6 Genericity in Spatial Vision

Marc K. Albert
Department of Information and Computer Science, University of California

Donald D. Hoffman
Department of Cognitive Science, University of California

ABSTRACT

One principle the brain uses to construct spatial interpretations of retinal images is *genericity.* We describe this principle and illustrate its operation in our perceptions of line drawings, object parts, and subjective surfaces.

INTRODUCTION

Your visual system reports on your environment: its objects, shapes, colors, motions, and spatial layout. You might expect from this report the same objectivity you expect from the local *Times.* A good reporter, you know, never creates news, but just reports it. Opinions and speculations get labeled as such and quarantined to their own section. The front page reports objective facts, free of reporter biases.

You might expect this objectivity from vision, but you will not get it. What you get instead resembles an opening statement from the local DA: a carefully constructed story, part fact, part supposition, clearly biased, sometimes downplaying or ignoring evidence to the contrary.

Why? Your visual system needs to tell a three-dimensional (3D) story about objects, shapes, colors, and motions. The only evidence it has to construct this story are photon catches at receptors laid out in a two-dimensional (2D) mosaic on the retina. The gap between the evidence given and the story to be constructed is enormous, as anyone will testify who has tried to build a working machine vision system: the evidence at the retina is logically compatible with innumerable different stories. Careful detective work is required to bridge the gap.

The problem is so difficult your brain devotes roughly 10 billion neurons to it. Wired into these neurons are many procedures and biases for story construction.

We discuss one of them. Visual psychologists and researchers in computer vision sometimes call it the principle of "genericity" or "nonaccidentalness" (Biederman, 1985; Binford, 1981; Hoffman & Richards, 1984; Koenderink, 1990; Lowe, 1985; Lowe & Binford, 1981; Ullman, 1979; Witkin & Tenenbaum, 1983). We discuss how this principle shapes our perceptions of line drawings, object parts, and subjective surfaces.

GENERICITY AND LINE DRAWINGS

The principle of genericity, in its simplest form, says to reject any 3D interpretation of the retinal image that would place the eye in an "unstable" viewing position. One way to define an unstable viewing position is as follows: it is a viewing position which, if perturbed slightly, would lead to a change in the topological or first order differential structure of the image. Two examples will help.

First a topological case. Suppose the image contains an L junction, i.e., two line segments which meet at a vertex as in the letter *L*. Consider a 3D interpretation consisting of two disconnected line segments at different depths in space. Under this 3D interpretation the reason the image contains an L junction, instead of two separated line segments, must be that your eye is viewing the line segments from a special vantage which makes their endpoints look connected. If you were to move your eye slightly, the image of the L junction would become an image of two separate line segments. (This separation actually occurs in the familiar "Ames chair illusion," in which a set of disconnected sticks in space look like a chair from only one special viewpoint, and otherwise appear to be disconnected. See Kilpatrick, 1952.) This introduction of a gap is a topological change in the structure of the image. Therefore the principle of genericity says to reject this 3D interpretation.

Now a first-order differential example. Suppose the image contains a single line segment. Consider a 3D interpretation consisting of two line segments in space meeting to form a right angle. Under this 3D interpretation the reason the image contains a single line segment, instead of two line segments meeting to form an L junction, must be that your eye is viewing the right angle from a special vantage which hides the vertex. If you were to move your eye slightly, the image of the line would become an image of two lines meeting at an L junction. This introduction of a tangent discontinuity is a change in the first-order differential structure of the image. Therefore the principle of genericity says to reject this 3D interpretation.

The genericity principle has been used by various researchers to justify rules for interpreting images (see, e.g., Lowe, 1985). Some examples are the following:

FIG. 6.1. Th
described by
er in 1832.

Rule 1. Points collinear in

Rule 2. Points smoothly c
the world.

Rule 3. Points symmetric i

Rule 4. Curves terminatin
common point in t

Rule 5. Three or more cu
intersect in a com

These rules greatly constra
Consider, for instance, the Ne
segments in this figure must b
Rules 1 and 2. The reasoning is
in a line segment are smoothl
connected in space. And accor
are collinear in the image they
form a line in space.

One can cast this reasoning
er the conditional probability
actually arose from the project
as P(wiggly in world|S in imag
$P(W \mid S)$ is zero. By Bayes' r

$$P(W$$

Here $P(W)$ and $P(S)$ are the pr
space and of the straight line
estimate $P(S)$. We can estimat
to a value which normalizes t
first factor in the numerator,

hly 10 billion neurons to it.
ases for story construction.
ıd researchers in computer
:y" or "nonaccidentalness"
:hards, 1984; Koenderink,
n, 1979; Witkin & Tenen-
es our perceptions of line

/VINGS

to reject any 3D interpreta-
: in an "unstable" viewing
›sition is as follows: it is a
ld lead to a change in the
image. Two examples will

ins an L junction, i.e., two
,. Consider a 3D interpreta-
t different depths in space.
ıtains an L junction, instead
·ye is viewing the line seg-
›ints look connected. If you
junction would become an
tion actually occurs in the
isconnected sticks in space
ınd otherwise appear to be
›n of a gap is a topological
inciple of genericity says to

:he image contains a single
ıg of two line segments in
nterpretation the reason the
› line segments meeting to
ing the right angle from a
move your eye slightly, the
:s meeting at an L junction.
:e in the first-order differen-
genericity says to reject this

researchers to justify rules
ıe examples are the follow-

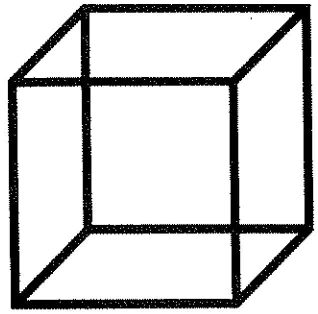


FIG. 6.1. The Necker cube first described by Louis Albert Necker in 1832.

Rule 1. Points collinear in an image are collinear in the world.

Rule 2. Points smoothly connected in an image are smoothly connected in the world.

Rule 3. Points symmetric in an image are symmetric in the world.

Rule 4. Curves terminating at a common point in an image terminate at a common point in the world.

Rule 5. Three or more curves intersecting at a common point in an image intersect in a common point in the world.

These rules greatly constrain the possible interpretations of line drawings. Consider, for instance, the Necker cube shown in Figure 6.1. The straight line segments in this figure must be interpreted as straight lines in space based on Rules 1 and 2. The reasoning is as follows. According to Rule 2, since the points in a line segment are smoothly connected in the image they must be smoothly connected in space. And according to Rule 1, since the points in a line segment are collinear in the image they must be collinear in space. Therefore they must form a line in space.

