COGNITION

Cognition 63 (1997) 29–78

Salience of visual parts

Donald D. Hoffman*, Manish Singh

Department of Cognitive Science, University of California, Irvine, CA 92697, USA

Received 20 February 1996, final version 30 October 1996

Abstract

Many objects have component parts, and these parts often differ in their visual salience. In this paper we present a theory of part salience. The theory builds on the *minima rule* for defining part boundaries. According to this rule, human vision defines part boundaries at negative minima of curvature on silhouettes, and along negative minima of the principal curvatures on surfaces. We propose that the salience of a part depends on (at least) three factors: its size relative to the whole object, the degree to which it protrudes, and the strength of its boundaries. We present evidence that these factors influence visual processes which determine the choice of figure and ground. We give quantitative definitions for the factors, visual demonstrations of their effects, and results of psychophysical experiments. ©1997 Elsevier Science B.V.

1. Introduction

When you glance at a fan or a futon and recognize it, you do so with apparent ease. In fact, however, during that glance billions of neurons labor in concert to transform, step by step, the shower of photons hitting each eye into recognized objects such as fans and futons. The ease of recognition, like the ease of an Olympic skater, is deceptive. Recognizing futons from photons is no small task.

Indeed it is not one task, but many. Color, shading, shape, motion, texture, and context are all typically used in the process. It is natural then for the theorist, faced with this complexity, to choose a strategy of divide and conquer. And fortunately, as Fig. 1 shows, there are natural ways to divide the problem of recognition. Notice that the silhouettes in this figure are easily recognized. But there is no color, shading, motion, or texture. Nor is context of any help; your present location, the time of day, and other contextual factors could not help you to predict

* Corresponding author. E-mail: ddhoff@uci.edu.

PII S0010-0277(96)00791-3

Fig. 1. Some easily recognized silhouettes. This illustrates that color, motion, texture, shading, and context are often not necessary for successful recognition. Shape alone is often sufficient.

what is in the figure. The figure contains only shape, and that of a restricted type, namely silhouettes.

Thus, in many cases, shape alone permits successful recognition of objects. Indeed we can recognize thousands of objects entirely by their shapes. These large numbers raise a major obstacle to successful recognition, namely indexing—efficiently searching one's memory of familiar objects to seek a best match to a given image. What Fig. 1 shows is that shape by itself is enough, in many cases, to index with success. Somehow human vision can represent silhouettes in a way that provides a useful first index into its memory of shapes. This first index is computed ''bottom up'' and might not be right on target, but it must be close enough so that any top-down searches it triggers can converge quickly to the right answer.

In light of these remarks we restrict attention, throughout this paper, to recognition by shape.

How is shape represented to provide a useful first index? Research to date yields few firm conclusions. However, there is growing consensus that representing shapes in terms of their parts may aid the recognition process in human vision (Baylis and Driver, 1995a,b; Bennett and Hoffman, 1987; Beusmans et al., 1987; Biederman, 1987; Biederman and Cooper, 1991; Braunstein et al., 1989; Driver and Baylis, 1995; Hoffman, 1983a,b; Hoffman and Richards, 1984; Marr, 1977, 1982; Marr and Nishihara, 1978; Palmer, 1975, 1977; Siddiqi et al., 1996; Stevens and Brookes, 1988; Todd et al., 1995; Tversky and Hemenway, 1984; but see Cave and Kosslyn, 1993). Parts may aid as well in computer vision (Binford, 1971; Brooks, 1981; Dickinson et al., 1992; Guzman, 1971; Pentland, 1986; Siddiqi and Kimia, 1995; Terzopoulos et al., 1987; Winston, 1975). The idea is that for you to recognize some shape in an image as, say, a cat, you must first decompose the shape into parts—for three reasons. First, cats are opaque, and second, cats can hide behind other opaque objects, as when a cat peeks from behind a chair. For both reasons you can't see all of a cat in a single glance. Thus to recognize a cat you must find and represent its parts that are visible in your image. The visible parts permit a first index into your catalogue of shapes, starting further routines which result in recognizing the cat. But the cat poses another problem. It walks, thereby moving its body nonrigidly. Here again, parts can come to the rescue. If you find the right parts of the cat, say such rigidly moving subshapes as its legs and feet, and represent them and their (changing) spatial relationships, again you might just recognize the cat. So the opacity of most objects and the nonrigidity of some makes parts a useful, perhaps essential, approach to recognition.

But one might ask, Which parts? How shall I find a cat's tail if I don't yet have a cat? We can solve this problem in two ways. We can define a priori a set of *basic shapes* that are the possible parts. Our task is then to find these basic shapes in images. Or we can instead define, by means of general computational rules, the *boundaries* between parts, that is, those points on a shape where one part ends and the next begins. Our task is then to find these boundaries in images.

Proponents of basic shapes have studied many alternatives: polyhedra (Roberts, 1965; Waltz, 1975; Winston, 1975), generalized cones and cylinders (Binford, 1971; Brooks, 1981; Marr and Nishihara, 1978), geons (Biederman, 1987), and superquadrics (Pentland, 1986). In each case the basic shapes reveal two drawbacks. They are ad hoc in origin and limited in scope. They are limited in scope since many objects are not composed solely of geons, polyhedra, superquadrics, generalized cylinders, or some combination: consider, for instance, a face or a shoe. They are ad hoc in origin because either (1) they are not derived from first principles or (2) they are derived from first principles but are not the entire set of basic shapes that follow from these principles. Polyhedra, superquadrics, and generalized cylinders make no appeal to first principles. Geons, in contrast, do appeal to the principle of ''nonaccidental properties'' (Witkin and Tenenbaum, 1983; Lowe, 1985). They are defined using three-dimensional (3D) features which, generically, survive under projection. Some examples are features like straight versus curved (only by an accident of view could a curve in 3D project to a straight line in an image) and parallel versus nonparallel (only by an accident of view could lines not parallel in 3D look parallel in the image). However, geons are not the entire set of primitives that follow from the principle of nonaccidental properties. Geons can only end, for instance, either in points, like a sharpened pencil, or in truncations, like a new unsharpened pencil (Biederman, 1987). They don't have rounded tips, although the distinction between tips that are rounded, pointed, and truncated is one that survives projection, and is critical to the proper recognition of fingers, toes, and peeled bananas (which have rounded tips that are not, even roughly, pointed or truncated). Thus every collection of basic shapes that has heretofore been proposed has in fact been ad hoc. Nevertheless they may be useful as qualitative descriptors of parts, rather than as an algorithm for parsing objects into parts.

Proponents of boundaries have also studied many alternatives. Rules for defining part boundaries include ''deep concavities'' (Marr and Nishihara, 1978), ''sharp concavities'' (Biederman, 1987), ''concave regions'' (Biederman, 1987), ''limbs and necks'' (Siddiqi and Kimia, 1995), and the ''minima rule'' (Hoffman and Richards, 1984). According to the minima rule, human vision defines part boundaries at negative minima of curvature on silhouettes, and along negative minima of the principal curvatures on surfaces. The other rules (except for limbs and necks, which we mention later) are similar in spirit but weaker in precision. The minima rule states precisely what they try to capture. Moreover converging experimental evidence suggests that human vision does in fact break shapes into parts as per the minima rule (Baylis and Driver, 1995a,b; Braunstein et al., 1989; Driver and Baylis, 1995; Hoffman, 1983a,b; Hoffman and Richards, 1984), and that it finds parts *preattentively* (Baylis and Driver, 1995a,b; Driver and Baylis, 1995). You find them early and you can't stop yourself. Therefore in what follows we build on the minima rule.

Although the minima rule gives precise points for carving shapes into parts, it neither describes the resulting parts nor compares their visual salience (Hoffman and Richards, 1984). In this paper we study salience. Our plan, in brief, is as follows. Since the minima rule is our point of departure, we first review its definition. This keeps the paper self contained. We then develop rules for part salience: first in 2D, then in 3D.

What do we mean by the salience of a part? As we said before, parts help us index our memory of shapes. Their salience determines, in part, their *efficacy* as an index. This efficacy might be measured in reaction times, error rates, confidence ratings, and judgments of figure and ground. For instance, as we demonstrate in Section 8, a low salience part sometimes has *no* efficacy as an index. This can happen because human vision prefers, ceteribus paribus, to choose figure and ground so that figure has the more salient parts. Thus a less salient part is not even used as an index whenever it loses in this figure–ground competition. Our argument here follows a pattern seen many times before in the literature. It has been shown that ''good'' parts provide better retrieval cues for recalling shapes (Bower and Glass, 1976), that ''good'' parts are themselves better recalled (Palmer, 1977), and that they are more easily identified in mental images (Reed, 1974). The geometrical theory of part salience developed here provides a de facto starting definition of part salience, to be refined in the light of psychophysical experiments.

Our goal here differs in two respects from the description of shape provided by codons (Richards and Hoffman, 1985; Richards et al., 1986). A codon is a segment of a silhouette's outline which is (1) bounded on either side by a minimum of curvature and (2) assigned to one of six classes based on its maxima and inflections. A minimum of curvature can of course be either positive or negative in sign. So codons are not, in general, bounded by two negative minima of curvature and they do not, in general, correspond to parts of a shape. This is the first respect in which codons are not relevant to our undertaking here; we are studying parts, and codons do not, in general, correspond to parts. The second respect is this: the codon description is qualitative, assigning segments to one of six categories, whereas we here seek a quantitative account of the geometric factors affecting part salience.

Our goal here also differs from the interesting project of describing silhouettes in terms of causal processes (Leyton, 1987, 1988, 1989, 1992). This project uses curvature extrema and symmetry to infer a description in terms of four processes: protrusion, indentation, squashing, and internal resistance. It provides an account of the perceived genesis of shapes. We, however, seek instead quantitative factors that determine part salience.

Finally, our goal differs from the three-dimensional interpretations of silhouettes provided by Richards et al. (1987). To obtain these interpretations they use theorems regarding the geometry of smooth surfaces and their image projections, together with assumptions regarding general position and generic surfaces. Given a silhouette they find a small set of three-dimensional interpretations and decompositions into parts. However, they do not describe the salience of these parts, which is our concern here.

We view this paper as follows. We are searching a space of hypotheses about geometric factors which affect part salience. Part salience is a subject where there is as yet little empirical work to guide us. Thus our intention is to survey the territory, provide an initial map, and point to directions we think interesting. Some of the hypotheses are, we suspect, less plausible than others, and we indicate this when we discuss them. But we include them for more completeness in our survey of theoretical possibilities. We state the hypotheses precisely enough that they can be tested by psychophysical experiment. When we are done with our survey, we then select those hypotheses that we find most plausible, and present them as our theory of part salience. This theory suggests intriguing visual effects, and we close by demonstrating these effects and presenting psychophysical experiments.

2. The minima rule

Any subset of a shape could be considered one of its parts. But for the task of recognition not just any parts will do. The parts must satisfy certain principles (Hoffman and Richards, 1984; Marr and Nishihara, 1978; Sutherland, 1968). They must be computable from images (else they cannot be obtained), defined on any shape (else they will not help us recognize some shapes), and invariant under generic perturbations of viewpoint (else we shall see new parts each time we move, which would defeat indexing). These principles suggest that to define the part boundaries of a shape we should use its intrinsic geometry. They don't, however, dictate which aspects of this geometry to use.

The minima rule bridges this gap by appeal to one further principle: transversality. Since this principle is key to our account of part salience, we briefly review it here (for more details, see Guillemin and Pollack, 1974; Hoffman, 1983a,b; Hoffman and Richards, 1984). Consider Fig. 2. On the left are two shapes which, since they are separate in space, are distinct parts of the visual scene. On the right the two are joined to form a new object, and look like natural parts of this object. Transversality says that we can, almost surely,[1] find these parts

[1] To say that something is true ‘‘almost surely’’ is to say that it is true ‘‘everywhere, except possibly on sets whose measure is zero’’. See, for example, Guillemin and Pollack (1974) and Halmos (1950).

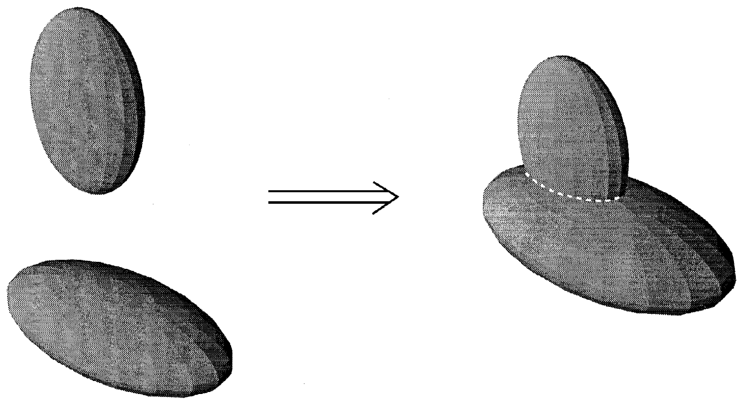


Fig. 2. A transversal intersection of two objects. Such an intersection leads to a concave crease at almost every point where their surfaces intersect.

by finding *concave creases*[2] in the surface of the new object. These creases are marked by the dashed contour in the figure. Concave creases point *into* the object, and differ from convex creases, such as the edges of a cube, that point *outwards*.

Transversality has broad application. It applies to the case shown in Fig. 2 where separate objects join to form a new object. But it also applies to other cases. For instance, it applies when one object extrudes from another, as when a finger grows out of a hand. In this case the transversal intersection appears, as we shall soon see, in smoothed form. Thus, since transversality is an abstract principle of mathematics, not an account of part genesis, it applies to any type of part genesis.

Transversality leads to the following rule for drawing part boundaries on 3D shapes:

Rule of concave creases: Divide 3D shapes into parts at concave creases.

The Schroeder staircase, shown in Fig. 3, illustrates this rule. At first glance one sees a staircase ascending to the left, with two dots on a single step, one on its rise and one on its tread. Each step is bounded on each side by a concave crease, as per the rule. Look a bit longer and the staircase suddenly reverses. The dots now lie on the rise and tread of distinct steps. You see a new set of steps, but again each step

[2] Technically, a crease on a surface is a set of points at each of which there are two tangent planes. One might wonder how differential notions, such as tangent planes and curvature, can be defined on real-world objects given that, strictly speaking, these notions don't apply: real objects exhibit different structures at different scales of resolution, and real images are discrete, not continuous. One can solve this problem by filtering images at various scales, and approximating each filtered image with piecewise continuous functions (see, for example, Koenderink, 1984; Witkin, 1983). Another solution to this problem is discussed by Hoffman (1983a). Here, and throughout, we will assume that some scale of resolution and some piecewise continuous approximation has been chosen. This has the effect, for instance, that our definitions of factors which determine part salience need not explicitly take into account wiggles at smaller scales. At the appropriate scale such wiggles are filtered out.

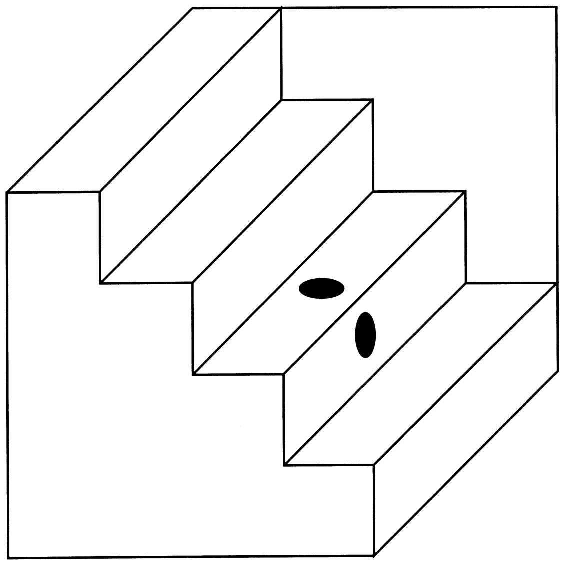


Fig. 3. The reversing staircase of Schroeder. The perceived organization into steps obeys the rule of concave creases, even when figure and ground appear to reverse.

is bounded on each side by a concave crease. The perceived reversal is in fact a switch of figure and ground. This switch turns concave to convex, and vice versa. And, per the rule, a boundary is seen at each (new) concave crease. Thus the new steps. For further examples see Hoffman and Richards (1984).

