potter, D.D., Mistlin, A.J., Head, A.S., Milner, A.D., 1984. Neurons responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity, and relation to perception. Human Neurobiology 3, 197–208.
Peterson, M.A., Gibson, B.S., 1994. Must figure-ground organization precede object recognition? An assumption in peril. Psychological Science 5, 253–259.
Phillips, R.J., 1972. Why are faces hard to recognise in photographic negative? Perception and Psychophysics 12, 425–426.
Poggio, T.A., Hurlbert, A., 1994. Observations on cortical mechanisms for object recognition and learning. In: Koch, C., Davis J.L. (Eds.), Large-Scale Neuronal Theories of the Brain. MIT Press, Cambridge, MA.
Rock, I., 1984. Perception. W.H. Freeman, New York.
Shashua, A., 1997. On photometric issues in 3D visual recognition from a single 2D image. International Journal on Computer Vision 21 (1–2), 99–122.
Snodgrass, J.G., Feenan, K., 1990. Priming effects in picture fragment completion: support for the perceptual closure hypothesis. Journal of Experimental Psychology: General 119 (3), 276–296.
Srinivas, K., 1993. Perceptual specificity in nonverbal priming. Journal of Experimental Psychology: Learning Memory and Cognition 19 (3), 582–602.
Stevens, K.A., 1981. The visual interpretation of surface contours. Artificial Intelligence 17, 47–74.
Street, R.F., 1931. A Gestalt completion test. PhD Thesis, Teachers College at Columbia University, New York.
Tarr, M.J., Bulthoff, H.H., 1995. Is human object recognition better described by geon structural descrip-

tions or by multiple views? Comment on Biederman and Gerhardstein (1993). Journal of Experimental Psychology: Human Perception and Performance 6, 1494–1505.
Waltz, D., 1975. Understanding line drawings of scenes with shadows. In: Winston, P.H. (Ed.), The Psychology of Computer Vision. McGraw-Hill, New York.
Witkin, A.P., Tenenbaum, J.M., 1983. On the role of structure in vision. In: Beck, J., Hope, B., Rosenfeld, A. (Eds.), Human and Machine Vision. Academic Press, New York.
Yonas, A., Kuskowski, M.A., Sternfels, S., 1979. The role of frames of reference in the development of responsiveness to shading information. Child Development 50, 495–500.
Yuille, A.L., 1991. Deformable templates for face recognition. Journal of Cognitive Neuroscience 31, 59–70.