One can cast this reasoning in a Bayesian format, assuming no noise. Consider the conditional probability that a given straight-line segment S in an image actually arose from the projection of a wiggly curve in space. We can write this as P(wiggly in world|S in image) or more simply $P(W \mid S)$. We wish to show that $P(W \mid S)$ is zero. By Bayes' rule we can write

$$P(W \mid S) = \frac{P(S \mid W)P(W)}{P(S)}. \qquad (6.1)$$

Here $P(W)$ and $P(S)$ are the prior probabilities, respectively, of wiggly curves in space and of the straight line segment S. For our purposes, we do not need to estimate $P(S)$. We can estimate the numerator, viz., $P(S \mid W)P(W)$, then set $P(S)$ to a value which normalizes the numerator to a probability. Consider, then, the first factor in the numerator, the so-called likelihood $P(S \mid W)$. $P(S \mid W)$ is the

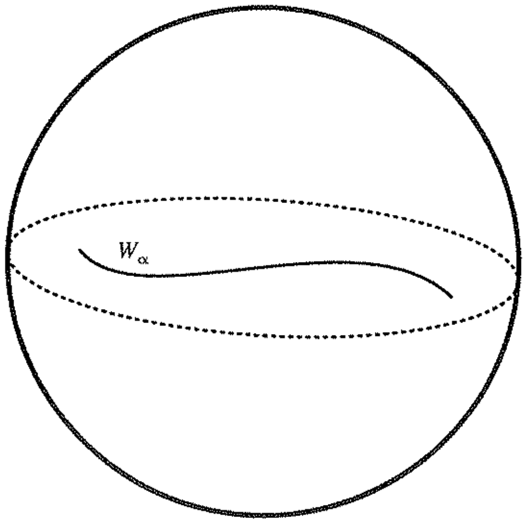


FIG. 6.2. The sphere of viewing directions around a wiggly line, with nongeneric views indicated by dashed lines.

conditional probability that a wiggly curve in space will project to the straight-line segment S. To determine $P(S \mid W)$, genericity says to assume that all viewpoints are equally likely. The set of viewing directions for a wiggly curve W_α in space is illustrated in Figure 6.2 by a sphere surrounding W_α. Since the sphere has finite area, it is possible to place a finite uniform measure on it; in the language of groups we can say that since SO(3) is a compact group it admits finite Haar measures. Given this assumption, the measure of a set of viewing directions is proportional to the area of the set. The set of viewing directions for which W_α projects to S or a scaled version of S is indicated by the dashed great circle on the sphere. This set is one dimensional; therefore, it has no area and, in consequence, zero probability. Moreover *every* wiggly curve W_α in space that can project to the line segment S, or a scaled version of S, can do so only from viewpoints on this *same* great circle. One can see this by noting that S and W_α must be entirely coplanar for W_α to project to S. Thus, not only is $P(S \mid W_\alpha) = 0$ for each α, but so also is $\int P(S \mid W_\alpha)P(W_\alpha)\, d\alpha$. Consequently, $P(S \mid W)$ is zero, which by (1) implies $P(W \mid S)$ is also zero. We conclude that genericity entails the rule: if the prior probability of 3D straight lines is nonzero, interpret any straight line in an image as straight in 3D. A wiggly 3D interpretation is nongeneric; a slight movement of the eye would reveal the wiggles.

Now consider Figure 6.3. Your initial impression is probably that this depicts some sort of pinwheel. You initially see it as flat. With difficulty you might also be able to see it as another view of the Necker cube, seen from a nongeneric view in which two vertices of the cube are precisely aligned. Why is it hard to see the Necker cube? As we have just shown, all line segments in the image must be interpreted as straight lines in space. Moreover, according to Rule 5, each of the

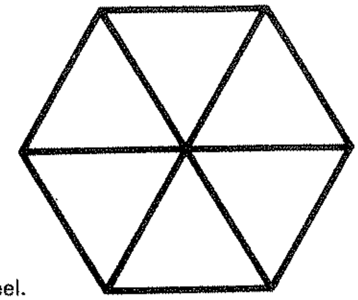


FIG. 6.3. A pinwheel.

three lines that intersect in the center of the figure must be interpreted as intersecting at a common point in space. This precludes a Necker cube interpretation.

You might argue that symmetry, not genericity, is the main principle underwriting a pinwheel interpretation in this figure. This would be the explanation given by the Gestalt principle of Prägnanz: That interpretation is to be made which is simplest. The 2D pinwheel is already highly symmetrical in 2D, so there is no need to go to a 3D interpretation.

But Figure 6.4 suggests that this is not right. Here we see a 3D shape from one generic and one nongeneric view. The generic view leads to a 3D interpretation. The nongeneric view usually does not, even though the 2D interpretation that is seen is much less simple or symmetric. We can explain the differing perceptions based on the rules derived from genericity, just as we did with the pinwheel. We cannot explain the difference by appeal to symmetry.

Genericity is not the only principle used by human vision to interpret images, and in many cases it cannot, by itself, force a unique interpretation. But it is a powerful principle. We will shortly mention other principles that interact with genericity in the generation of 3D interpretations. But Figure 6.5 indicates just

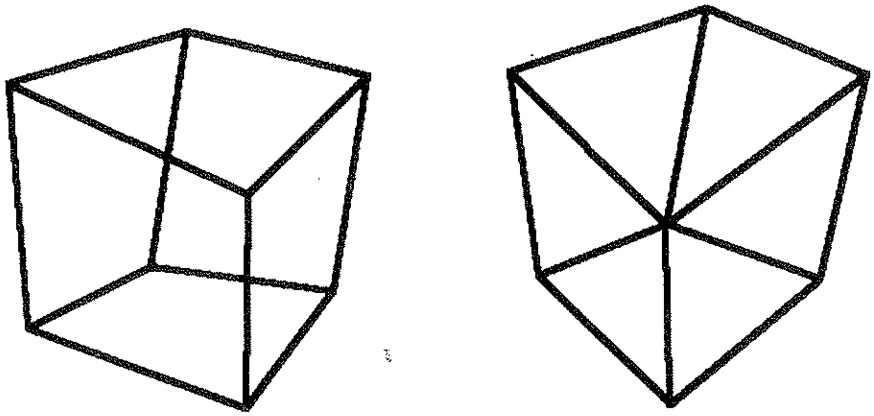


FIG. 6.4. One generic and one nongeneric view of a 3D shape (adapted from Kanizsa, 1975).