The rule of concave creases has obvious limitations. Smooth objects have no concave creases and yet can appear to have parts. An example is the cosine surface shown in Fig. 4. Here we see circular hills separated by valleys. The boundaries between one hill-shaped part and the next are marked by dashed contours. And these boundaries are smooth—not concave creases. Are these part boundaries seen only because we have drawn dashed contours? Not at all. If you turn the figure upside down you see new parts; the dashed contours which before sat between

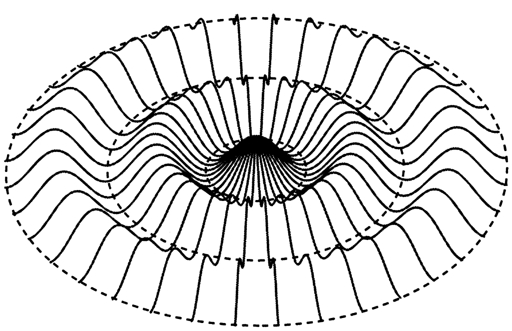


Fig. 4. The reversing cosine surface. We divide this surface into parts even though there are no concave creases.

hills now sit on top of hills. This suggests that human vision can divide smooth shapes into parts, and that prior experience with a given shape is not required to do so.

We can generalize the rule of concave creases. If one smooths a surface in the neighborhood of a crease, then each point of the crease becomes an extremum of curvature (Bennett and Hoffman, 1987). The idea of smoothing is apt if you consider the generation of many shapes in nature. For instance, the sharp articulations between bones in the body appear smoothed by the overlying muscles, fat, and skin. An example of smoothing is shown in Fig. 5. Here we see a side view of a step boundary on the Schroeder staircase. Next to it we see four successive smoothings of the boundary, with the curve closest to the step having the least smoothing. Notice that smoothing decreases the magnitude of curvature at the extremum.[3] This will be a key later to assessing the salience of part boundaries.

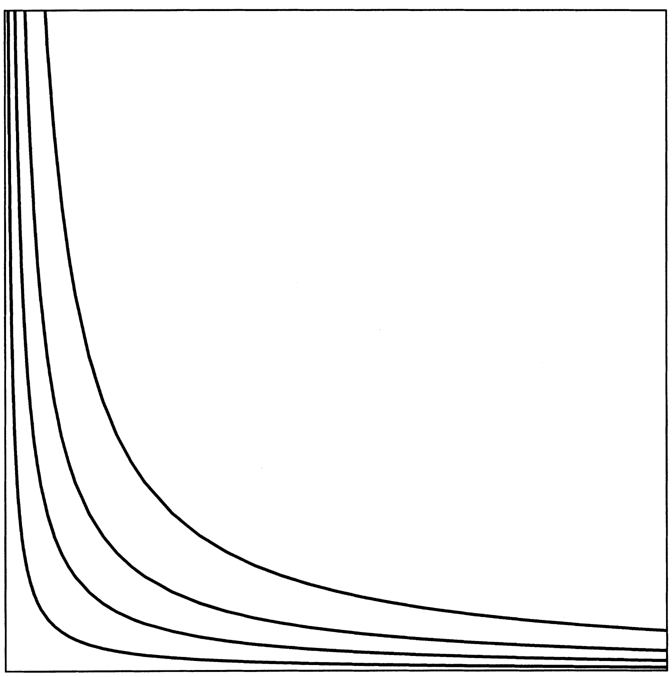


Fig. 5. Smoothing of a transversal intersection. A greater degree of smoothing leads to a lower magnitude of curvature at the curvature extremum.

[3] A detailed treatment of this is in Bennett and Hoffman (1987). But some brief remarks here might help. Using standard cartesian coordinates x and y for the plane, we can describe the neighborhood of the step boundary in Fig. 5 as the set of points where $f(x,y)=xy=0$; this ‘‘zero level set’’ of f is in fact the lines $x=0$ and $y=0$. We can describe successive smoothings of the boundary by zero sets of the functions $g(x,y)=xy-s$, where $s>=0$. Here s is a smoothing parameter. As s approaches 0, the zero sets of the functions g approach that of f. Indeed to generate Fig. 5 we plotted g with different values of s. One can show that, as s approaches zero, the magnitude of curvature at the negative minimum of curvature goes to infinity, that is, to that of the crease.

Intuitively, as shown in Fig. 5, the salience of a boundary increases as the magnitude of its curvature increases, that is, as it gets ''closer'' to the transversal case. The most salient boundaries are those with infinite curvature, that is, concave creases, which indicate unsmoothed transversal intersections. Thus, in short, smoothing a boundary decreases its salience. We shall be more precise about this shortly.

Smoothing is the key to a more general rule for defining part boundaries on shapes. But before we can state the rule, we must recall some concepts from the geometry of curves and surfaces (see, for example, Do Carmo, 1976; Hoffman and Richards, 1984).

In the case of plane curves we recall the concepts of normal vectors and curvature. These are illustrated in Fig. 6.

Normal vectors are unit length vectors pointing perpendicular to the curve. Observe that we can choose the normal vectors to be pointing to the left, as shown in Fig. 6, or to the right. Thus there are two distinct choices for the field of normals on a curve. There are also two distinct choices for the assignment of figure and ground on a curve. We will adopt a convention: the normals point to the side of the curve that is taken to be figure. Thus in Fig. 6, figure is to the left of the curve and ground is to the right.

The second concept is curvature. Intuitively, its magnitude at a point on a curve describes how quickly the normals change direction around that point. (Equivalently, the magnitude of curvature is the reciprocal of the length of the radius of the best fitting circle.) The sign of curvature is positive in those regions of the curve that are convex with respect to the choice of figure, and negative in regions that are concave. Thus the point labeled ''A'' has negative curvature, and is in fact

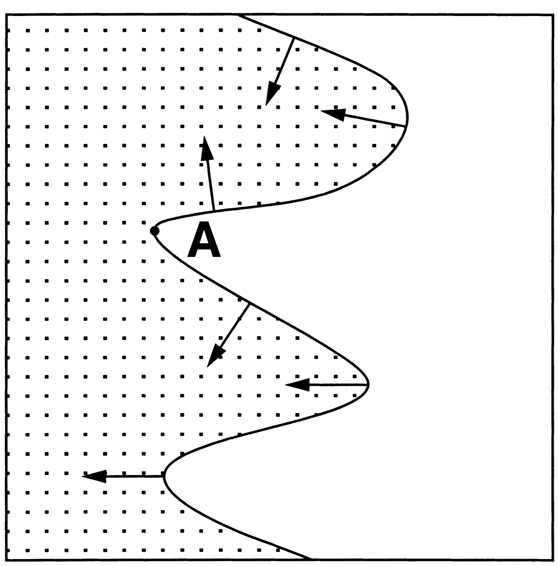


Fig. 6. Normals and curvature on a plane curve. We adopt the convention that the normals point to the side of the curve which is ''figure''.

a negative minimum of curvature. Were the field of normals drawn instead to the right, thereby changing the assignment of figure and ground as per our convention, the point ‘‘A’’ would be a positive maximum of curvature.

In the case of surfaces we recall the concepts of normal, principal curvatures, and lines of curvature. Normals are illustrated in Fig. 7(a). They are unit length vectors pointing perpendicular to the surface. Observe that, as with a curve, there are two distinct choices for the field of normals to a surface (except for strange surfaces like Moebius strips). We again adopt a convention: the normals point to the side of the surface that is taken to be figure. Thus in Fig. 7(a), figure is inside the cylinder, ground is outside.

Curvature on a surface is more complex than on a curve. This is illustrated in Fig. 7(b). Imagine standing on the cylinder at the point labeled ‘‘P’’. As you look around at the surface you find that the rate at which it curves away from you depends on which direction you look. It curves away fastest in the direction

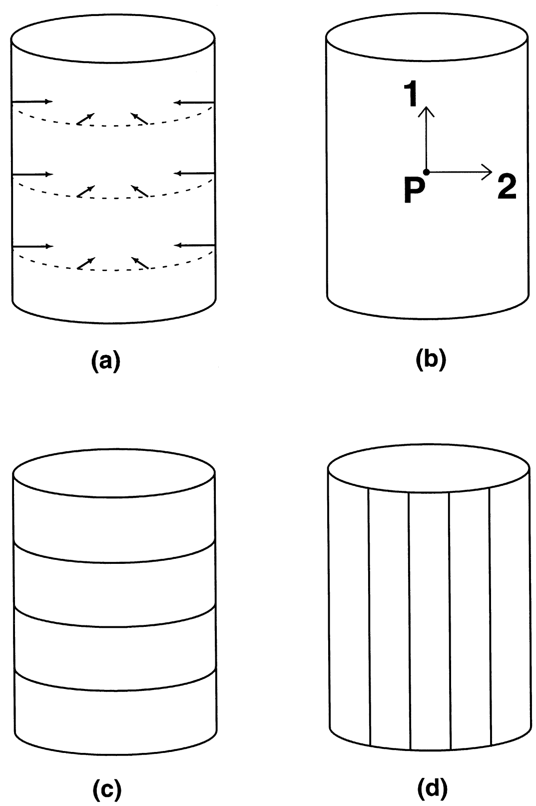


Fig. 7. (a) Normals on a surface. Again we adopt the convention that the normals point to the side of the surface which is ‘‘figure’’. (b) The principal directions of curvature for a point on a surface. At every point of any surface, the direction in which the surface curves most is always orthogonal to the direction in which it curves least. (c,d) Lines of curvature on a surface. The lines of greatest curvature are always locally orthogonal to the lines of least curvature.

labeled ‘‘2’’; it looks flat in the direction labeled ‘‘1’’. These two directions are, respectively, the directions of greatest and least curvature. They are called the principal directions, and the curvatures in these directions are called the principal curvatures. The 18th-century Swiss mathematician Leonhard Euler proved that the principal directions are always perpendicular to each other at every point on every smooth surface (for spheres and planes this statement is trivially true; for other surfaces it holds nontrivially).

Lines of greatest curvature are illustrated in Fig. 7(c). Each line is obtained by always moving in the direction of greatest curvature. Lines of least curvature, illustrated in Fig. 7(d), are obtained by always moving in the direction of least curvature. A principal curvature is positive in those regions of its line of curvature that are convex with respect to the choice of figure and ground. It is negative in those regions that are concave. So, for example, on bumps both principal curvatures are positive, on dents both are negative, and on saddles one is positive and the other negative. If figure and ground are reversed on a surface, bumps and dents change roles, and the signs of the principal curvatures change. With these concepts we can now state a more general rule for defining part boundaries on 3D shapes:

> *Minima rule:* All negative minima of the principal curvatures (along their associated lines of curvature) form boundaries between parts.[4]

The minima rule derives, as we have seen, from the principle of transversality and the application of smoothing. It can explain the parts we see in the cosine surface (Fig. 4). The cosine curves shown in this figure are in fact lines of curvature. (The other family of lines of curvature are circles running orthogonal to the cosine curves. These have no associated negative minima.) The negative minima of the principal curvatures along the cosine curves are indicated by the dashed contours. As the minima rule predicts, these are the perceived part boundaries. When you turn the figure upside down, your perceptual assignment of figure and ground reverses. As noted before, this entails that the signs of the principal curvatures must change. Negative minima become positive maxima and vice versa. Thus, according to the minima rule, the part boundaries must move to the new negative minima. And this is why we see the new organization into parts. More examples can be found in Hoffman (1983a) and Hoffman and Richards (1984).

The minima rule is a competence theory for part boundaries on 3D shapes. That is, the minima rule states in principle what these boundaries are, leaving open how in practice they may be computed from images despite noise and limited resources. The minima rule also motivates a competence theory for defining part boundaries on 2D silhouettes, by considering them as projections of the 3D case and using the

[4] A set of rules which generalize the minima rule are hinted at in Hoffman and Richards (1984) and spelled out in Beusmans et al. (1987). These handle ‘‘negative parts’’, such as a depression in a sphere, as well as the ‘‘positive parts’’ handled by the minima rule. Negative parts are beyond the scope of this paper. We restrict attention, therefore, to the minima rule.

tools of projective geometry (Beusmans et al., 1987; Hoffman and Richards, 1984). This theory motivates the following rule:

> *Minima rule for silhouettes:* For any silhouette, all negative minima of curvature of its bounding curve are boundaries between parts.

According to this rule, the plane curve in Fig. 6 has a part boundary at the point labeled ‘‘A’’. Switch figure and ground on this curve, and ‘‘A’’ is no longer a part boundary (it is now a positive maximum, not a negative minimum). This rule explains an effect described by Attneave (1971) and illustrated in Fig. 8. One simply draws a curve through the middle of a disk and separates the two halves. The halves look quite different, even though their wiggly bounding curves are, by construction, identical. The reason, according to the minima rule, is that the parts we see in one half differ entirely from the parts we see in the other. Figure and ground are switched in the left half as compared to the right. What are negative minima for the left are positive maxima for the right, and vice versa. Consequently the part boundaries are different, and the perceptual representations of the same curve in the two halves are different. They have different parts and hence they look different. (More examples of this phenomenon can be found in Baylis and Driver, 1995a,b; Braunstein et al., 1989; Driver and Baylis, 1995; Hoffman and Richards, 1984.)

In what follows it will be important to distinguish between *part boundaries* and *part cuts* (Beusmans et al., 1987). This is illustrated in Fig. 9 for the case of silhouettes. In part (a) is shown the outline of a silhouette with its part boundaries marked. In (b) and (c) are two possible part cuts, which connect the part boundaries in two different ways. As the figure illustrates, part *boundaries* on silhouettes are isolated points, whereas part *cuts* on silhouettes are curves that partition the silhouette. In the case of 3D objects, part *boundaries* are curves (Fig. 9(d)) and part *cuts* are surfaces (Fig. 9(e)). The minima rule defines part

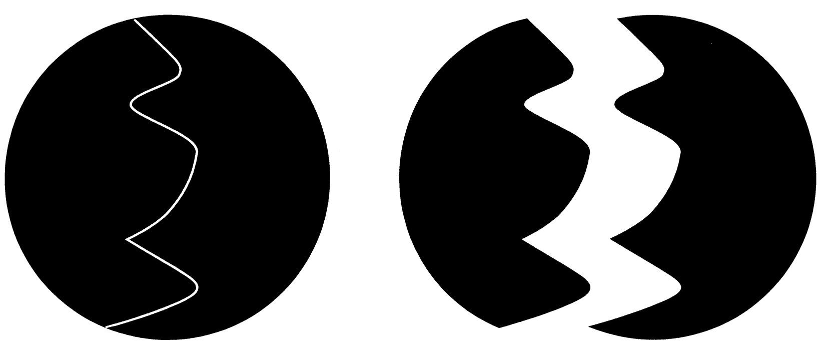


Fig. 8. Attneave’s (1971) disk. In accordance with the minima rule, the two sides are organized differently into parts. Thus they appear very different, despite the fact that the bounding curves are, by construction, identical.

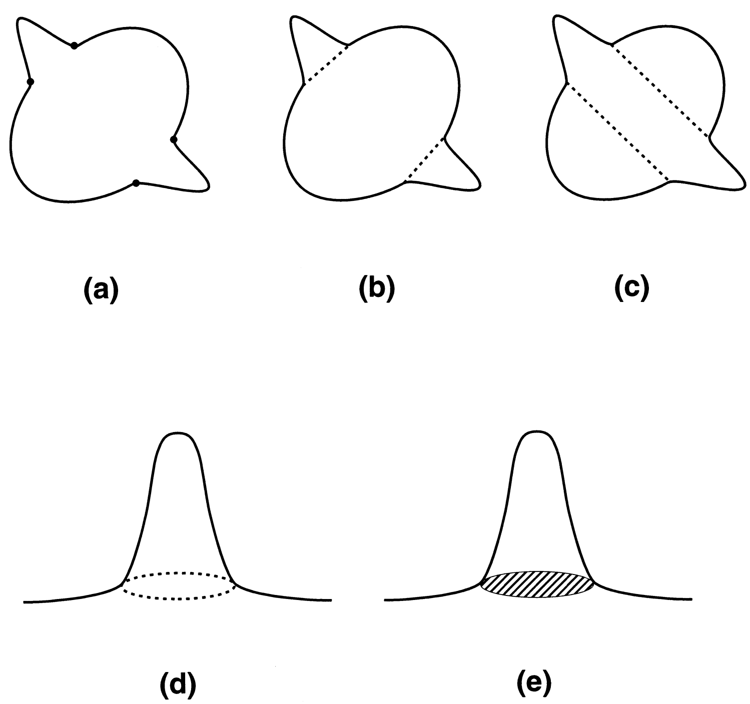


Fig. 9. (a–c) Part boundaries and part cuts on a silhouette. Part boundaries are isolated points (as shown in a), whereas part cuts are plane curves which partition the silhouette (as in b and c). The minima rule defines part boundaries on a silhouette, but does not dictate, in general, how these part boundaries are to be joined to form part cuts. (d,e) Part boundaries and part cuts on a 3D shape. Part boundaries are space curves (as shown in d) and part cuts are surfaces which partition the 3D shape (as shown in e).

boundaries, not part cuts. However, it constrains how parts can be cut by requiring cuts to pass through the boundaries it defines. But that is all it says about part cuts. It does not, for instance, say that a part corresponds to a segment between two consecutive, negative minima along the shape's outline (cf. Kurbat, 1994; Leyton, 1992; Siddiqi and Kimia, 1995). Instead, to specify part cuts precisely a separate theory is required (Beusmans et al., 1987; Siddiqi and Kimia, 1995; Singh et al., 1996). Such a theory will use the boundaries defined by the minima rule, plus other boundaries that might be required, and will interpolate cuts between these boundaries.

For example, Siddiqi and Kimia (1995), building on their earlier work (Kimia et al., 1991, 1992; Siddiqi et al., 1996), propose two kinds of part cuts, namely, ''limbs'' and ''necks''. They define a limb as ''a part-line going through a pair of negative curvature minima with co-circular boundary tangents on (at least) one side of the part-line'' (p. 243). Their ''part-line'' is what we call a ''part cut''. Two tangents are ''co-circular'' if and only if they are both tangent to the same circle (Parent and Zucker, 1989, p. 829). Siddiqi and Kimia (1995) define a neck as ''a part-line which is also a local minimum of the diameter of an inscribed circle'' (p. 243). Investigating the theoretical and empirical implications of these definitions is an interesting direction for further research.

3. Salience in 2D: the case of cusps

Having reviewed the minima rule, we now study part salience. We start, for simplicity, with silhouettes. Moreover, to simplify even further, we assume that the only part boundaries are cusps (i.e., sharp points). Such silhouettes can be represented by plane curves that are closed (i.e., no gaps or loose ends), simple (i.e., no self intersections), oriented (i.e., traced in a definite direction), and whose only negative minima of curvature are concave cusps (which have curvature of infinite magnitude). We later consider silhouettes with smooth part boundaries.

3.1. Boundary strength

The salience of a part of an object depends, we propose, on its size relative to that of the whole object, the degree to which it protrudes from the object, and the strength of its part boundaries. This sounds plausible enough. And indeed the intuitions here go back at least as far as Gestalt psychologists such as Rubin (1958). The point of this paper is not just to give new intuitions, but also to make old intuitions precise—to cast them as precise hypotheses that can be tested by psychophysical experiments. In this section we begin this process by studying the strength of part boundaries that are concave cusps.

Two geometric properties determine this strength: the turning of normals at the cusp and the cusp type. We consider first the turning of normals.

Again we start with transversality. If two objects interpenetrate at random then, almost surely (refer to footnote 1), at every point where they intersect there will be a concave crease. If, at some point, the objects do not meet transversally then their surfaces, at that point, are perfectly aligned. This is illustrated in Fig. 10(a). On the left are the silhouettes of two objects. On the right they intersect nontransversally at every point, so that the silhouette is smooth at every point where they meet.

Now suppose that we rotate one of the objects, so that the surfaces are barely misaligned as shown in Fig. 10(b). We now have a transversal intersection and, in consequence, a concave crease. Thus we have a part boundary. But it is not a strong boundary, for the crease is not pronounced. We are still very close to the nontransversal case in which the surfaces are aligned. We can alter this by rotating one of the objects still further, so that the misalignment of the two surfaces is more pronounced, as in Fig. 10(c). This takes us further from the nontransversal case and thereby increases the salience of the part boundary.

A natural way to measure this aspect of salience is by means of normal vectors (introduced in Fig. 6). Fig. 10(d) shows a transversal intersection, with the region of interest circled. In Fig. 10(e) this region is blown up, and normal vectors to the two curves at their point of intersection are shown. Recall that, for a given choice

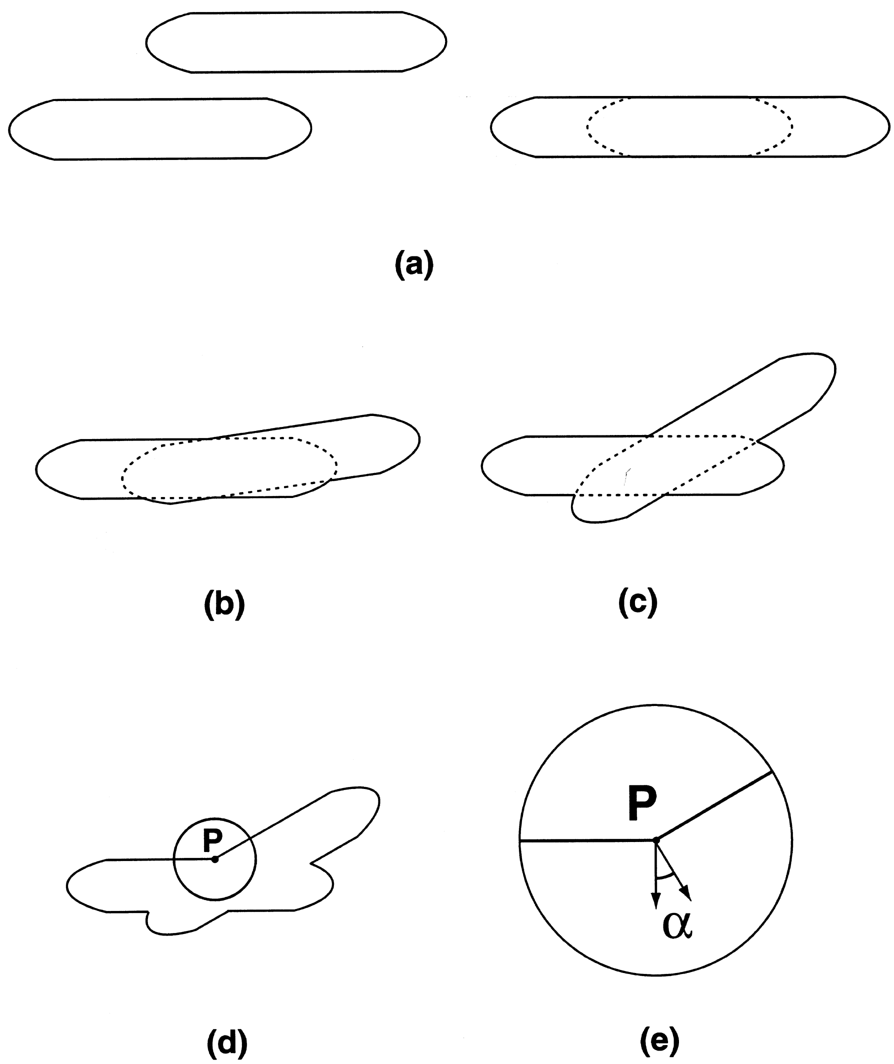


Fig. 10. (a) A nontransversal intersection of two silhouettes. Such an intersection is smooth. (b,c) A sequence of increasingly pronounced transversal intersections. (d,e) The turning angle at a transveral intersection. The greater this angle, the more pronounced is the intersection.

of figure and ground, each point on a smooth curve has one normal vector. But, as Fig. 10(e) shows, a cusp point is special in that it has two normal vectors (one obtained by approaching the cusp from the left and the other by approaching it from the right). The *turning angle* between these vectors measures the departure from the nontransversal. The turning angle is 0 degrees for a nontransversal intersection. It grows as the objects rotate away from the nontransversal case. This motivates the following hypothesis for the salience of part boundaries.

Hypothesis of turning normals: The salience of a cusp boundary (and the salience of the parts for which the cusp is a boundary) increases as the magnitude of the turning angle between its two normal vectors increases.

A more precise understanding of this hypothesis involves three geometric concepts: the orientation of a curve, the figure–ground assignment for a curve, and the normal map. The orientation of a curve is the direction in which it is traced. Thus a curve has two possible orientations. There are also two possible assignments of figure and ground along a curve. As mentioned before, we adopt the convention that the choice of orientation and the choice of figure and ground are linked, such that figure is always to the left of the curve, and ground to the right, in the chosen orientation. Fig. 11(a,b) shows a curve with its two possible orientations (indicated by double arrow heads) and two corresponding choices of figure and ground (indicated by stippled figure and unstippled ground). Observe that, in order to have a consistent physical interpretation, a plane curve must have a globally consistent orientation. There must be no points where the orientation of the curve reverses. If the orientation reverses, then the figure–ground assignment reverses and the curve cannot be consistently interpreted as the bounding contour of a single object. We highlight this obvious fact in the following principle:

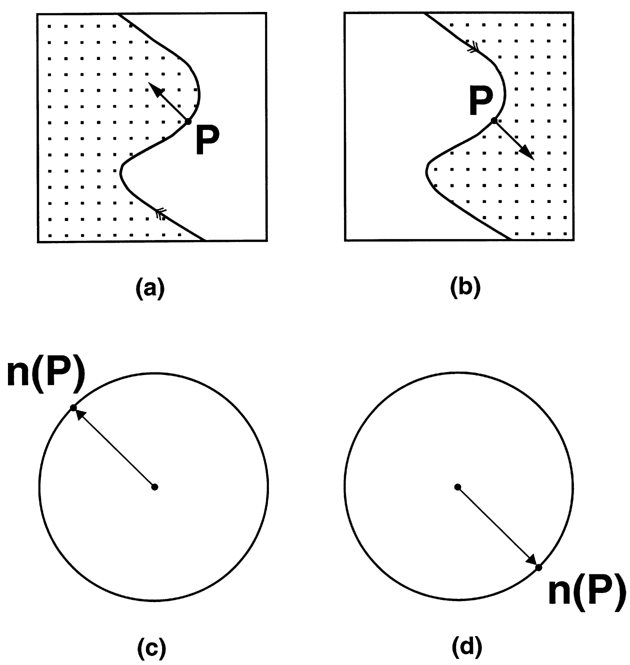


Fig. 11. (a,b) Orientation and figure/ground assignment on a plane curve. We adopt the convention that the orientation is chosen so that figure is to the left. (c,d) The normal map for a plane curve. The normal map takes each point on a plane curve to that point on the unit circle which has the same normal vector direction.

Principle of consistent orientation: To have a consistent physical interpretation, a plane curve must have a globally consistent orientation (i.e., it must have no points where its orientation reverses).

Physical consistency, as we shall see, is not invariably respected by human vision. Other principles may override it and lead to physically inconsistent interpretations.

The normal map takes points on a plane curve into points on the unit circle in a simple fashion. Think of each point of the unit circle as indicating a *direction*, namely the direction from the center of the circle to that point. The unit circle then represents all possible directions in the plane. Now each point on a smooth oriented curve also has a canonical direction, namely the direction of its normal vector. The normal map simply takes each point of the plane curve to that point on the unit circle representing this canonical direction. Reversing the orientation of the curve changes this direction by 180 degrees. Fig. 11(c,d) shows the normal map of point P (of Fig. 11(a,b)) in each of its two possible orientations.

Fig. 12(a,b) illustrates that as one traverses a curve through a region of positive curvature (e.g., from point A to point B) the normal turns counterclockwise. We shall call this a positive turning, or a turning through a positive angle. Through a region of negative curvature (e.g., from point B to point C) the normal turns clockwise. We shall call this a negative turning, or a turning through a negative angle. And at points where curvature changes sign, the normal reverses its turning direction.

At a cusp the normal map jumps, as shown in Fig. 12(c,d). In this case there is a two-way ambiguity in specifying the angle of the jump: as a turn clockwise or counterclockwise. One angle will always be less than 180 degrees, the other always greater. The smaller angle is the one of interest. It is the angle between the surfaces of two objects whose interpenetration would form the visible cusp. The larger angle has no such geometric or physical significance. Thus we stipulate that *turning angles* at cusps must have magnitudes less than 180 degrees, with positive values corresponding to counterclockwise turns and negative values corresponding to clockwise turns. We highlight this in the following principle:

Principle of least turning: Turning angles at cusp boundaries must have magnitudes less than 180 degrees.

The principle of least turning has certain consequences for simple closed plane curves with cusps. In particular, the total turning of the normal (summed over both smooth portions and cusps) as one completely traverses such a curve once is precisely 360 degrees. This provides a global error check for all the local turnings that one computes while traversing the curve.

For a nontransversal intersection, the turning angle is 0 degrees (the smallest possible turning angle) and, according to the turning of normals hypothesis, the salience is zero (Fig. 10(a)). Rotating slightly one object leads to a cusp with

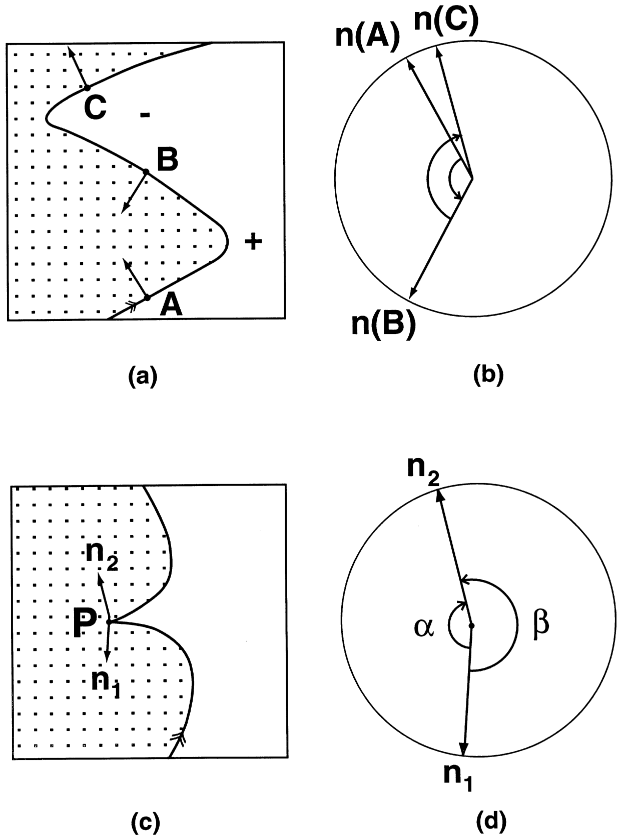


Fig. 12. (a,b) Positive and negative turnings of a plane curve. Counterclockwise turnings are positive, clockwise turnings are negative. (c,d) Jumps of the normal map at cusp points. A cusp point has two normals. These two normals determine two distinct angles, whose sum is 360 degrees. The turning angle between the normals is the smaller of these two angles.

turning magnitude slightly greater than 0; the salience is now small, but not zero (Fig. 10(b)). Rotating much further leads to a cusp with turning magnitude much larger than 0; the salience is now much larger (Fig. 10(c)).

The turning of normals hypothesis is supported by several demonstrations involving ambiguities of figure and ground. But to interpret these demonstrations as we would like requires an additional hypothesis relating part salience with the choice of figure and ground.

> *Hypothesis of salient figures:* Other things being equal, that choice of figure and ground is preferred which leads to the most salient parts for the figure object.

To understand this hypothesis, it helps to recall that many factors affect one's perception of figure and ground. There are cognitive factors, such as attentional focus or verbal suggestion. There are low-level factors, such as symmetry (Bahnsen, 1928; Hochberg, 1964), size and contrast (Rubin, 1958; Koffka, 1935), and convexity[5] (Kanisza and Gerbino, 1976; Stevens and Brookes, 1988). Part salience, according to the hypothesis of salient figures, is one of many such factors. Since part salience, like convexity and symmetry, is determined by geometry, it too may be a low-level factor. Indeed experiments that we discuss shortly suggest that it is. If so, the hypothesis of salient figures suggests that human vision proceeds as follows in its early processing of a visual contour (see Fig. 13(a)). First it finds all points whose magnitude of curvature is locally maximal. Then it uses these extrema to find competing partitions of the contour (Fig. 13(b)). Finally it chooses that assignment of figure and ground which leads to the most salient parts for the figure side (Fig. 13(c)). The assignment of figure and ground obtained in this fashion might not be globally consistent (Fig. 13(d)). This account differs from that of Baylis and Driver (1995a), (1995b), who propose that shape descriptions are computed only after figure and ground are assigned, and only for the figure.