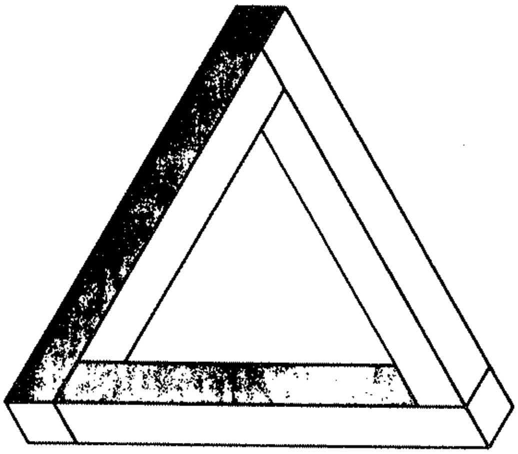


FIG. 6.5. The Penrose triangle.

how committed to genericity we can be. This figure shows the well-known Penrose triangle. At first glance this looks like a normal 3D model of a triangle. A closer look reveals that what you perceive is physically impossible. No real triangle could be built which would project to this image. However, there is a different 3D object which could be built and which would project to this image. It is a triangle broken at one corner with the two edges twisted away from each other. This has been constructed by Gregory (1970). However, to get the image shown in Figure 6.5, one must photograph this 3D model from exactly one viewpoint. Move the camera slightly and the image of the Penrose triangle is ruined. Since this physically possible 3D interpretation requires a special viewpoint, human vision rejects it. We prefer, in this case, to see a 3D interpretation which is physically impossible but satisfies genericity, rather than to see one which is physically possible and violates genericity.

The preceding analyses did not take into account the fact that real-world visual systems have only finite resolution and must tolerate noise. These limitations imply that nongeneric interpretations of images by human vision will have "small" but not zero probability. For this reason the "rules" of image interpretation based on genericity are really like cues: they can be overruled, even by a visual system that is ideal in the sense that it always infers the "most probable" interpretation of the images presented to it. Image interpretation using cues is based on a comparision of the collective weight of the cues (evidence) favoring each interpretation.

Jepson and Richards (1992) have presented counterexamples to the hypothesis that human vision always interprets images in accordance with the generic viewpoint assumption. For example, in Figure 6.6a (due to Jepson and Richards) the bottom edge of the small block on the left appears to be collinear with the bottom edge of the large block on the right. However, if the figure is rotated clockwise

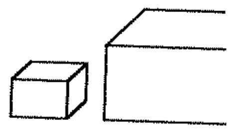


FIG. 6.6. The

by 90° then the interpre
collinear in space. Inste
than the large block. T
rotated counterclockwis
A proximity rule of
this display. According
should be interpreted as
another instance of the
bilities): If two features
viewpoints would place
appears to explain the ap
the block.
In Figure 6.6a the bo
collinear in the image.
each other in the image a
edge of the large block.
blocks are at the same h
that they are at the same
depth assignment. This i
imity rule, which would
be at approximately same
to have been overruled
collinearity rules in this
However, when the di
rule would be expected t
blocks rather than down
top surfaces of the blocks
to overrule the collineari
clockwise by 90°, the hei
be seen behind the smal

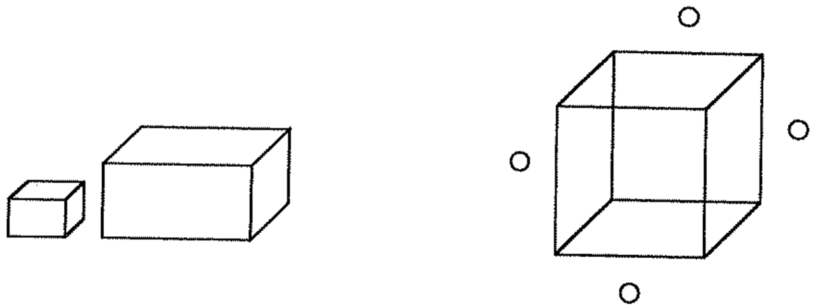


FIG. 6.6. The interaction of genericity with other cues to depth.

by 90° then the interpretation changes. Now those edges are not interpreted as collinear in space. Instead the small block appears to be closer to the observer than the large block. The effect is even stronger when the original figure is rotated counterclockwise by 90°.

A proximity rule of depth assignment seems to be affecting our perception of this display. According to this rule features that are near each other in an image should be interpreted as being near each other in space. The proximity rule is another instance of the principle of genericity (in the sense of "small" probabilities): If two features are widely separated in space, then only a small range of viewpoints would place them near each other in an image. In Figure 6.6b this appears to explain the apparent depth of the small circles relative to the edges of the block.

In Figure 6.6a the bottom edges of the blocks (in the original orientation) are collinear in the image. However, the features on the blocks that are nearest to each other in the image are the right rear edge of the small block and the left front edge of the large block. In the original orientation the bottom edges of the two blocks are at the same height in the visual field. This supports the interpretation that they are at the same depth in space using the "height-in-the-field" rule of depth assignment. This interpretation is also supported by collinearity. The proximity rule, which would predict that the right rear edge of the small block would be at approximately same depth as the left front edge of the large block, appears to have been overruled by the combination of the "height in the field" and collinearity rules in this case.

However, when the display is rotated clockwise by 90° the height-in-the-field rule would be expected to be inoperative since the observer is looking up at the blocks rather than down at them from above; the bottom surfaces rather than the top surfaces of the blocks are now visible. In this case the proximity rule appears to overrule the collinearity rule. When the original display is rotated counterclockwise by 90°, the height-in-the-field rule predicts that the large block should be seen behind the small block, in agreement with the proximity rule. The combination of these rules strongly overrules collinearity. Thus there appears to be a good deal of both cooperation and competition among the various genericity-based rules and other rules of depth assignment.

GENERICITY AND PARTS

Genericity, as we have seen, helps to guide the assignment of 3D structures to 2D images. To recognize these 3D structures as objects, further processing is required. One aspect of this further processing is the decomposition of 3D structures into simpler subunits, or "parts."

Part decompositions aid the recognition process by allowing recognition despite occlusions and despite nonrigid motions of parts, such as legs or arms (Biederman, 1987; Hoffman & Richards, 1984; Marr & Nishihara, 1978). Ideally a part decomposition should be (1) easily computed from images, (2) applicable to all classes of objects, and (3) independent of viewing geometry.

Genericity motivates an approach to part decompositions that is close to ideal. Figure 6.7 illustrates the basic idea. On the left of the figure are two objects. On the right the two have been generically intersected to form a single composite object. Since the intersection is generic, the tangent planes to the surfaces of the two objects are almost never parallel at the points where the two surfaces meet. The two surfaces almost everywhere meet in a concave discontinuity. This is illustrated by the dashed circular contour in Figure 6.7. That surfaces generically intersect in concave discontinuities follows from a transversality theorem of the field of differential topology (Guillemin & Pollack, 1974).

This motivates a simple rule for decomposing 3D shapes into parts: Divide shapes into parts at contours of concave discontinuity (Hoffman & Richards, 1984).

An application of this rule is illustrated in Figure 6.8. On the left is the well-known Schröder staircase. At first this appears to be an ascending staircase, and the two dots appear to lie on the rise and tread of a single step. Note that all the steps are bounded by lines of concave discontinuity, as dictated by genericity.

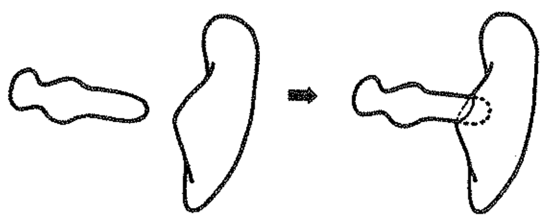


FIG. 6.7. Generic intersections of the surfaces of objects leads to concave discontinuities.

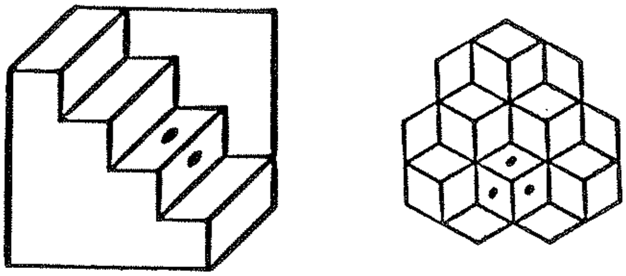


FIG. 6.8. The Schroder staircase and stacked cubes.

Contours of convex discontinuity separate the rise and tread of a single step, but do not serve to carve the staircase into steps. Upon further inspection, this staircase will appear to reverse figure and ground, so that concave discontinuities become convex, and vice versa. Our rule for part decomposition therefore predicts a new organization into parts, with part boundaries along the new lines of concave discontinuity. You can check for yourself that this prediction is fulfilled. The two dots which appeared to be on the rise and tread of a single step now appear to be on two distinct steps. Similar comments hold for the stacked cubes on the right. The cubes are all separated along contours of concave discontinuity, with the three dots at first appearing to lie on a single cube. Upon further inspection, figure and ground reverse and one gets new part boundaries, and new cubes, as predicted by genericity.

Smoothing contours of concave discontinuity leads to extrema of surface curvature, specifically negative extrema in one of the principal curvatures. (An extremum of curvature is a negative extremum if it is in a concave region of the surface.) This suggests that for smooth objects we use negative extrema of the principal curvatures to delineate parts (Hoffman & Richards, 1984). An example of the parts given by this rule is shown for the "cosine surface" illustrated in Figure 6.9. The dashed circular contours indicate the negative extrema of surface curvature and, therefore, the part boundaries. These boundaries organize the surface into a succession of ring-shaped hills. If you turn this illustration upside down, you will notice that the dashed circular contours no longer work as part boundaries. Now they appear to lie in the middle of the hills, instead of between the hills. Your organization of the cosine surface into parts has changed. The reason is that turning the illustration upside down causes your visual system to reverse the choice of figure and ground on the cosine surface. (We reverse figure and ground because, apparently, we prefer to see the surface lying below us rather than floating above us.) This reversal of figure and ground turns concavities into convexities, and vice versa. Thus, negative extrema of the principal curvatures become positive extrema, and vice versa. And, since negative extrema determine the part boundaries, these boundaries must move to the new negative extrema. Consequently we see new parts.

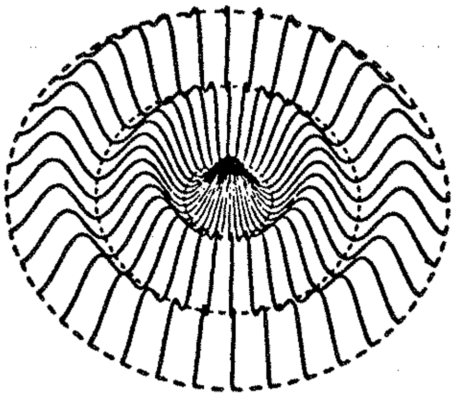


FIG. 6.9. The cosine surface.

Koenderink (1990) has proposed a theory of object recognition based on the idea of generic versus accidental views. In this theory the ambient space of possible viewpoints on a scene is divided into "cells." The cell that contains a particular viewpoint is the largest connected region of the ambient space within which all viewpoints give rise to "qualitatively" equivalent images. The "cell walls" in this theory define surfaces in space. When an observer crosses a cell wall, the qualitative structure of the image changes. If we are considering the case of orthographic projection, then the cells become just patches on the sphere of viewing directions, and the cell walls are the curves that bound those patches. Koenderink's claim is that much of the quantitative, metric information in images is not used by the visual system and that, for most purposes, object recognition proceeds using only qualitative information.

GENERICITY AND ILLUSORY CONTOURS

Genericity turns out to have important implications for the perception of surfaces, including illusory surfaces. Nakayama and Shimojo (1990, 1992) used it to explain phenomena in the area of stereoscopic perception of untextured surfaces, and Kellman and Shipley (1991) used it in their "discontinuity" theory of perceptual unit formation.

When Figure 6.10 (Nakayama & Shimojo, 1990, 1992) is cross-fused (by crossing one's eyes so that the left and right figures superimpose in the middle) most people perceive a horizontal bar overlaying a vertical bar, along with illusory contours in the central region that complete the boundary of the horizontal bar. In other words, the black region appears to split into two distinct surfaces, giving a 3D segmentation. Since the cross in this display is untextured, no disparity information is available in its interior. Also, the horizontal edges carry no disparity information since any point along such an edge in one eye could

FIG. 6.10. Stereo cross (adapted from Nakayama and Shimojo, 1990).

match any point along the corresponding edge in the other eye. (These different correspondences reflect the fact that a physical edge that projects to a straight horizontal edge in both eyes could still oscillate in depth in an arbitrary way, as discussed earlier.) Only the vertical edges carry unambiguous horizontal disparity information.

Nakayama and Shimojo's explanation for this perception is based on the generic viewpoint assumption. This assumption predicts the interpretation perceived by most people and eliminates others that would be predicted by other algorithms. Their argument is essentially the following: According to the generic viewpoint assumption edges that are straight in an image are straight in space, and edges that are collinear in an image are collinear in space. Since the disparity information in the image tells the observer that the outer endpoints of the horizontal edges of the horizontal rectangle are at the same depth (because the vertical edges *do* carry horizontal disparity information), these rules imply that the horizontal edges must be frontoplanar.