An example of the role of this hypothesis is shown in Fig. 14(a). One can see

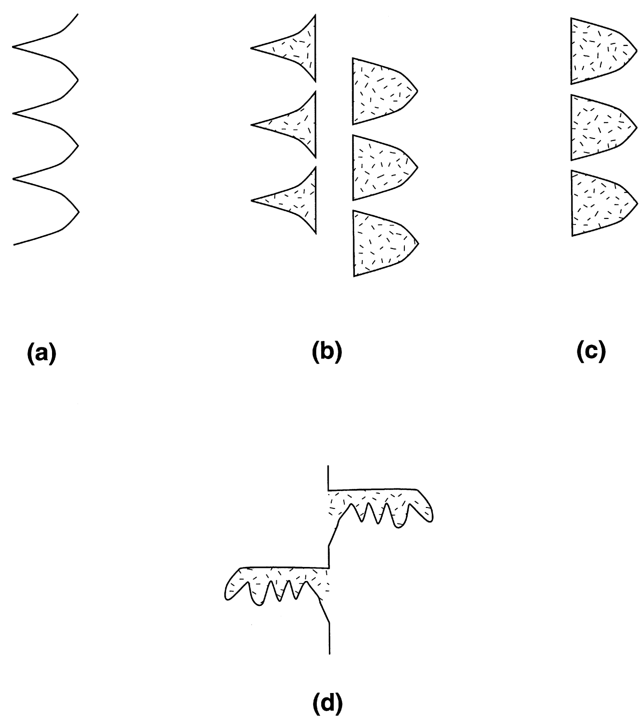


Fig. 13. Part salience as a determiner of the assignment of figure and ground.

[5] Here we mean the technical sense of convex and concave: A shape is convex if each pair of its points is connected by a line segment which never goes outside the shape; otherwise it is concave.

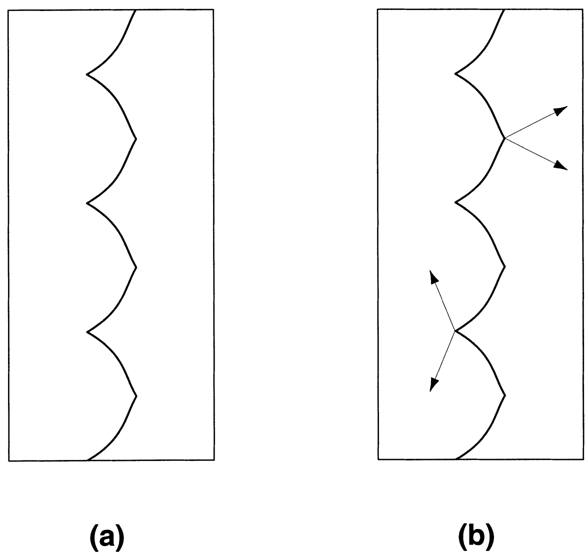


Fig. 14. (a) Bullets and tepees. Bullets are the more salient parts. (In this and in subsequent figures we use just a curve drawn against a white background, instead of coloring one side of the curve, to avoid any bias that such color contrasts might induce on the perception of figure and ground.) (b) Turning of normals for the bullets and tepees. Bullets are more salient, in part, because they have the greater turning of normals at their boundaries.

this figure either as bullets pointing right or as tepees pointing left. Bullets, for most observers, are the dominant interpretation (see Experiment 3 in Section 8). This cannot be due to such factors as symmetry and contrast, since these are equal in the two interpretations. It could be due in part to convexity, since the bullets have convex tips and the tepees concave (refer to footnote 5). It could also be influenced by the turning of normals since, as shown in Fig. 14(b), the part boundaries for bullets have greater turning of normals than those for tepees, and therefore are more salient. Moreover, the bullets have larger area than the tepees which, as we shall see, also contributes to their greater salience.

The turning angle at a cusp is not the only factor determining its salience as a part boundary. Also important is the cusp *type*. There are 12 basic types of cusps (six types noted by Stevens and Brookes, 1988, plus their figure–ground reversals; see also Brady and Asada, 1984, Fig. 12). These are illustrated in Fig. 15. The first six types are part boundaries; the last six are part tips. Each cusp can be distinguished by the sign of curvature of each of its two arcs and by the choice of figure and ground. Following Stevens and Brookes, we label these cusps as follows. Type 1 has two arcs of positive curvature, and is therefore labeled $+/+$. Type 2 has one arc of zero curvature (a straight line) and one of positive curvature, and is therefore labeled $0/+$. For analogous reasons, type 3 is labeled $0/0$, type 4 $-/+$, type 5 $-/0$, and type 6 $-/-$. Types 7 through 12 are the figure–ground reversals, respectively, of types 1 through 6 (so the respective signs of curvature are flipped). Stevens and Brookes note that, of the first six cusp types, type 1 cusps

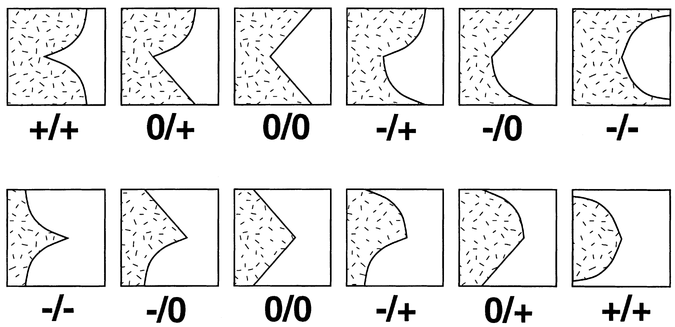


Fig. 15. The six cusp types (Stevens and Brookes, 1988), plus their figure–ground reversals. The cusp type of a boundary affects its salience.

are ‘‘most suggestive of a part boundary’’ and that each successive type is less suggestive. The reason, they say, is that ‘‘intersecting convex silhouettes produce concave cusp discontinuities’’ (i.e., type 1 cusps) at each point of intersection. Put simply, human vision prefers figure–ground assignments that lead to figures with convex parts, and type 1 cusps are strong (though not perfect) indicators of convex parts. Stevens and Brookes (1988) give experimental evidence that the type 1 cusp is most salient as a part boundary. Thus cusp type, in addition to turning angle, seems to be a factor in determining boundary salience. Both factors are, as one should wish, scale invariant.

3.2. Size and protrusion

The salience of a part depends not only on the strength of its boundaries, but also on its area relative to the whole object. The importance of relative area for visual salience has long been recognized by Gestalt psychology (e.g., Rubin, 1958), and more recently by research into the notion of scale (e.g., Koenderink, 1984; Kimia et al., 1990; Siddiqi and Kimia, 1995).

For a part defined by two part boundaries, the area is that of the region bounded, on one side, by a curve joining the boundaries and, on the other, by the outline of the silhouette (see Fig. 16(a)). The curve joining the boundaries is the *base* of the part, and constitutes the *part cut*. As mentioned earlier, there is as yet no general theory of part cuts and therefore no theory of the exact geometry of these part bases. In the absence of such a theory, we take part cuts to be straight lines joining part boundaries. In some cases, such as the elbow of Fig. 16(b), points in addition to negative minima of curvature may have to serve as part boundaries. If a part is defined by more than one part cut, as is the circular part in Fig. 16(c), the area of the part is defined as the area of the region bounded by (1) the part cuts and (2) any portions of the silhouette outline which delimit the part.

The relative area of a part can be quantified as the ratio of its visible area to the visible area of the whole object. The larger this ratio the more salient, one would predict, is the part. We highlight this in the following hypothesis:

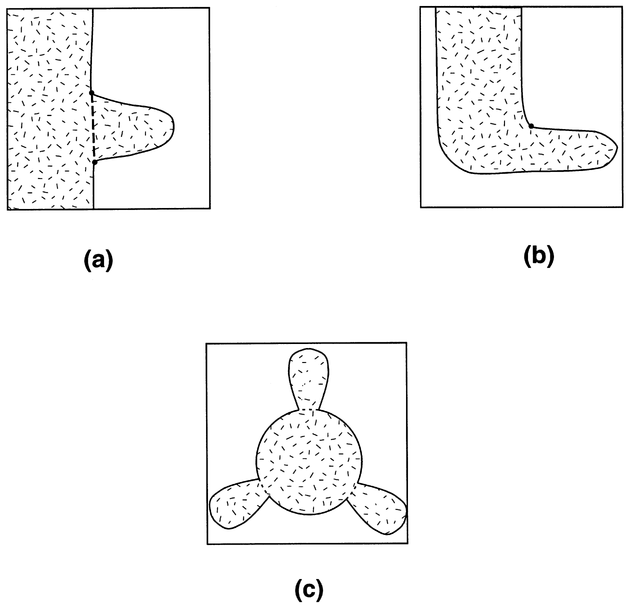


Fig. 16. Part bases. (a) The base is a curve (indicated by dashes) joining two negative minima part boundaries. (b) The base sometimes must join a negative minimum of curvature with another point which is not a negative minimum. (c) Some parts, like the central part of this figure, have more than one part cut and therefore more than one base.

> *Hypothesis of relative area:* The salience of a part increases as the ratio of its visible area to the visible area of the whole silhouette increases.

The relative area is invariant under translations, rotations, and uniform scalings in the plane. It is not, in general, invariant if the 3D object that gives rise to the silhouette is rotated in depth, or if this 3D object is partially occluded. Thus, if this hypothesis is correct, one predicts human judgments of part salience to share these same properties. Some examples are shown in Fig. 17. In Fig. 17(a) is shown the silhouette of an object with a prominent part. In Fig. 17(b) the silhouette has been translated, rotated in the plane, and uniformly scaled. The hypothesis of relative area predicts that subjects will see no change in the salience of the part. In Fig. 17(c) the object has been rotated in depth and in Fig. 17(d) partially occluded. For these changes the hypothesis of relative area predicts that subjects will see the part as less salient.

There are, of course, other proposals one might consider. For instance, the salience of a part might depend on its absolute retinal area or on its absolute apparent area, unscaled by that of the whole object. But these hypotheses seem unlikely, especially in light of Fig. 17.

Area and boundary strength are apparently not the only determinants of part salience. Two parts with the same relative area and same boundary strengths may,

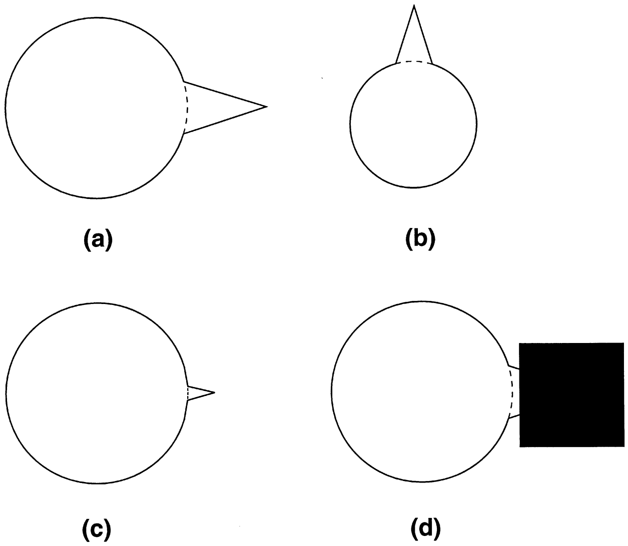


Fig. 17. The effects of relative size on the salience of a part.

nonetheless, vary in perceived salience. Another factor (at least) is involved. Intuitively, this factor is the degree to which a part ''sticks out'' from its object. Parts that stick out more seem to be more salient.

Several quantifications of ''sticking out'' are possible. For instance, one might compute the inertia tensor of the part, and take the magnitude of that eigenvector associated with the smallest eigenvalue. This gives a measure of the principal elongation of the part. But unfortunately this measure is not scale invariant. One might make it so by taking the ratio of the magnitudes of the two eigenvectors. But a large value of this ratio does not distinguish short and wide parts (which do not ''stick out'' much) from those long and narrow (which do). The same problem defeats a measure of protrusion computed using the ratio of the square of the perimeter of the part to its area.

Siddiqi and Kimia (1995) propose a measure of ''sticking out'' that they call ''extent''. They compute it as follows: ''extent is computed as the ratio of major to minor axes of the largest ellipse that can be inscribed in the shape, with the part-line as its minor axis, but restricted to lie within the part-line's local neighborhood *N*'' (p. 247). Their ''part-line'' is what we call a ''part cut'', and the local neighborhood *N* is a disk centered on the part cut and having a diameter ''somewhat larger'' than the part cut. However for most parts it is not possible to inscribe an ellipse with the property that the part cut is the minor axis of the ellipse.

A different measure which captures what we want is the ratio of the perimeter of the part (excluding the base) to the length of the base. This is shown in Fig. 18(a). (Again, we take the base to be the straight line joining the two boundary points which define the part. The perimeter is the arc length of the bounding curve

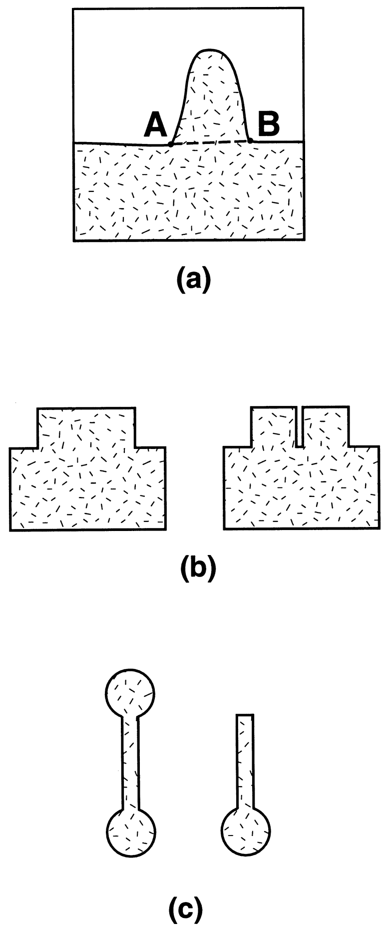


Fig. 18. Part protrusion measured as the ratio of perimeter to base(s).

of the silhouette between the two boundary points.) In the general case a part may have more than one base, as in Fig. 16(c). Therefore we define the protrusion of a part to be the ratio of its perimeter length (excluding all its bases) to the *sum* of all its base lengths. This definition has several nice properties. First, it is scale invariant. Second, it distinguishes parts that are long and narrow from those short and wide. Third, as illustrated in Fig. 18(b), it has the consequence that the protrusion of the large bottom part remains roughly the same if the top part on the left figure is divided into two parts as shown in the right figure. Fourth, as illustrated in Fig. 18(c), it has the consequence that the protrusion of the middle part of the dumbbell on the left is half that of the corresponding part on the right figure. Fifth, as illustrated later in Fig. 30(a), the protrusion is zero for a part (the central part of Fig. 30(a)) that meets the bounding curve of the silhouette only in isolated points. Given this definition of protrusion, we consider the following hypothesis:

> *Hypothesis of silhouette protrusion:* The salience of a silhouette part increases as its protrusion increases.

4. Salience in 2D: smooth boundaries

The definitions of protrusion and relative area do not change if the part boundaries are smooth. However, the analysis of boundary strength does change, and to this we now turn our attention.

For a silhouette, each negative minimum of curvature on its bounding curve is, by the minima rule, a part boundary. The primary difference between a smooth boundary and cusp boundary is that the cusp has two normals, whereas the smooth has, by definition, only one. Therefore the hypothesis of turning normals, which measures the salience of a cusp boundary by the angle between its two normals, cannot be used for the salience of smooth boundaries.