Nakayama and Shimojo point out that no simple disparity-spreading scheme is consistent with this perceptual interpretation. The idea here is that if an interpolation algorithm were used to assign depth to the interior of the cross using the disparity information available at its boundaries, then the horizontal arms should be smoothly interpolated in depth between the depths the vertical edges of the horizontal rectangle and the vertical edges of the vertical rectangle (see Figure 6.10). (This assumes that disparity signals generated by the boundaries of a homogeneous connected region can spread freely within that region.)

Kellman and Shipley (1991) have proposed a theory of "perceptual unit formation" using the principle of transversality. The term "perceptual unit formation" refers to illusory surface and contour formation, as well as occluded (amodal) surface and contour formation. According to their theory a necessary condition for perceptual unit formation is the presence of tangent discontinuities in the boundaries of regions in the image. If a pair of contours leading into distinct tangent discontinuities are "relatable," meaning that their extensions beyond the tangent discontinuitites intersect at an angle greater than or equal to 90°, then interpolation occurs between the contours. In this way perceptual units are formed.

The justification for the role given to tangent discontinuities in Kellman and Shipley's theory is based on the concept of generic occlusion: The transversality principle implies that tangent discontinuities almost always (in the technical sense) occur when occlusion is present. Essentially, the concept of generic occlusion in terms of transversality clarifies and justifies in a formal way the use of T junctions to infer interposition.

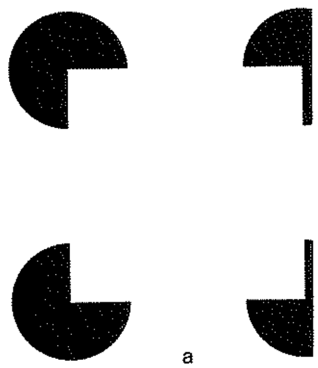
Later we will present a different proposal about the role of tangent discontinuities in the perception of illusory surfaces. Our proposal will allow for the fact that some examples of illusory surfaces do not give the impression of being interposed in front of their inducers. This will help to explain some data obtained by Shipley and Kellman (1990) that was inconsistent with their theory.

GENERICITY AND ICOs

In this section we will investigate some implications of genericity for the perception of illusory contours (ICs). In particular, we will propose necessary conditions for the perception of ICs in which the illusory surface appears to partially occlude its inducers. We will refer to these as "ICOs," which is short for "illusory contours that occlude." It will be shown that the generic viewpoint assumption places restrictions on the topological and first order differentiable structure of displays in which ICOs are perceived. However, as mentioned earlier, our analyses will assume infinite resolution and no noise. Human vision, with its finite resolution and inevitable noise can be expected to treat our "necessary conditions" as biases rather than strict rules.

For the case of ICOs induced by "blobs" we will use the principle of transversality. Consider Figure 6.11a. In this display most people see a white square that stands out from the surrounding white area. The white square appears to be in front of and partially occluding black disks. Notice that the principle of transversality, as applied to occlusion, is obeyed here: the tangents of the circles differ from the tangents of illusory square at the points where the contours meet. The occlusion is generic. However, in Figure 6.11b we have smoothed out the sharp convex corners of the "pacmen" in Figure 6.11a. If an ICO were seen in this display, then the occlusion would not be generic. In fact, most people do see a weak IC in Figure 6.11b, but they do not describe it as an ICO. Most people perceive the blobs to be pushed up against the side of the illusory square, as though the blobs were made of a soft, flexible material that has been deformed to fit the square's shape.

Thus, we propose that tangent discontinuities are a necessary condition for ICOs. But we do not claim that they are necessary for all ICs. Shipley and Kellman (1990) found in their experiments that subjects do perceive ICs in displays in which the tangent discontinuities have been removed, although the ICs were usually rated by subjects as weaker than when discontinuities were present. However, the IC in Figure 6.12 is rated by most subjects as relatively strong, although it contains no tangent discontinuities and the short line segments by themselves do not produce a signficant IC. There is not much brightness



a

FIG. 6.11. (a) Tange
necessary for the p
Smooth the discontir
ing illusory surface i

enhancement in this figur
(e.g., Kennedy, 1988) dist
ness enhancement, since it
not always accompanied b
Giannini, & Bonaiuto, 19
Purghe & Katsaras, 1991)
the distinction between IC
arguments for the necessity
generic occlusion only ap
The concept of generi
induced by the ends of line
line drawings are useful i
typical example of an IC

FIG. 6.12. F
lusory cont
figures with
tinuities, bu
tours do no

concept of generic occlu-
a formal way the use of

role of tangent discon-
oposal will allow for the
the impression of being
plain some data obtained
vith their theory.

genericity for the percep-
propose necessary condi-
rface appears to partially
hich is short for "illusory
ric viewpoint assumption
lifferentiable structure of
entioned earlier, our an-
nan vision, with its finite
at our "necessary condi-

the principle of transver-
le see a white square that
square appears to be in
the principle of transver-
gents of the circles differ
e the contours meet. The
e smoothed out the sharp
n ICO were seen in this
ct, most people do see a
as an ICO. Most people
of the illusory square, as
that has been deformed to

a necessary condition for
for all ICs. Shipley and
jects do perceive ICs in
n removed, although the
hen discontinuities were
nost subjects as relatively
nd the short line segments
is not much brightness

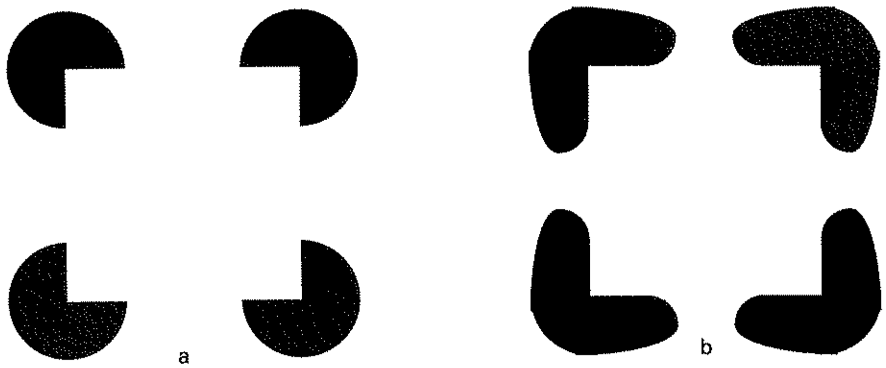


FIG. 6.11. (a) Tangent discontinuities at the "lips" of the pacmen are necessary for the perception of an occluding illusory surface. (b) Smooth the discontinuities at the "lips" of the pacmen and no occluding illusory surface is seen.

enhancement in this figure. However we, as well as many other researchers (e.g., Kennedy, 1988) distinguish ratings of IC strength from ratings of brightness enhancement, since it has been shown that strong illusions of contour are not always accompanied by enhanced brightness. Other researchers (Bonaiuto, Giannini, & Bonaiuto, 1991; Kennedy, 1978; Kennedy, 1988; Purghe, 1991; Purghe & Katsaras, 1991) have also emphasized the theoretical significance of the distinction between ICs and ICOs. And from a theoretical point of view, the arguments for the necessity of tangent discontinuities based on transversality and generic occlusion only apply to ICOs.