We need therefore a new method for measuring the salience of smooth boundaries, one analogous to the turning of normals for cusps. Differential geometry provides, of course, a standard method to describe the turning of normals at smooth points, namely, the *curvature*. But curvature is not scale invariant (e.g., its value doubles if the figure shrinks by half). Thus curvature does not provide, for our purposes, an appropriate measure of turning. This problem, however, can be solved.

To do so we need more than just local (differential) properties of the curve, and we need less than its global properties. We need something in between. We need properties of how the curve evolves in an appropriate region nearby (but not just infinitesimally nearby) a negative minimum of curvature. This region we call a *locale*.

How shall we define a locale? It should be a region over which we obtain a measure of turning that is stable, in the sense that small changes in size of the locale yield small changes in measured turning. This suggests the following. Consider a smooth boundary with inflections of curvature on both sides (see Fig. 19(a)). Inflections are, as you see in the figure, regions where the normal turns little. So if we measure the turning between these inflections, we get a stable result. Therefore we take the *locale* of a part boundary to extend to the nearest inflection on each side. A locale is *proper* if it contains no positive maxima of curvature. We are led to the following:

> *Hypothesis of locale turning:* The salience of a smooth boundary increases as the turning of normals in its (proper) locale increases.

By ‘‘turning of normals in its locale’’ we mean the turning of the normal from the first inflection point defining the locale to the second (traversing the curve in

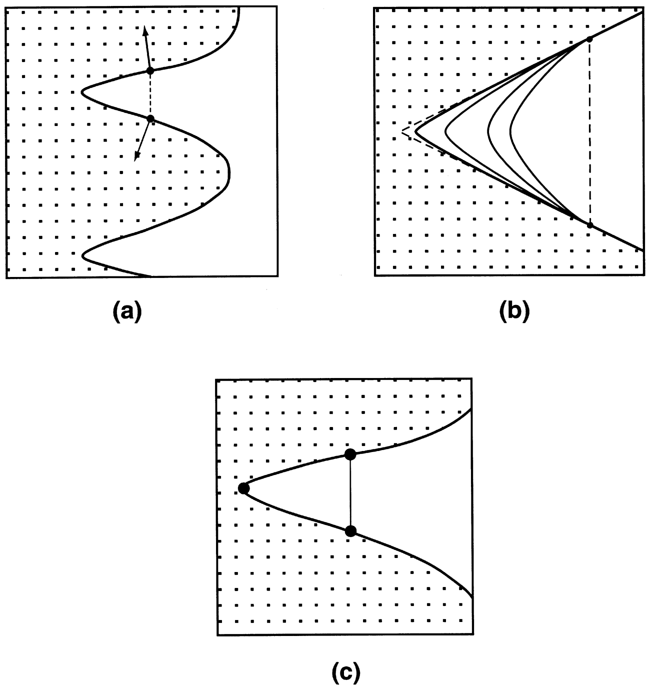


Fig. 19. (a) The locale of a part boundary. A canonical way to measure the turning of normals about a smooth boundary is from the nearest inflection on one side to the nearest inflection on the other. (b) Different part boundaries having the same locale turning. Clearly locale turning is not the only factor determining the salience of smooth part boundaries. (c) The normalized magnitude of curvature at a part boundary. Since curvature is not a scale-invariant quantity, we can make it scale invariant by multiplying by the length of a canonically chosen chord.

the direction appropriate to the choice of figure and ground). This hypothesis is an analog of the hypothesis of turning normals for cusp boundaries.

One can construct examples of boundaries for which locales do not exist or are not proper. For these one might use something like the *curvature contrast* at the boundary rather than locale turning. The curvature contrast at a point is the magnitude of the second derivative (with respect to arc length) of curvature at that point.

As can be seen in Fig. 19(b), the hypothesis of locale turning does not capture all variation in a curve that might influence the salience of a boundary. Smooth boundaries can differ in magnitudes of curvature, yet all have the same locale turning. Undoubtedly these boundaries differ in salience: the higher curvature are more salient than the lower. One reason, as we noted before, is that higher curvature boundaries better approximate the ideal transversal boundary (Fig. 5). Moreover their curvature differs more from that of neighboring points on the curve (Fig. 19(b)) and is therefore more easily detected; this is the intuition captured by curvature contrast.

Curvature varies with scale. One can, however, transform curvature into a scale-invariant quantity, as shown in Fig. 19(c). First find the chord joining the two inflections which define the locale. Then multiply the curvature at the boundary by the length of this chord. The resulting *normalized curvature* does not change with scale (e.g., if the figure shrinks to half size, the curvature doubles but the chord halves, so their product is constant). It is this normalized curvature which, we suggest, captures those changes in boundary salience missed by locale turning.

> *Hypothesis of normalized curvature:* The salience of a smooth boundary increases as the magnitude of normalized curvature at the boundary increases.

This hypothesis implies that human subjects are sensitive to changes in curvature. There is evidence for such sensitivity in the perception of 2D curves (Triesman and Gormican, 1988; Wilson and Richards, 1985, 1989; Wolfe et al., 1992), 3D surfaces presented in structure-from-motion or motion parallax (Cornilleau-Peres and Droulez, 1989; Koenderink, 1986; Norman and Lappin, 1992; Norman and Todd, 1993; Rogers and Graham, 1983; Saidpour and Braunstein, 1994; Todd, 1984; Todd and Norman, 1991), and 3D surfaces presented in stereo or motion parallax (Rogers, 1986; Rogers and Cagenello, 1989; Rogers and Collett, 1989; Rogers and Graham, 1983).

Why do we need both a measure of locale turning and a measure of normalized curvature to characterize the strength of a smooth boundary? Because the two can vary independently. As we have seen in Fig. 19(b), different curves with the same locale turning can have different normalized curvatures. Similarly, different curves with the same normalized curvature can have different locale turnings.

Locale turning, normalized curvature, and curvature contrast are distinct from ‘‘total curvature’’ and ‘‘curvature disparity’’, two notions discussed by Siddiqi and Kimia (1995). They define total curvature as follows: ‘‘Total curvature is the actual amount a tangent at one negative curvature minima [sic] has to bend to align with the tangent at the second negative curvature minima [sic]’’ (p. 244). In other words, total curvature is computed along a part *cut*, whereas locale turning, normalized curvature, and curvature contrast are computed along the part *silhouette*. So the notions are unrelated. Siddiqi and Kimia (1995) describe curvature disparity as ‘‘a measure of how ‘thin’ the neck is’’ (p. 244). They do not define what the measure is, but it is clear that their notion is distinct from that of normalized curvature and curvature contrast.

5. Salience in 3D: the case of concave creases

We now consider the salience of parts of 3D objects. Many ideas discussed for silhouettes still apply, but with added complexity. In this section we consider parts whose boundaries are concave creases. Later we consider parts whose boundaries are smooth.

5.1. Boundary strength

For silhouettes, each cut intersects the outline in isolated boundary points. If such a point is a cusp then it has, by the principle of least turning, a unique turning angle. This simplifies the analysis of boundary salience. But on surfaces each part boundary is a continuous, one-dimensional family of points—that is, a curve. Thus, whereas there is a unique turning angle for each cusp boundary on a silhouette, there is an entire one-dimensional family of *possibly different* turning angles for each crease boundary on a surface. Each point on a crease has two normal vectors (see Fig. 20). The angle between these normals is the turning angle at that point.[6] Clearly, in the general case, this angle can change from point to point along a crease.

However, in the simplest case all points on the crease have the same turning (as in Fig. 20). Such a crease is *uniform*. For uniform creases we consider the following.

> *Hypothesis of turning normals for uniform creases:* The salience of a crease boundary which is uniform grows as the magnitude of its turning angle increases.

Nonuniform creases are not as simple. We must combine *different* turning angles along the crease into measures of boundary salience. Precisely how will

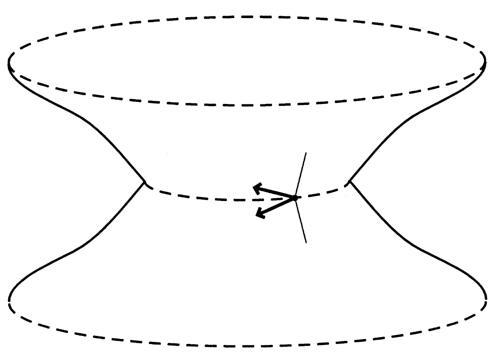


Fig. 20. Turning angles on a surface boundary defined by a concave crease. There is an entire one-dimensional family of possibly different turning angles. This complicates the measurement of boundary salience.

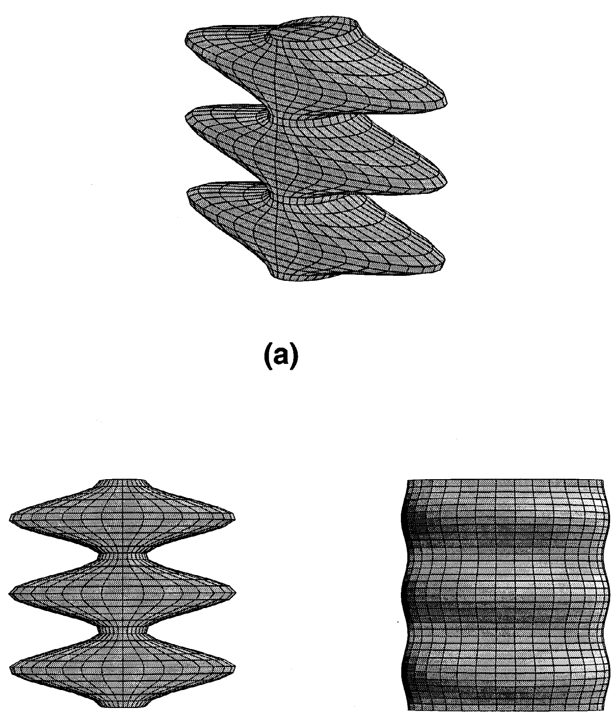
[6] Denoting the two normals by $\boldsymbol{n}$ and $\boldsymbol{m}$, the desired angle, t, can be computed easily by the formula $\mathrm{Cos}\, t = (n.m)/(|\boldsymbol{n}|\ |\boldsymbol{m}|)$, where $\boldsymbol{n}$ and $\boldsymbol{m}$ are now three-dimensional vectors (with the constraint, from the principle of least turning, that $|t| < 180$ degrees).

depend, in part, on the nature of our representation of the shape—in particular, on whether it is viewer dependent or viewer independent.

If it is viewer independent, then we can compute measures of boundary strength that do not vary with viewpoint. A natural such measure is the *mean turning angle*, averaged over the entire boundary; another measure is the *maximum turning angle* over the entire boundary. This latter measure would model the case in which the overall salience of a part boundary is determined by its most salient portion.

It seems likely, however, that the salience of a boundary depends on viewpoint. Consider, for instance, Fig. 21(a). This is a surface of revolution, flattened so that any horizontal cross-section is an ellipse. Fig. 21(b),(c) shows it from two orthogonal views. The strengths of the boundaries of the middle part clearly differ in these two views. So if our representation of shape is viewer centered (like Marr's $2\frac{1}{2}$-D sketch), then we expect measures of boundary strength that vary with viewpoint. In this case, we consider the following two hypotheses.

Hypothesis of visible mean (maximum) turning: The salience of a crease



(a)

(b) **(c)**

Fig. 21. A flattened surface of revolution, given parametrically by ((3+.1 Cos[z]) Cos[t], 4 (1.5+ Cos[z]) Sin[t],z). The salience of the parts depends on the direction in which the surface is viewed.

boundary grows as the magnitude of the mean (maximum) turning angle, averaged over the visible portion of the boundary, increases.

5.2. Relative size and protrusion

As we have seen, for parts of a silhouette we need a notion of part base before we can define a part's relative size and protrusion. The same is true for parts of 3D objects. But, as we mentioned before, we do not yet have a theory of part cuts, in 2D or 3D, and therefore no principled way to define the part bases. In the case of silhouettes we simply used a straight line joining two boundary points to define the base of a part (Fig. 16(a)). We need a similar, though necessarily more complex, construction for the base of a part in 3D.

Consider the part shown in Fig. 22(a). We wish to define its base so that we can then define its relative size and protrusion. Again, the precise definitions will depend on the assumed nature of the visual representation (viewer dependent versus viewer independent).

Consider first a viewer-independent representation. We can take an *invariant*

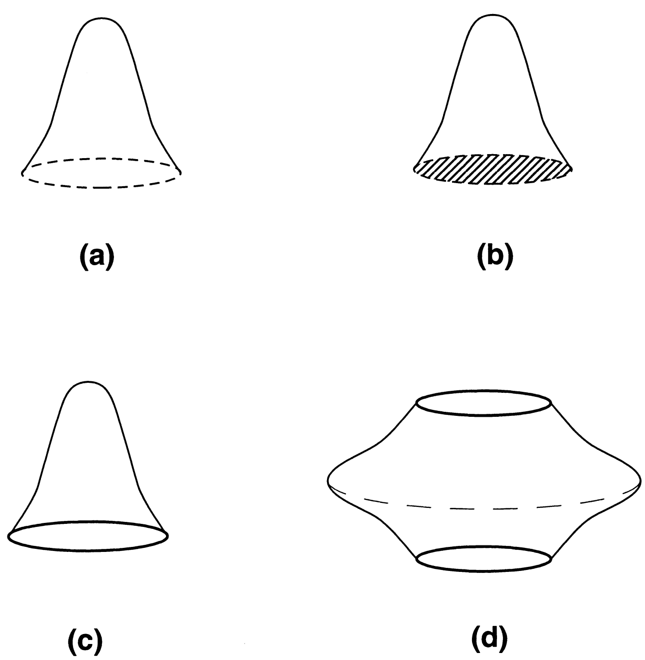


Fig. 22. (a) A part defined by a single closed boundary curve. (b) Invariant base for a part defined by a single closed boundary curve. The invariant base is used to define the relative volume and protrusion of a part. (c,d) Surface areas of parts defined by (c) a single closed boundary curve and (d) several closed boundary curves.

base of the part to be the minimal surface delimited by its boundary curve, as shown in Fig. 22(b).[7] Intuitively, this surface is the soap film obtained by dipping the boundary (by itself) into liquid soap and then lifting it out (see, for example, Do Carmo, 1976). If the boundary curve is planar, then the base is the planar region it encloses.

The invariant volume of the part is the volume enclosed between the surface of the part and its invariant base(s). The *invariant relative volume* of the part is the ratio of the invariant volume of the part to the invariant volume of the whole object. The larger this ratio, the more salient, one might predict, is the part.

The invariant surface area of a part having one boundary curve is the area of the surface within that boundary curve (excluding the invariant base, see Fig. 22(c)). The *invariant protrusion* of the part is the ratio of the area of its invariant surface and its invariant base.

If the part has more than one boundary curve and therefore more than one invariant base, as shown in Fig. 22(d), then we define the invariant protrusion of the part to be the ratio of its invariant surface area (excluding the areas of the invariant bases) to the sum of the areas of its invariant bases.

However, it seems likely that the perception of protrusion and relative size of a part vary with viewpoint. One reason is that the back of each (opaque) part is not visible in a single view, so that the invariant measures described above cannot, in practice, be obtained from a single view. So human vision can adopt either of two strategies: it can use just the visible surfaces of objects and their parts, or it can try to complete the shape of the (not visible) back sides of objects and their parts. We consider these possibilities in turn.

First, human vision might take a minimal approach and assume that there is no back to the part. What you see of the part is all there is. This approach is illustrated in Fig. 23. Fig. 23(a) shows the case of a part defined by one boundary, and Fig. 23(b) shows the case of a part defined by two boundaries. Fig. 23(c,d) illustrates the resulting *cropped bases* for the respective cases. The *visible surface areas* for the two cases are just the areas of their respective visible surfaces. (One can similarly define the visible surface area of the entire object.) The resulting *visible*

[7] This seems a natural choice, given that we don't yet have a theory of part cuts. Minimal surfaces for 3D parts, and straight lines for silhouette parts, are the simplest ways to cut the parts. However a more sophisticated theory of part cuts for 3D objects might use the geometry of the object in the neighborhood of its part boundaries. For instance, instead of interpolating a minimal surface between a part boundary, it might interpolate a surface which is smoothly continuous with a part's surface at the part boundary and which, say, minimizes quadratic variation (Grimson, 1981) or other functionals (e.g., Blake, 1984; Blake and Zisserman, 1986, 1987; Marroquin, 1985; Terzopoulos, 1984, 1986; Weiss, 1990). Similar remarks hold, mutatis mutandis, for part cuts on silhouettes (Siddiqi and Kimia, 1995). Whatever the ultimate theory of part cuts is, we expect that it will not significantly affect the proposals regarding part salience, for these proposals are about ratios, for example, the ratio of the perimeter of a part to its base, and about how salience varies monotonically with these ratios. Small changes in the definition of base will have little effect on all this.