The concept of generic occlusion can also be applied to the case of ICs induced by the ends of lines. The rules described earlier for the interpretation of line drawings are useful in this connection. Consider Figure 6.13a. This is a typical example of an IC induced by line endings. In this display most people

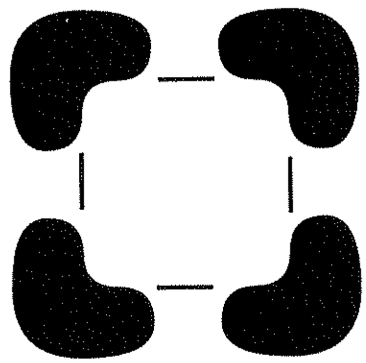


FIG. 6.12. Relatively strong illusory contours can be seen in figures without tangent discontinuities, but these illusory contours do not appear to occlude.

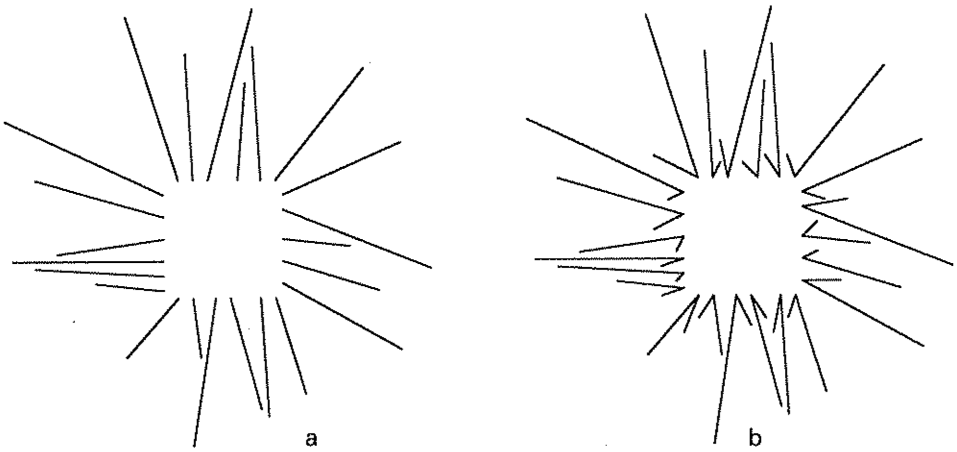


FIG. 6.13. (a) An illusory surface induced by line segments. (b) Adding more line segments to form vertices reduces the strength of the illusory surface and makes it appear to be nonoccluding.

perceive an illusory white square that is interposed in front of the lines. The lines appear to continue underneath the illusory square. Now, in Figure 6.13b we added short line segments to Figure 6.13a to make L junctions at the points where the inducing line endings occurred in Figure 6.13a. This change has weakened the IC, but it has also produced a *qualitative* change in the IC. Its apparent depth has moved back to approximately the same depth as the L junctions of the inducers. The inducing lines no longer appear to continue underneath the illusory surface. Instead, the L junctions of the inducers appear to be stuck into the side of the illusory surface. The IC is not an ICO.

We can account for this qualitative change in appearance using the concept of generic occlusion. The IC in Figure 6.13b passes right through the L junctions of the inducers. So, including the IC, there is in fact a K junction at each point where the IC meets an inducer. From Rule 5 for line-drawing interpretation, we know that if our viewpoint is generic then at a K junction all contours must be at the same depth. Therefore, the IC cannot be an ICO if genericity is obeyed.

Similarly, comparing Figure 6.14a with Figure 6.14b, most observers perceive an ICO in Figure 6.14a, and no IC or a nonoccluding IC in Figure 6.14b. Again, this can be explained by the K junctions formed at the intersection points of the inducers and the potential IC. What is interesting about this example is that the tangents of the inducing lines agree at the vertices. When attention is restricted to a small area around a vertex, it appears more like a termination of an isolated line than an L junction. Yet, the IC is strongly affected. This appears to be inconsistent with the predictions of the line-end-contrast theory of Frisby and Clatworthy (1975) as well as the neural network theory of Grossberg and Mingolla (1985).

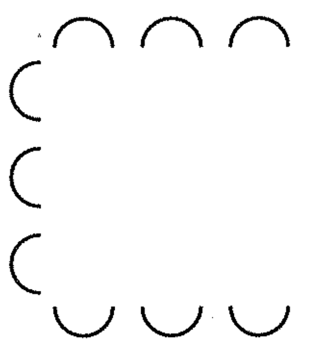


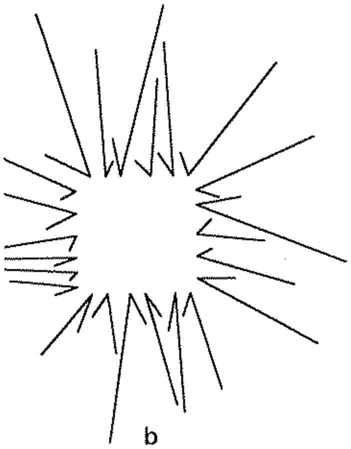
FIG. 6.14. (a) An
more curves to fo
face and makes i

WHY OUTL

Genericity also provide
"blobs," when drawn
question has been wide
Kanizsa, 1974; Rock,
servers report seeing or
short line segments in
strong IC is seen.

We can understand,
displays are perceived
seen in Figure 6.15a. T
occluded blob-shaped
very thin edge of thos
1991). On the other ha
the fact that they are ly
at the same depth as the
of viewpoint. Now the
by Rule 4, the short li
circular arcs at the junc
same depth as the circ
similar argument can b

Intuitively the idea i
being occluded by an il
system would have to
where the circular arcs



ıe segments. (b) Add-
əs the strength of the
ıccluding.

n front of the lines. The lines
. Now, in Figure 6.13b we
ke L junctions at the points
ure 6.13a. This change has
itative change in the IC. Its
he same depth as the L junc-
ppear to continue underneath
: inducers appear to be stuck
an ICO.
ɔearance using the concept of
ght through the L junctions of
t a K junction at each point
ıe-drawing interpretation, we
nction all contours must be at
CO if genericity is obeyed.
: 6.14b, most observers per-
ccluding IC in Figure 6.14b.
med at the intersection points
ting about this example is that
rtices. When attention is re-
more like a termination of an
ıgly affected. This appears to
-contrast theory of Frisby and
ıeory of Grossberg and Min-

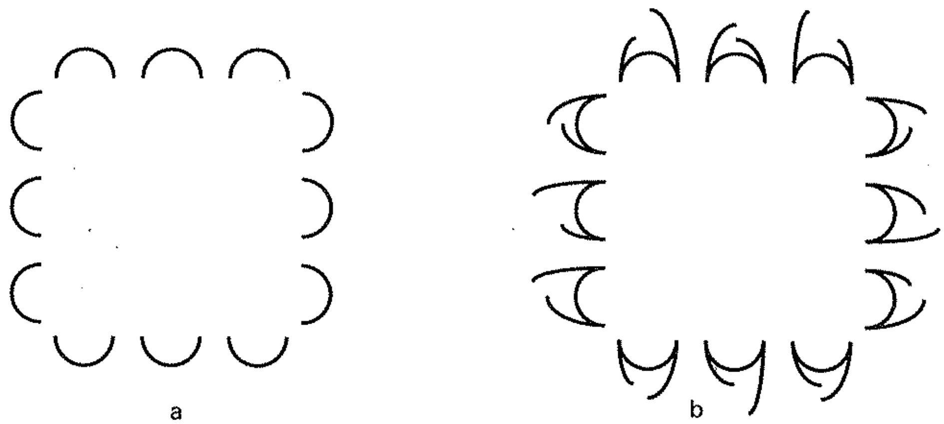


FIG. 6.14. (a) An illusory surface induced by semicircles. (b) Adding more curves to form vertices reduces the strength of the illusory surface and makes it appear to be nonoccluding.

WHY OUTLINES OF BLOBS DO NOT INDUCE ICOs

Genericity also provides an interesting new perspective on the question of why "blobs," when drawn only in outline, do not produce significant ICs. This question has been widely discussed among researchers in IC perception (e.g., Kanizsa, 1974; Rock, 1987). For example, consider Figure 6.15a. Most observers report seeing only a very weak if any IC in this display. However, if the short line segments in Figure 6.15a are removed, as in Figure 6.15b, then a strong IC is seen.

We can understand, at least in part, the difference in the way these two displays are perceived by using genericity: Assume that an illusory surface is seen in Figure 6.15a. Then the short line segments cannot be viewed as partially occluded blob-shaped elements since it would be highly improbable that just a very thin edge of those blobs would be visible (also see Kellman & Shipley, 1991). On the other hand, if they are viewed as unoccluded line segments, then the fact that they are lying on (or directly next to) the IC means that they must be at the same depth as the IC. Otherwise it would imply an improbable coincidence of viewpoint. Now the short line segments coterminate with the circular arcs. So by Rule 4, the short line segments must also be seen at the same depth as the circular arcs at the junction points. Therefore, the potential IC must be seen at the same depth as the circular arcs at the junction points, so it cannot be an ICO. A similar argument can be used in the case of outlines of pacman inducers.

Intuitively the idea is that if the circular arcs in Figure 6.15a were perceived as being occluded by an illusory surface, as they are in Figure 6.15b, then the visual system would have to "wonder" why the short line segments terminate exactly where the circular arcs pass underneath the illusory surface in the image.

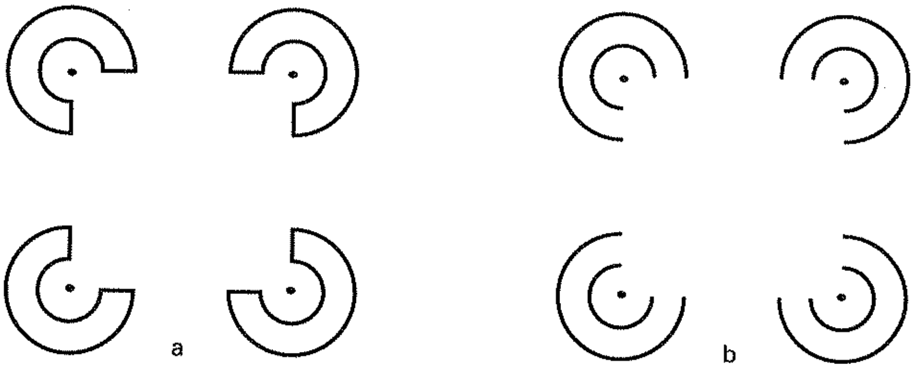


FIG. 6.15. (a) Blob outlines do not induce illusory contours. (b) Removing the short line segments in part (a) allows an ICO to emerge.

Kanizsa (1974) has discussed displays very similar to these in comparing his theory of IC perception with that of Gregory: "According to Gregory the sense data are used by the brain according to certain strategies, in order to decide which object has the highest probability of being present. But then, comparing the perceptual effects of Figures 12.26a and 12.26b [similar to our figures 6.15b and 6.15a, respectively], one should conclude that for the brain [a corner of the type in Figure 12.26b] is more probable than [a corner of the type in Figure 12.26a], a conclusion that seems to me rather implausible."

In our view it is not that the inducers in Figure 6.15a are more probable than those in Figure 6.15b, but that those in Figure 6.15a would be highly *improbable* if there were a surface (the potential illusory figure) in front, whereas those in Figure 6.15b would not. In other words, given that an ICO occurs in Figure 6.15b our theory helps us to understand why, despite the similarity between Figures 6.15a and 6.15b, an ICO does not occur in the former.

Genericity is also helpful in understanding the role of line segments that run along the length of an IC, as in Figure 6.12 discussed earlier. Consider Figure 6.16, in which the short line segments of Figure 6.15a have been moved away from the circular arcs. Most observers perceive an IC in this figure that is as strong or stronger than the one in Figure 6.15b. Note that the circular arcs appear to be partially occluded, whereas the line segments that lie along the IC do not. The line segments appear to be closer than the circular arcs, lying at the same depth as the IC. This is what would be predicted on the basis of the genericity arguments given above. The line segments are usually described by observers as entities of some sort attached to the edge of the illusory surface. Generally, in displays in which some of the inducers of an illusory figure are consistent with generic occlusion and some are not, the former are seen as partially occluded, while the latter are seen as unoccluded and are pulled forward to the same depth as the IC.

FIG. 6.16. A n
Figure 6.15a tha
strong IC.