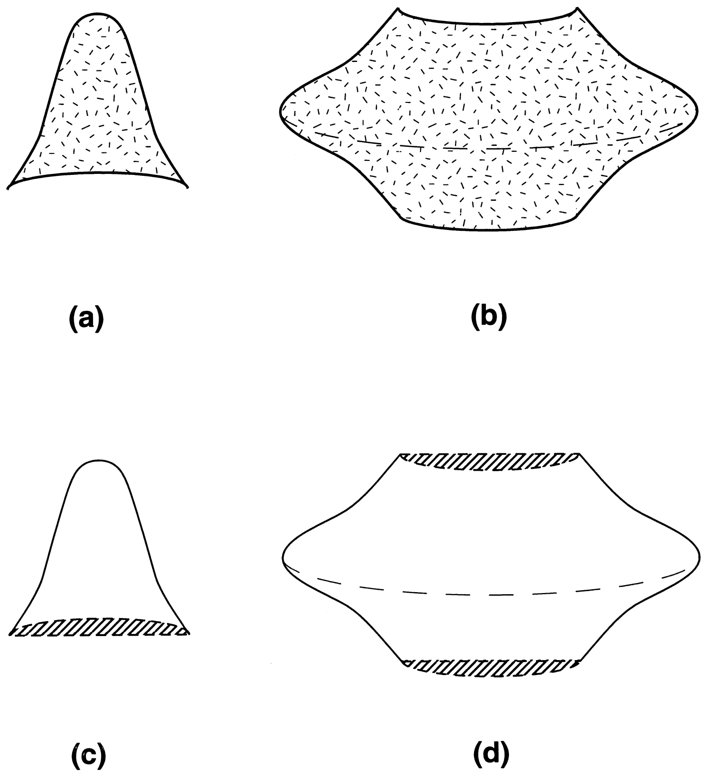


Fig. 23. (a,b) Parts defined by (a) a single closed boundary curve and (b) two closed boundary curves. If human vision makes no inference about the unseen back of a part, then the measures of its salience must depend only on its visible aspects. (c,d) Cropped bases for the parts shown in (a) and (b). These are the bases one obtains by making a minimal inference about the unseen backside of a part.

volume of the part is the volume enclosed[8] by the visible surface and cropped base(s). The *visible protrusion* of the part, then, is defined to be the ratio of its visible surface area (excluding the area(s) of its visible base(s)) to the (summed) area of its cropped base(s). The *visible relative volume* of a part is defined to be the ratio of the visible volume of the part to the visible volume of the entire object. We are led then to consider the following two hypotheses:

> *Hypothesis of visible relative volume:* The salience of a 3D part increases as the magnitude of its visible relative volume increases.

[8] By the volume ''enclosed'' by a visible surface and cropped base we mean the volume that is bounded on the back by a minimal surface. Obtaining the visible volume for an entire object is more difficult due to the possibility of internal occluding contours. One must add the visible volumes for all of the individual parts of the object.

Hypothesis of visible protrusion: The salience of a 3D part increases as the magnitude of its visible protrusion increases.

Second, rather than restrict itself to the visible surfaces and volumes of parts, human vision might try to complete the parts. For instance, it might use characteristics of the visible surface of a part and assumptions about symmetry to smoothly fill in its back (see, for example, Terzopoulos et al., 1987, for a theoretical account of how this might be done). How precisely human vision completes 3D shapes (if it does) is not known, and speculation on this is beyond the scope of this paper. Once the completion is done, we can compute the completed protrusion and completed relative volume of a part, precisely as we did, respectively, the invariant protrusion and invariant relative volume. Hence, we consider the following hypotheses.

Hypothesis of completed relative volume: The salience of a 3D part increases as the magnitude of its completed relative volume increases.
Hypothesis of completed protrusion: The salience of a 3D part increases as the magnitude of its completed protrusion increases.

Finally we should note that our discussion in this section has assumed that at least some of the relevant part boundaries are visible. If none of these boundaries are visible, human vision could either estimate their locations or use just the visible surface of the part to compute its salience. These two cases lead to different empirical predictions. If human vision estimates the hidden part boundaries, then one would expect that the salience of parts will be less affected by foreshortening (an effect due to projection). If human vision uses only the visible surfaces of parts then its measures of part salience should be more affected by foreshortening. Thus this issue is amenable to psychophysical tests in which objects are rotated in depth and judgments of part salience are measured as a function of foreshortening.

6. Salience in 3D: smooth surfaces

We now consider smooth surfaces in 3D. There is not much new to be said about them over what has already been said in the other cases. The theoretical measures of protrusion and size in this case are identical to those for crease boundaries. And the theoretical measures of boundary salience follow straightforwardly from the corresponding smooth silhouette measures (viz., normalized curvature and locale turning) by extending them to the one-dimensional smooth boundaries of the 3D case. The only added twist here is that each point on the boundary is uniquely associated to a line of curvature, namely, the line of curvature for which that point is a negative minimum of curvature. Thus for each point on the boundary the computation of its normalized curvature, locale, and locale turning must be done with respect to its own line of curvature.

7. A theory of part salience

We have considered some factors affecting part salience, given precise definitions of these factors, and cast them as hypotheses. We now propose a concrete theory of part salience by stating which of these hypotheses we think best (i.e., most plausible as a psychological account). The theory describes how relative size, protrusion, and boundary strength affect perceived salience. This gives a clear target for theorists to refine and experimentalists to test.

The theory has two parts: a theory for 2D silhouettes and a theory for 3D shapes. We state these theories separately below. But first we reiterate one hypothesis that we claim holds in both cases.

> *Hypothesis of salient figures:* Other things being equal, that choice of figure and ground is preferred which leads to the most salient parts for the figure object.

7.1. Theory for silhouettes

1. *Hypothesis of turning normals:* The salience of a cusp boundary (and the salience of the parts for which the cusp is a boundary) increases as the magnitude of the turning angle between its two normal vectors increases.
2. *Hypothesis of relative area:* The salience of a part increases as the ratio of its visible area to the visible area of the whole silhouette increases.
3. *Hypothesis of silhouette protrusion:* The salience of a part increases as its protrusion increases.
4. *Hypothesis of locale turning:* The salience of a smooth boundary increases as the turning of the normals in its locale increases.
5. *Hypothesis of normalized curvature:* The salience of a smooth boundary increases as the normalized curvature at the boundary increases.

7.2. Theory for 3D shapes

1. *Hypothesis of turning normals for uniform creases:* The salience of a crease boundary which is uniform grows as the magnitude of its turning angle increases.
2. *Hypothesis of visible mean turning:* The salience of a crease boundary grows as the magnitude of the mean turning angle, averaged over the visible portion of the boundary, increases.
3. *Hypothesis of completed relative volume:* The salience of a 3D part increases as the magnitude of its completed relative volume increases.
4. *Hypothesis of completed protrusion:* The salience of a 3D part increases as the magnitude of its completed protrusion increases.
5. *Hypothesis of normalized curvature for uniform boundaries:* The salience of a smooth boundary which is uniform increases as the magnitude of its normalized curvature increases.

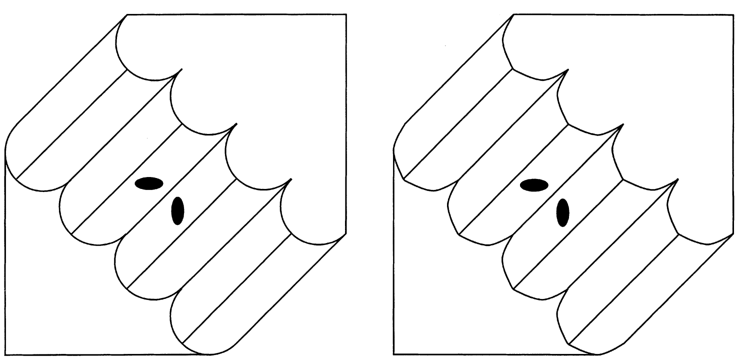
6. *Hypothesis of locale turning for uniform boundaries:* The salience of a smooth boundary which is uniform increases as the magnitude of its locale turning increases.
7. *Hypothesis of visible mean normalized curvature:* The salience of a smooth boundary grows as the magnitude of the mean normalized curvature, averaged over the visible boundary, increases.
8. *Hypothesis of visible mean locale turning:* The salience of a smooth boundary grows as the magnitude of the mean locale turning, averaged over the visible boundary, increases.

8. Visual demonstrations and psychophysical experiments

We have proposed a theory of the geometric factors that influence judgments of part salience. Its plausibility can be suggested by demonstrations but determined, of course, only by experiments. To such experiments and demonstrations we now turn.

EXPERIMENT 1

Consider the demonstration shown in Fig. 24(a). It consists of two variations of the Schroeder staircase shown earlier in Fig. 3. Recall that the Schroeder staircase can be seen in two different interpretations, either as a normal ascending staircase (the ''below'' interpretation), or as a strange inverted staircase that one would dare not climb (the ''above'' interpretation). One usually prefers to see the staircase below. But in Fig. 24(a) we have altered the staircase so that the below



(a) (b)

Fig. 24. Two modifications of the Schroeder staircase. By altering the relative salience of the steps in the two figure–ground interpretations of each staircase, we make one interpretation more easily seen. Here the inverted interpretation is more easily seen in both figures. These figures were used as stimuli in Experiment 1.

interpretation has the least salient part boundaries, namely, arcs of circles, while the above interpretation has highly salient part boundaries, namely, concave creases with large turning angles. According to the theory of part salience proposed here, the above interpretation should now be preferred, since it has the most salient parts. You can check whether your perception accords with this theory: it does if you prefer to see the dots as lying on two distinct round steps; it doesn't if you prefer to see the dots as lying on a single step. If you continue to see the dots as lying on two distinct round steps even while slowly rotating the staircase through 360 degrees, then your perception strongly accords with this theory.

A similar story holds for Fig. 24(b). Here the staircase has been altered so that the below interpretation has part boundaries with a smaller turning angle (and therefore weaker salience), whereas the above interpretation has part boundaries with a larger turning angle (and therefore stronger salience). Again you can check to see if your perception accords with the theory: it does if you prefer to see the dots as lying on two distinct steps, even while rotating the staircase; it doesn't if you prefer to see the dots as lying on a single step. These predictions are tested by this experiment.

9. Method

9.1. Subjects

The subjects were 10 students from the University of California, Irvine. The subjects were volunteers naive to the purposes of the experiment. All had normal or corrected to normal acuity.

9.2. Stimuli

The stimuli were the standard Schroeder staircase (Fig. 3) and its two modifications (Fig. 24). The staircases were viewed at a distance of about 1 m and subtended about 7 degrees of visual angle.

9.3. Design

Each staircase was shown an equal number of times upright and inverted, and with dots as depicted in Figs. 3 and 24 or with the dots shifted by one face. The dots were shifted so that (especially for the staircases of Fig. 24) the dots would straddle the stronger part boundary half the time and the weaker part boundary half the time. There were a total of 3 (staircases) $\times$ 2 (inversions) $\times$ 2 (dot placements) $\times$ 6 (repetitions) = 72 trials. These were preceded by 24 practice trials.

9.4. Apparatus

The figures were displayed on a high-resolution (1024×768) monitor by a Macintosh Quadra computer using the SuperLab program. Subjects used a keyboard to respond.

9.5. Procedure

Subjects were instructed as follows: ''On each trial you will see a staircase having two dots. Sometimes the dots will appear on the same step, sometimes on different steps. Press the 'S' button if they appear on the same step, or the 'K' button if they appear on different steps. The first 24 trials are practice. Please respond as quickly and accurately as possible. In making your judgments please attend not just to the dots but also to the steps.''

The staircases were presented in random order. Each trial consisted of a fixation dot for 500 ms, then a staircase for 50 ms, and finally a mask composed of black dots randomly placed on a white background. The mask remained until the subject responded. The experiment was sufficiently brief that no rest periods were required.

10. Results and discussion

Consider again Fig. 24(b). If a subject responds ''different'' to this figure then we know that the subject sees the most salient negative minima as step boundaries. Thus we record the response as ''most salient''. If a subject responds ''same'' to this figure then we record the response as ''least salient''. For the version of Fig. 24(b) in which the dots are shifted by one step face, we record just the opposite: a response of ''different'' is recorded as '' least salient'' and a response of ''same'' as ''most salient''. This holds true whether the figure is shown upright or inverted. We record responses to Fig. 24(a) and 3 in like manner, although for Fig. 3 this is pure convention since all boundaries are of equal salience.

For Fig. 24(a), 65.8% of all subject responses were ''most salient''. This value is the mean over all four versions of Fig. 24(a) that were shown: inverted and upright, dots straddling a strong boundary and dots straddling a weak boundary. The values for Fig. 24(b) and 3 were, respectively, 67.9% and 48.3%. A one-factor analysis of variance (ANOVA) on these means shows a significant effect of boundary strength in the predicted direction: $F(2, 18)=6.55$, $p<.008$. Post hoc contrasts showed a significant difference between Figs. 3 and 24(b), $p<.004$, but not between Fig. 24(b) and 24(a).

This supports the hypothesis that subjects choose figure and ground so that figure has the more salient part boundaries. These results obtained with presentations of 50 ms. This suggests that salience of negative minima of curvature affects the perception of figure and ground very early in visual processing. This also suggests that the interpretation of line drawings is a rapid early visual process (Enns and Rensink, 1991, 1992, 1995).

EXPERIMENT 2

Consider the face–vase illusion shown in Fig. 25(a). In Fig. 25(b) it is altered so that part boundaries for the faces are cusps and are therefore more salient than part boundaries for the vase. By the theory of salience proposed here one should prefer to see faces. In Fig. 25(c) it is altered so that part boundaries for the vase are cusps and are therefore more salient than part boundaries for the faces. Now one should prefer to see the vase. (One might argue that we prefer the vase in Fig. 25(c) because the faces look strange. But so do the faces in Fig. 25(b), and yet we prefer to see them.) These predictions are tested by this experiment.

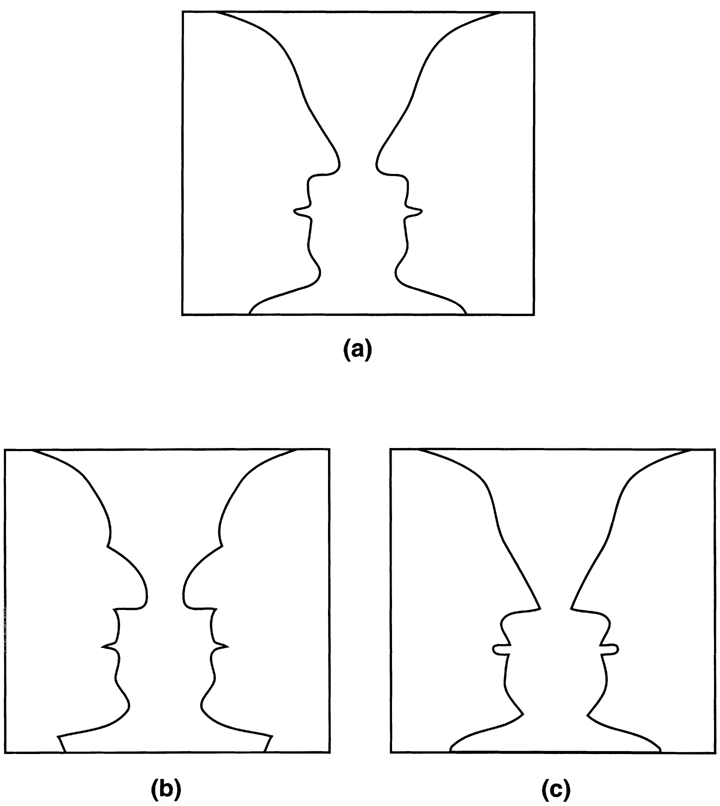


Fig. 25. Modifications of the face–vase illusion. In (a) is shown the outline of a typical face–vase figure. In (b) the part boundaries for the faces are more salient, so the faces should be more easily seen. In (c) the part boundaries for the vase are more salient, so the vase should be more easily seen. A similar effect can be obtained using T-junctions at part boundaries (Hoffman and Richards, 1982). These figures were used as stimuli in Experiment 2.

11. Method

11.1. Subjects

The subjects were 10 students from the University of California, Irvine. The subjects were volunteers naive to the purposes of the experiment. All had normal or corrected to normal acuity.

11.2. Stimuli

The stimuli were the face–vase illustrations of Fig. 25. On each trial one such illustration was shown. The face–vase illustrations were viewed at a distance of about 1 m and each subtended about 9 degrees of visual angle.

11.3. Design

There were three face–vase illustrations. Each was shown 16 times for a total of 48 trials. These were preceded by 12 practice trials.

11.4. Apparatus

The figures were displayed on a high-resolution (1024×768) monitor by a Macintosh Quadra computer using the SuperLab program. Subjects used a keyboard to respond.

11.5. Procedure

Subjects were instructed as follows: ''On each trial you will be shown a figure which can be seen either as a vase or as two faces. If you see it as a vase, press the 'V' button. If you see it as two faces, press the 'F' button. The first 12 trials are practice. Please respond as quickly and accurately as possible.''

The face–vase illustrations were presented in random order. Each trial consisted of a fixation dot for 500 ms, then a face–vase figure for 250 ms, and finally a screen with the words ''Please respond with 'V' for vase or 'F' for faces.'' This screen remained until the subject responded. The experiment was sufficiently brief that no rest periods were required.

12. Results and discussion

The percentage, averaged over 10 subjects, of ''vase'' responses was 44.4% for Fig. 25(a), 25% for Fig. 25(b), and 74.4% for Fig. 25(c). (Data from an eleventh subject, who never reported the vases interpretation for Fig. 25(a), was excluded from the analysis.) A one-factor ANOVA of these means shows a significant effect of boundary strength in the predicted direction: $F(2, 18)=11.57$, $p<.001$. This supports the hypothesis that subjects choose figure and ground so that figure has the more salient part boundaries.

EXPERIMENT 3

Consider the patterns, motivated by experiments of Stevens and Brookes (1988), shown in Fig. 26. To create each one, we start with a plane curve given by a periodic function, so that all turnings at part boundaries for one choice of figure and ground have one common value, and all turnings for the other choice of figure and ground have another (possibly different) common value. The ratio of these two values (larger to smaller) we call the *turning ratio*. We take the plane curve and produce multiple parallel copies of it. This results in a uniformly repeated pattern which is perceived as a set of overlapping figures, much like the scales on a fish.

Each pattern in Fig. 26 is ambiguous. One could see overlapping ''fish scales'' pointing either to the right or to the left. Which interpretation will predominate? According to the hypothesis of salient figures, the one with the more salient parts. In Fig. 26(a) we have made the salience of the two interpretations equal, by making their boundary strengths, areas, and protrusions all equal. As a result, both interpretations of this figure should be equally preferred. In Fig. 26(b), we have made the salience of the right-pointing parts greater, by increasing their areas and turning angles. We predict that the right-pointing interpretation is now more easily seen. You can check this prediction by looking at the dot in Fig. 26(b). Notice that it is near a curve, and that the curve is the boundary of a fish scale. Now check: Is the dot on that scale or just off it? If most of the time it looks to be on the scale, then your perception agrees with our prediction. These casual observations are tested in this experiment, using a technique introduced by Stevens and Brookes (1988).

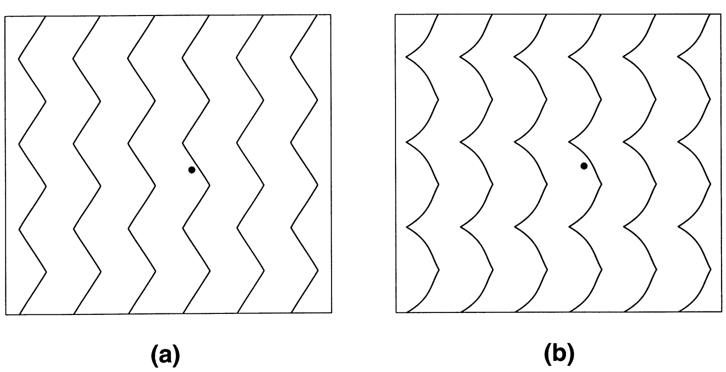


Fig. 26. Two demonstrations of part salience (adapted from Stevens and Brookes, 1988). Each figure can be seen in two different interpretations. The degree to which one interpretation dominates the other depends on the relative salience of their respective parts. These figures were used as stimuli in Experiment 3.

13. Method

13.1. Subjects

The subjects were 10 students from the University of California, Irvine. The subjects were volunteers naive to the purposes of the experiment. All had normal or corrected to normal acuity.

13.2. Stimuli

The stimuli were three fish-scale patterns of the type shown in Fig. 26, each with its own turning ratio and area ratio. The three turning ratios were (in degrees) $120:120$ (as in Fig. 26(a)), $120:90$, and $120:60$ (as in Fig. 26(b)) with corresponding area ratios of $1:1$, $1.25:1$, and $1.67:1$. On each trial one such pattern was shown. The illustrations were viewed at a distance of about 1 m and each was about 7 degrees tall and 9 degrees wide.

13.3. Design

Each fish-scale pattern was shown on half the trials in an original orientation, and on half the trials with a 180 degree rotation. On half the trials the dot was to the left of its nearest curve, and on half the trials to the right. On half the trials the dot was placed at a higher vertical location and on half the trials at a lower vertical location. In each case the dot appeared half way between two consecutive cusps (like the dot in Fig. 26(b)). There were a total of 3 (turning ratios)$\times$2 (orientations)$\times$2 (left–right placements)$\times$2 (vertical locations)$\times$3 (repetitions)= 72 trials. These were preceded by 24 practice trials.

13.4. Apparatus

The figures were displayed on a high-resolution (1024×768) monitor by a Macintosh Quadra computer using the SuperLab program. Subjects used a keyboard to respond.

13.5. Procedure

Subjects were instructed as follows: ''On each trial you will be shown a pattern of fish scales. On the pattern will be a single dot, near a curve of the pattern. Decide whether the dot is on or just off the scale bounded by that curve. Press the 'F' button if the dot is just off the scale, or the 'N' button if the dot is on the scale. The first 24 trials are practice. Please respond as quickly and accurately as possible.'' The fish-scale patterns were presented in random order. Each trial consisted of a fixation dot for 500 ms, then a fish-scale pattern for 250 ms, and finally a mask composed of black dots randomly placed on a white background.

The mask remained until the subject responded. The experiment was sufficiently brief that no rest periods were required.

14. Results and discussion

Consider again Fig. 26(b). If a subject responds that the dot is ''on'' the scale then we know that the subject sees the most salient negative minima as scale boundaries. Thus we record the response as ''most salient''. If a subject responds ''off'' to this figure then we record the response as ''least salient''. Similar interpretations hold, mutatis mutandis, for the versions of Fig. 26(b) with the other dot placements. We did the same for responses to the 120 : 90 figure, and for Fig. 26(a), although for Fig. 26(a) this is pure convention since all boundaries are of equal salience.

The percentage, averaged over 10 subjects, of ''most salient'' responses was 98.3% for Fig. 26(b) (with ratio 120 : 60), 92.5% for the stimulus with ratio 120 : 90, and 54.6% for Fig. 26(a) (with ratio 120 : 120). A one-factor ANOVA on these means shows a significant effect of boundary strength in the predicted direction: $F(2, 18) = 142$, $p < .0001$. This supports the hypothesis that subjects choose figure and ground so that figure has the more salient part boundaries.

DEMONSTRATION 1

A simple modification of the Stevens and Brookes stimuli leads to other demonstrations (such as Fig. 27) of the effects of part salience. Simply use the same generating curves to create *ruled surfaces*, by placing several copies of the generating curve side by side, but much more closely than in the Stevens and Brookes stimuli, so that one gets an impression of a surface in three dimensions. The surface depicted in Fig. 27 has two choices for figure and ground. In the first choice all parts have greater volume and larger turning angles at their boundaries than in the second choice. Thus we predict that this first choice is more easily seen. You can judge the prediction for yourself. If you see the dot in Fig. 27 as lying on a hill most of the time, then your perception is in line with our prediction. Otherwise it is not.

DEMONSTRATION 2

A modification of the cosine surface (Fig. 4) leads to another demonstration (Figs. 28 and 29). We use a plane curve to generate a surface of revolution. Each surface of this type has two choices for figure and ground. Which choice is most easily seen depends on the salience of the parts for the two choices. In Fig. 28(a,b) the parts have identical salience in both choices. In fact, Fig. 28(b) is just Fig. 28(a) turned upside down. Nevertheless, the two figures look quite different, because we usually see Fig. 28(b) in the opposite choice of figure and ground than we see Fig. 28(a) (due to a preference, ceteris paribus, to see figure below). Fig.

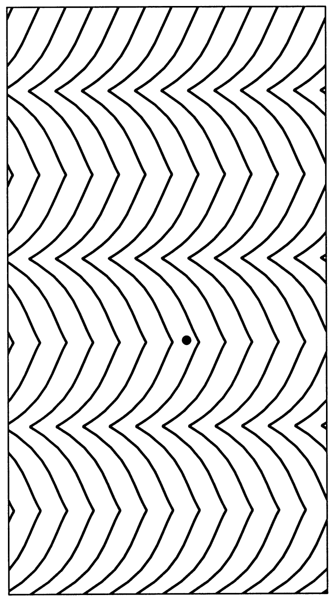


Fig. 27. A ruled surface. This surface can be seen in two different interpretations. The degree to which one interpretation dominates depends on the relative salience of their respective parts.

29(a,b) is just like Fig. 28(a,b), except that now the parts for one choice of figure and ground, namely, the choice you first see in Fig. 29(a), have greater volume and larger turning angles at their boundaries than do the parts for the other choice, namely, the choice you first see in Fig. 29(b). To demonstrate to yourself that the interpretation you see in Fig. 29(a) is, as predicted, more dominant than the one you see in Fig. 29(b), try the following. Slowly turn the page of the figure through a full circle. You will see that there are some orientations of the page in which both of the two figures are seen in the ''dominant'' interpretation (viz., the interpretation you see of Fig. 29(a) when the page is upright). You will also see that there are *no* orientations of the page in which both of the two figures are seen in the

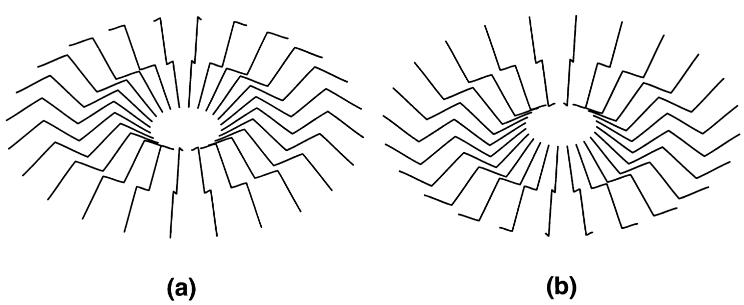


Fig. 28. Two surfaces of revolution. Each surface can be seen in two different interpretations. Here the parts in each interpretation have identical salience, so that one interpretation is not more dominant than the other.

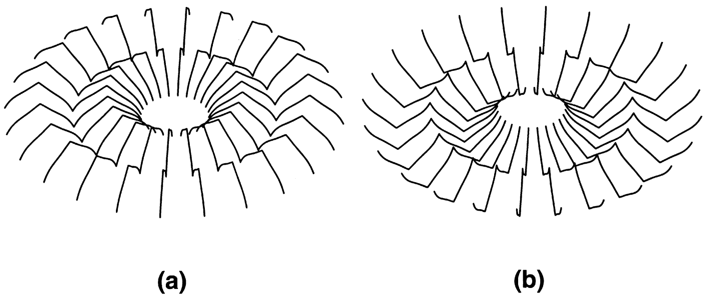


Fig. 29. Two more surfaces of revolution. Here the parts in one interpretation (that seen in (a) when the page is upright) are more salient than the parts in the other interpretation (that seen in (b) when the page is upright).

''nondominant'' interpretation (viz., the interpretation you see of Fig. 29(b) when the page is upright). This demonstrates that the Fig. 29(a) interpretation is indeed more dominant, as is predicted by the greater salience of its parts. If you try the same rotation test on Fig. 28, you will find that at all orientations of the page the two surfaces appear to have figure–ground interpretations that are the *opposite* of each other. This is due to the equal salience of the parts in the two interpretations.

15. Concluding remarks

We have proposed that the salience of a part depends on (at least) three factors: its size relative to the whole object, the degree to which it protrudes, and the strength of its boundaries. We have considered precise definitions for these factors, and presented visual demonstrations and psychophysical tests of their effects. And in Experiment 1 we have given evidence that part salience affects the early visual processing of figure and ground.

Assessing the salience of visual parts is one small aspect of the general problem of object recognition. However, it is one that cannot be avoided, since psychophysical evidence now suggests that human vision breaks shapes into parts, and does so preattentively (Baylis and Driver, 1995a,b; Driver and Baylis, 1995). That parts are computed preattentively is surprising and significant: parts and their salience may affect not just high-level visual processes such as object recognition, but also lower-level processes, such as figure/ground organization. (Peterson and Gibson, 1993, discuss how recognition can affect the choice of figure/ground.)

As mentioned earlier, there is much psychophysical work to be done on the perception of part salience. There are also several directions for further theoretical work. First, there are factors affecting part salience that we have not discussed here, such as symmetry and ecological significance (see, for example, Wenderoth, 1994), part orientation, and whether a part appears alone or in a group of similar parts. Second, as we have noted before, we need a theory of part cuts both for silhouettes and for 3D shapes. The minima rule for defining part boundaries is but a first step to a complete theory of part cuts. Some progress has been made in the case of silhouettes (Beusmans et al., 1987; Siddiqi and Kimia, 1995; Singh et al., 1996), but more is needed. Progress thus far reveals that although almost all negative minima are part boundaries, not all part boundaries are negative minima. Therefore measures for the salience of these other boundaries will need to be developed. Third, we need a theoretical understanding of parts for which relative size and protrusion (but not boundary strength) appear to be the primary or sole determinants of salience. Two examples are shown in Fig. 30(a,b). To understand these examples it may be helpful to distinguish three kinds of parts: main, core, and peripheral. A peripheral part has only a single part cut, whereas a core part has more than one part cut (Fig. 30(c,d)). A main part is, roughly speaking, much larger than any other part (Fig. 30(e,f)). (We say ‘‘roughly speaking’’ because a

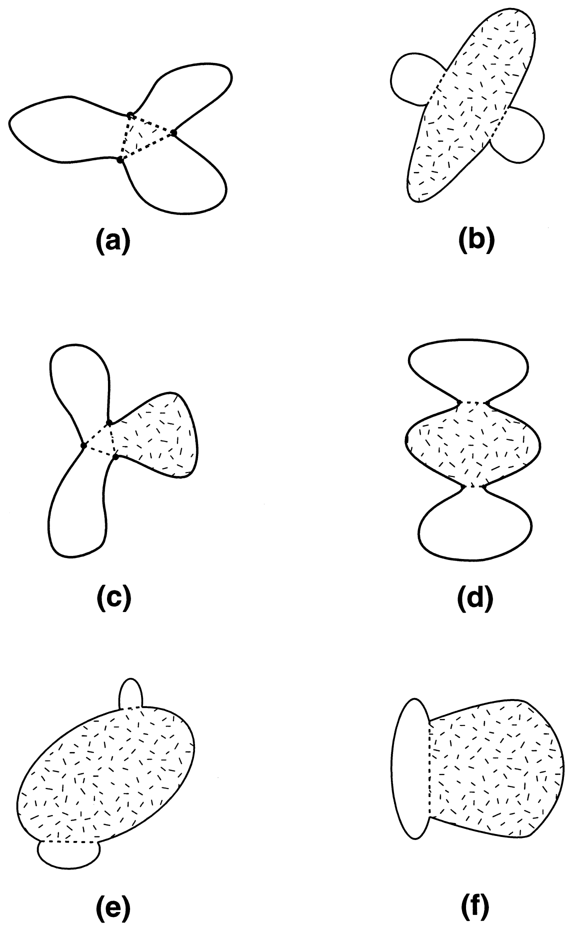


Fig. 30. (a) A part (shown stippled) whose salience is determined primarily by its relative size. (b) A part whose salience is determined primarily by its relative size and protrusion. (c) A peripheral part. (d) A core part. (e,f) Main parts.