Kanizsa (1974) has argued
with line-end inducers. Suppo
the effects seen in the displays
be a more satisfactory explana
predicts perceived depth relat

Genericity can be applied i
the "neon color spreading" s
with a display that produces n
are added that intersect the c
result is that the color spreadir
ency disappears (see Albert a

The reports of our visual syst
of the world. Rather these rep
have been wired into the billi
ples underlie these inference
principle of genericity. As we
of line drawings, the parts of
genericity and of its interactic
is a promising direction for r

AC

We thank B. Bennett, M. Br
useful discussions, and M.
authors were supported by
N00014–88–K–0354.

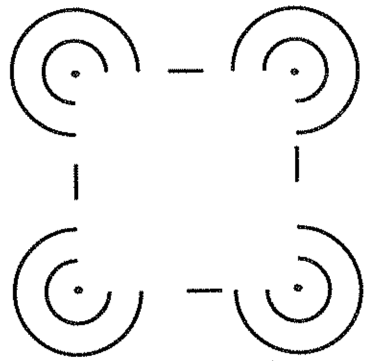


FIG. 6.16. A modification of Figure 6.15a that does induce a strong IC.

Kanizsa (1974) has argued that "closure" can explain the perception of ICs with line-end inducers. Supporters of this theory might claim that it can explain the effects seen in the displays in this section. However, we believe genericity to be a more satisfactory explanation, since it is a valid ecological constraint. It also predicts perceived depth relations, which closure does not.

Genericity can be applied in a similar way to obtain necessary conditions for the "neon color spreading" spreading effect (see Van Tuijl, 1975). Beginning with a display that produces neon color spreading using colored lines, more lines are added that intersect the original lines at their points of color change. The result is that the color spreading is greatly reduced and the perception of transparency disappears (see Albert and Hoffman, in press).

CONCLUSION

The reports of our visual systems are not unbiased accounts dictated by the state of the world. Rather these reports are the result of sophisticated inferences which have been wired into the billions of neurons which process vision. Many principles underlie these inferences. One of the most powerful and ubiquitous is the principle of genericity. As we have seen, genericity is integral to our perceptions of line drawings, the parts of objects, and subjective surfaces. Further study of genericity and of its interactions with other principles that shape our perceptions is a promising direction for research into human vision.

ACKNOWLEDGMENTS

We thank B. Bennett, M. Braunstein, J. Liter, C. Prakash, and S. Richman for useful discussions, and M. D'Zmura for comments on an earlier draft. The authors were supported by NSF grant DIR-9014278 and by ONR contract N00014–88–K–0354.

REFERENCES

Albert, M. K., & Hoffman, D. D. (in press). Generic visions: General position assumptions in visual perception. *Scientific American.*

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

Biederman, I. (1987). Recognition-by-components: A theory of human image interpretation. *Psychological Review, 94,* 115–147.

Binford, T. O. (1981). Inferring surfaces from images. *Artificial Intelligence, 17,* 205–244.

Bonaiuto, P., Giannini, A. M., & Bonaiuto, M. (1991). Visual illusory productions with or without amodal completion. *Perception, 20,* 243–257.

Frisby, J. P., & Clatworthy, J. L. (1975). Illusory contours: Curious cases of simultaneous brightness contrast? *Perception, 4,* 349–357.

Gregory, R. L. (1970). *The intelligent eye.* New York: McGraw-Hill.

Grossberg, S., & Mingolla, E. (1985). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review, 92,* 173–211.

Guillemin, V., & Pollack, A. (1974). *Differential topology.* Englewood Cliffs, NJ: Prentice-Hall.

Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition, 18,* 65–96.

Jepson, A., & Richards, W. (1992). *What makes a good feature?* (MIT AI Memo 1356).

Kanizsa, G. (1974). Contours without gradients or cognitive contours? *Italian Journal of Psychology, 1,* 93–112.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology, 23,* 141–221.

Kennedy, J. M. (1978). Illusory contours not due to completion. *Perception, 7,* 187–189.

Kennedy, J. M. (1988). Line endings and subjective contours. *Spatial Vision, 3,* 151–158.

Kilpatrick, K. P. (1952). Elementary demonstrations of perceptual phenomena. In K. Kilpatrick (Ed.), *Human behavior from the transactional point of view* (pp. 1–15). Hanover, NH: Institute for Associated Research.

Koenderink, J. (1990). *Solid shape.* Cambridge: MIT Press.

Lowe, D. (1985). *Perceptual organization and visual recognition.* Boston: Kluwer.

Lowe, D. G., & Binford, T. O. (1981). The interpretation of three-dimensional structure from image curves. *Proceedings of IJCAI-7* (pp. 613–618). Vancouver, British Columbia, Canada:

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London, B200,* 269–294.

Nakayama, K., & Shimojo, S. (1990). Towards a neural understanding of visual surface representation. *Cold Spring Harbor Symposium on Quantitative Biology, 40,* 911–924.

Nakayama, K., & Shimojo, S. (1992). Experiencing and perceiving visual surfaces. *Science, 257,* 1357–1363.

Purghe, F. (1991). Is amodal completion necessary for the formation of illusory figures? *Perception, 20,* 623–636.

Purghe, F., & Katsaras, P. (1991). Figural conditions affecting the formation of anomalous surfaces: Overall configuration versus single stimulus part. *Perception, 20,* 193–206.

Rock, I. (1987). A problem solving approach to illusory contours. In S. Petry & G. E. Meyer (Eds.), *The perception of illusory contours* (pp. 62–70). New York: Springer-Verlag.

Shipley, T. F., & Kellman, P. J. (1990). The role of discontinuities in the perception of subjective figures. *Perception and Psychophysics, 48,* 259–270.

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

van Tuijl, H.F.J.M. (1975). A new visual illusion: Neonlike color spreading and complementary color induction between subjective contours. *Acta Psychologica, 39,* 441–445.

Witkin, A. P., & Tenenbaum, J. M. (1983). On the role of structure in vision. In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and machine vision* (pp. 481–543). New York: Academic Press.

7 En M Cc Ar

Ehtibar N. Dzhafarc
University of Illinois

The "empirical meanin
priori restrictions on st
down as "empirically n
truth value (true or false
measurements of the q
statement that is "empi
equivalently reformulate
ful." This is achieved by
constants, whose numer
specified class. Provide
mechanics or psychoph
(such as mass in grams
in this chapter determin
the invariance of these
these measurements: th
dependent constants, is
which the invariance co
given statement is logic
one should be able to m
the aid of the same set
in physics, for instance
given set of physical l
motion) by demonstrat
(invariant under scaling
stants that have been e