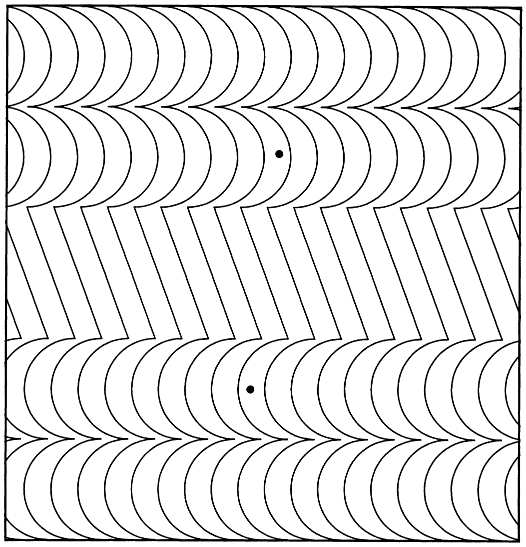


Fig. 31. A puzzle. Why does one prefer to see this as an impossible object in which figure and ground are not consistently maintained throughout (refer to footnote 9)?

precise definition of main part probably requires a hierarchical decomposition of a shape into parts, and then comparison of relative sizes within each specific level.) The salience of main parts is determined, it appears, primarily by their relative sizes and protrusions (Fig. 30(a)). The salience of certain core parts is determined, it appears, primarily by their relative sizes (Fig. 30b). But for peripheral parts all three factors appear to affect salience. Clearly, more theoretical work is needed on these issues.

Some of the fun of studying part salience is in creating new figures which play with salience. Sometimes, as in Fig. 31, this leads to visual puzzles.[9]

Acknowledgments

For discussions and suggestions, we thank Marc Albert, Bruce Bennett, Myron Braunstein, Mike D'Zmura, Ki-Ho Jeon, Jin Kim, Jeff Liter, Scott Richman, Asad Saidpour, and Jessica Turner. We especially thank Marc Albert for helping us refine our definition of part protrusion. And we thank three anonymous reviewers for helpful comments. This work was supported by National Science Foundation grant DIR-9014278.

[9] At the top, one choice of figure and ground gives more salient parts, but at the bottom the opposite choice does. As a result you probably see both of the dots as lying on bumps (rather than one on a bump and one on a dent). This figure pits the principle of most salient figures against the principle of physical possibility, and physical possibility often loses.

References

Attneave, F. (1971). Multistability in perception. *Scientific American*, *225*(6), 63–71.

Bahnsen, P. (1928). Eine untersuchung ueber symmetrie und asymmetrie bei visuellen wahrnehmungen. *Zeitschrift für Psychologie*, *108*, 355–361.

Baylis, G.C., & Driver, J. (1995a). One-sided edge assignment in vision. 1. Figure–ground segmentation and attention to objects. *Current Directions in Psychological Science*, *4*, 140–146.

Baylis, G.C., & Driver, J. (1995b). Obligatory edge assignment in vision: The role of figure and part segmentation in symmetry detection. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1323–1342.

Bennett, B.M., & Hoffman, D.D. (1987). Shape decompositions for visual shape recognition: The role of transversality. In W.A. Richards & S. Ullman (Eds.), *Image understanding* (pp. 215–256). Norwood, NJ: Ablex.

Beusmans, J., Hoffman, D.D., & Bennett, B.M. (1987). Description of solid shape and its inference from occluding contours. *Journal of the Optical Society of America*, *A*, *4*, 1155–1167.

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

Biederman, I., & Cooper, E.E. (1991). Priming contour-deleted images: Evidence for intermediate representations in visual object recognition. *Cognitive Psychology*, *23*, 393–419.

Binford, T.O. (1971, December). Visual perception by computer. *IEEE Systems Science and Cybernetics Conference*, Miami, FL.

Blake, A. (1984). Reconstructing a visible surface. In *Proceedings of AAAI Conference 1984* (pp. 362–365). Los Altos, CA: AAAI Press.

Blake, A., & Zisserman, A. (1986). Invariant surface reconstruction using weak continuity constraints. In *Proceedings of Computer Vision and Pattern Recognition* (pp. 62–67). Washington, DC: IEEE Computer Society Press.

Blake, A., & Zisserman, A. (1987). *Visual reconstruction*. Cambridge, MA: MIT Press.

Bower, G.H., & Glass, A.L. (1976). Structural units and the redintegrative power of picture fragments. *Journal of Experimental Psychology*, *2*, 456–466.

Brady, M., & Asada, H. (1984). Smoothed local symmetries and their implementation. *International Journal of Robotics Research*, *3*, 36–61.

Braunstein, M.L., Hoffman, D.D., & Saidpour, A. (1989). Parts of visual objects: An experimental test of the minima rule. *Perception*, *18*, 817–826.

Brooks, R.A. (1981). Symbolic reasoning among 3-D models and 2-D images. *Artificial Intelligence*, *17*, 205–244.

Cave, C.B., & Kosslyn, S.M. (1993). The role of parts and spatial relations in object identification. *Perception*, *22*, 229–248.

Cornilleau-Peres, V., & Droulez, J. (1989). Visual perception of surface curvature: Psychophysics of curvature detection induced by motion parallax. *Perception and Psychophysics*, *46*, 351–364.

Dickinson, S.J., Pentland, A.P., & Rosenfeld, A. (1992). From volumes to views: An approach to 3D object recognition. *Computer Vision, Graphics, and Image Processing: Image Understanding*, *55*(2), 130–154.

Do Carmo, M. (1976). *Differential geometry of curves and surfaces*. Englewood Cliffs, NJ: Prentice-Hall.

Driver, J., & Baylis, G.C. (1995). One-sided edge assignment in vision. 2. Part decomposition, shape description, and attention to objects. *Current Directions in Psychological Science*, *4*, 201–206.

Enns, J.T., & Rensink, R.A. (1991). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review*, *98*, 335–351.

Enns, J.T., & Rensink, R.A. (1992). A model for the rapid interpretation of line drawings in early vision. In D. Brogan (Ed.), *Visual search II* (pp. 73–89). London: Taylor and Francis.

Enns, J.T., & Rensink, R.A. (1995). Preemption effects in visual search. *Psychological Review*, *102*, 101–130.

Grimson, W.E.L. (1981). *From images to surfaces: A computational study of the human early visual system*. Cambridge, MA: MIT Press.

Guillemin, V., & Pollack, A. (1974). *Differential topology*. Englewood Cliffs, NJ: Prentice-Hall.
Guzman, A. (1971). Analysis of curved line drawings using context and global information. In *Machine intelligence* (Vol. 6). Edinburgh: Edinburgh University Press.
Halmos, P.R. (1950). *Measure theory*. New York: Springer-Verlag.
Hochberg, J. (1964). *Perception*. Englewood Cliffs, NJ: Prentice-Hall.
Hoffman, D.D. (1983a). *Representing shapes for visual recognition*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.
Hoffman, D.D. (1983b). The interpretation of visual illusions. *Scientific American*, *249*(6), 154–162.
Hoffman, D.D., & Richards, W.A. (1982). Representing smooth plane curves for recognition: Implications for figure–ground reversal. *Proceedings of the National Conference of the American Association for Artificial Intelligence* (pp. 5–8). Los Altos, CA: Kaufmann.
Hoffman, D.D., & Richards, W.A. (1984). Parts of recognition. *Cognition*, *18*, 65–96.
Kanisza, G., & Gerbino, W. (1976). Convexity and symmetry in figure–ground organization. In M. Henle (Ed.), *Art and artefacts* (pp. 25–32). New York: Springer.
Kimia, B.B., Tannenbaum, A.R., & Zucker, S.W. (1990). Toward a computational theory of shape: An overview. *Lecture Notes in Computer Science*, *427*, 402–407.
Kimia, B.B., Tannenbaum, A.R., & Zucker, S.W. (1991). *Entropy scale-space* (pp. 333–344). New York: Plenum Press.
Kimia, B.B., Tannenbaum, A.R., & Zucker, S.W. (1992). *The shape triangle: Parts, protrusions, and bends* (Tech. Rep. No. TR-92-15). Montreal: McGill University Research Center for Intelligent Machines.
Koenderink, J.J. (1984). The structure of images. *Biological Cybernetics*, *50*, 363–370.
Koenderink, J.J. (1986). Optic flow. *Vision Research*, *26*, 161–179.
Koffka, K. (1935). *Principles of gestalt psychology*. New York: Harcourt, Brace and World.
Kurbat, M.A. (1994). Structural description theories: Is RBC/JIM a general-purpose theory of human entry-level object recognition? *Perception*, *23*, 1339–1368.
Leyton, M. (1987). Symmetry–curvature duality. *Computer Vision, Graphics, and Image Processing*, *38*, 327–341.
Leyton, M. (1988). A process grammar for shape. *Artificial Intelligence*, *34*, 213–247.
Leyton, M. (1989). Inferring causal history from shape. *Cognitive Science*, *13*, 357–387.
Leyton, M. (1992). *Symmetry, causality, mind*. Cambridge, MA: MIT Press.
Lowe, D. (1985). *Perceptual organization and visual recognition*. Amsterdam: Kluwer.
Marr, D. (1977). Analysis of occluding contour. *Proceedings of the Royal Society of London, Series B*, *197*, 441–475.
Marr, D. (1982). *Vision*. San Francisco: Freeman Press.
Marr, D., & Nishihara, H.K. (1978). Representation and recognition of three-dimensional shapes. *Proceedings of the Royal Society of London, Series B*, *200*, 269–294.
Marroquin, J. (1985). *Probabilistic solution of inverse problems*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.
Norman, J.F., & Lappin, J.S. (1992). The detection of surface curvatures defined by optical motion. *Perception and Psychophysics*, *51*, 386–396.
Norman, J.F., & Todd, J.T. (1993). The perceptual analysis of structure from motion for rotating objects undergoing affine stretching transformations. *Perception and Psychophysics*, *53*, 279–291.
Palmer, S.E. (1975). Visual perception and world knowledge: Notes on a model of sensory cognitive interaction. In D.A. Norman & D.E. Rumelhart (Eds.), *Explorations in cognition* (pp. 279–307). San Francisco: Freeman.
Palmer, S.E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, *9*, 441–474.
Parent, P., & Zucker, S.W. (1989). Trace inference, curvature consistency, and curve detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*, 823–839.
Pentland, A.P. (1986). Perceptual organization and the representation of natural form. *Artificial Intelligence*, *28*, 293–331.
Peterson, M.A., & Gibson, B.S. (1993). Shape recognition inputs to figure ground organization in three-dimensional displays. *Cognitive Psychology*, *25*, 383–429.

Reed, S.K. (1974). Structural descriptions and the limitations of visual images. *Memory and Cognition*, *2*, 329–336.

Richards, W.A., & Hoffman, D.D. (1985). Codon constraints on closed 2D shapes. *Computer Vision, Graphics, and Image Processing*, *31*, 156–177.

Richards, W.A., Dawson, B., & Whittington, D. (1986). Encoding contour shape by curvature extrema. *Journal of the Optical Society of America*, *3*, 1483–1489.

Richards, W.A., Koenderink, J., & Hoffman, D.D. (1987). Inferring three-dimensional shapes from two-dimensional silhouettes. *Journal of the Optical Society of America*, *4*, 1168–1175.

Roberts, L.G. (1965). Machine perception of three-dimensional solids. In J.T. Tippett et al. (Eds.), *Optical and electrooptical information processing* (pp. 211–277). Cambridge, MA: MIT Press.

Rogers, B.J. (1986). The perception of surface curvature from disparity and motion parallax cues. *Investigative Ophthalmology and Visual Sciences*, *27*, 181.

Rogers, B.J., & Cagenello, R. (1989). Disparity curvature and the perception of three-dimensional surfaces. *Nature*, *339*, 135–137.

Rogers, B.J., & Collett, T.S. (1989). The appearance of surfaces specified by motion parallax and binocular disparity. *Quarterly Journal of Experimental Psychology*, *41*, 697–717.

Rogers, B.J., & Graham, M.E. (1983). Anisotropies in the perception of three-dimensional surfaces. *Science*, *221*, 1409–1411.

Rubin, E. (1958). Figure and ground. In D.C. Beardslee (Ed.), *Readings in perception* (pp. 194–203). Princeton, NJ: Van Nostrand. (Reprinted from *Visuell wahrgenommene figuren*, 1915, Copenhagen: Gyldenalske Boghandel).

Saidpour, A., & Braunstein, M.L. (1994). Curvature and depth judgments of the same simulated shape from motion parallax and structure from motion. *Investigative Ophthalmology and Visual Sciences*, *35*, 1316.

Siddiqi, K., & Kimia, B.B. (1995). Parts of visual form: Computational aspects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *17*, 239–251.

Siddiqi, K., Tresness, K.J., & Kimia, B.B. (1996). Parts of visual form: Psychophysical aspects. *Perception, 25,* 399–424.

Singh, M., Seyranian, G., & Hoffman, D.D. (1996). *Cuts for parsing visual shapes*. University of California, Irvine, Institute for Mathematical Behavioral Sciences Memo, 96–33.

Stevens, K.A., & Brookes, A. (1988). The concave cusp as a determiner of figure–ground. *Perception*, *17*, 35–42.

Sutherland, N.S. (1968). Outlines of a theory of visual pattern recognition in animals and man. *Proceedings of the Royal Society of London*, *B*, *171*, 297–317.

Terzopoulos, D. (1984). *Multi-level computation of visual surface representation*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Terzopoulos, D. (1986). Integrating visual information from multiple sources. In A.P. Pentland (Ed.), *From pixels to predicates: Recent advances in computational vision* (pp. 111–142). Norwood, NJ: Ablex.

Terzopoulos, D., Witkin, A., & Kass, M. (1987). Symmetry-seeking models and 3D object reconstruction. *International Journal of Computer Vision*, *1*, 211–221.

Triesman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, *95*, 15–48.

Todd, J.T. (1984). The perception of three-dimensional structure from rigid and nonrigid motion. *Perception and Psychophysics*, *36*, 97–103.

Todd, J.T., & Norman, J.F. (1991). The visual perception of smoothly curved surfaces from minimal apparent motion sequences. *Perception and Psychophysics*, *50*, 509–523.

Todd, J.T., Koenderink, J.J., van Doorn, A.J., and Kappers, A.M.L. (1996). Effects of changing viewing conditions on the perceived structure of smoothly curved surfaces. *Journal of Experimental Psychology—Human Perception and Performance, 22,* 695–706.

Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General*, *113*, 169–193.

Waltz, D. (1975). Generating semantic descriptions from drawings of scenes with shadows. In P. Winston (Ed.), *The psychology of computer vision* (pp. 19–91). New York: McGraw-Hill.

Weiss, I. (1990). Shape reconstruction on a varying mesh. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*, 345–362.

Wenderoth, P. (1994). The salience of vertical symmetry. *Perception*, *23*, 221–236.

Wilson, H.R., & Richards, W.A. (1985). Discrimination of contour curvature: Data and theory. *Journal of the Optical Society of America*, *A*, *2*(7), 1191–1198.

Wilson, H.R., & Richards, W.A. (1989). Mechanism of contour curvature discrimination. *Journal of the Optical Society of America*, *A*, *6*(1), 1006–1115.

Winston, P.A. (1975). Learning structural descriptions from examples. In P.H. Winston (Ed.), *The psychology of computer vision* (pp. 157–209). New York: McGraw-Hill.

Witkin, A.P. (1983). Scale-space filtering. *Proceedings of the International Joint Conference on Artificial Intelligence*, Karlsruhe, pp. 1019–1021.

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.

Wolfe, J.M., Yee, A., & Friedman-Hill, S.R. (1992). Curvature is a basic feature for visual search tasks. *Perception*, *21*, 465–480.