Cognition, 18 (1984) 65–96

Parts of recognition*

D.D. HOFFMAN
University of California, Irvine

W.A. RICHARDS
Massachusetts Institute of Technology

Abstract

We propose that, for the task of object recognition, the visual system decomposes shapes into parts, that it does so using a rule defining part boundaries rather than part shapes, that the rule exploits a uniformity of nature—transversality, and that parts with their descriptions and spatial relations provide a first index into a memory of shapes. This rule allows an explanation of several visual illusions. We stress the role inductive inference in our theory and conclude with a précis of unsolved problems.

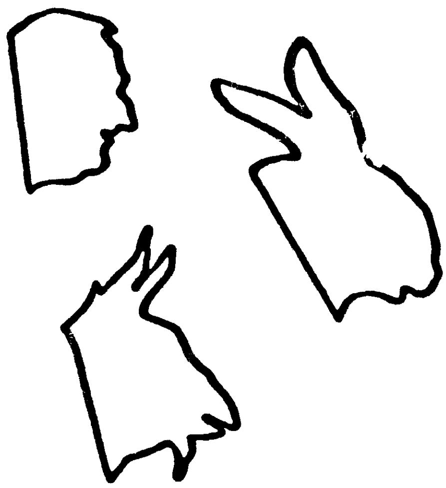
1. Introduction

Any time you view a statue, or a simple line drawing, you effortlessly perform a visual feat far beyond the capability of the most sophisticated computers today, through well within the capacity of a kindergartener. That feat is shape recognition, the visual identification of an object using only its shape. Figure 1 offers an opportunity to exercise this ability and to make several observations. Note first that, indeed, shape alone is sufficient to recognize the objects; visual cues such as shading, motion, color, and texture are not present in the figure. Note also that you could not reasonably predict the contents of the figure before looking at it, yet you recognized the objects.

*We are grateful to Thomas Banchoff, Aaron Bobick, Mike Brady, Carmen Egido, Jerry Fodor, Jim Hodgson, Jan Koenderink, Jay Lebed, Alex Pentland, John Rubin, Joseph Scheuhammer, and Andrew Witkin for their helpful discussions and, in some cases, for reading earlier drafts. We are also grateful to Alan Yuille for comments and corrections on the mathematics in the appendices. Preparation of this paper was supported in part by NSF and AFOSR under a combined grant for studies in Natural Computation, grant 79-23110-MCS, and by the AFOSR under an Image Understanding contract F49620-83-C-0135. Technical support was kindly provided by William Gilson; artwork was the creation of Julie Sandell and K. van Buskirk. Reprints may be obtained from D. Hoffman, School of Social Sciences, University of California, Irvine, CA 92717, U.S.A.

0010-0277/84/$10.10

Figure 1. *Some objects identifiable entirely from their profiles.*



Clearly your visual system is equipped to describe the shape of an object and to guess what the object is from its description. This guess may just be a first guess, perhaps best thought of as a first index into a memory of shapes, and might not be exactly correct; it may simply narrow the potential matches and trigger visual computations designed to narrow them further.

This first guess is more precisely described as an inference, one the truth of whose premises—the descriptions of shape—does not logically guarantee the truth of its conclusion—the identity of the object. Because the truth of the conclusion does not follow logically from the truth of the premises, the strength of the inference must derive from some other source. That source, we claim, is the regularity of nature, its uniformities and general laws. The design of the visual system exploits regularities of nature in two ways: they underlie the mental categories used to represent the world and they permit inferences from impoverished visual data to descriptions of the world.

Regularities of nature play both roles in the visual task of shape recognition, and both roles will be examined. We will argue that, just as syntactic analysis decomposes a sentence into its constituent structure, so the visual system decomposes a shape into a hierarchy of parts. Parts are not chosen arbitrarily; the mental category 'part' of shapes is based upon a regularity of nature discovered by differential topologists—transversality. This is an example of a regularity in the first role. The need arises for a regularity in the second role because although parts are three-dimensional, the eye delivers only a two-dimensional projection. In consequence the three-dimensional parts must be inferred from their two-dimensional projections. We propose that this inference is licensed by another regularity, this time from the field of singularity theory.

2. Why parts?

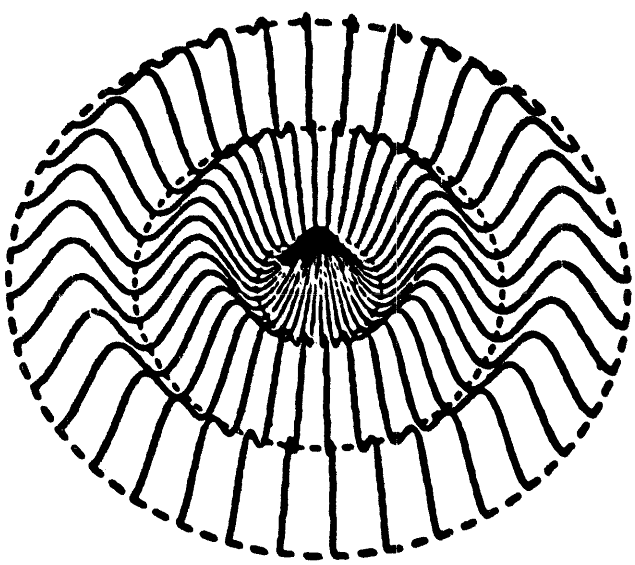
Before examining a part definition and its underlying regularity, we should ask: Given that one wants to recognize an object from its shape, why partition the shape at all? Could template matching or Fourier descriptors rise to the occasion? Possibly. What follows is not so much intended to deny this as to indicate the usefulness of parts.

To begin, then, an articulation of shapes into parts is useful because one never sees an entire shape in one glance. Clearly the back side is never visible (barring transparent objects), but even the front side is often partially occluded by objects interposed between the shape and the observer. A Fourier approach suffers because all components of a Fourier description can change radically as different aspects of a shape come into view. A part theory, on the other hand, can plausibly assume that the parts delivered by early vision correspond to the parts stored in the shape memory (after all, the contents of the shape memory were once just the products of early visual processing), and that the shape memory is organized such that a shape can be addressed by an inexhaustive list of its parts. Then recognition can proceed using the visible parts.

Parts are also advantageous for representing objects which are not entirely rigid, such as the human hand. A template of an outstretched hand would correlate poorly with a clenched fist, or a hand giving a victory sign, etc. The proliferation of templates to handle the many possible configurations of the hand, or of any articulated object, is unparsimonious and a waste of memory. If part theorists, on the other hand, pick their parts prudently (criteria for prudence will soon be forthcoming), and if they introduce the notion of spatial relations among parts, they can decouple configural properties from the shape of an object, thereby avoiding the proliferation of redundant mental models.

The final argument for parts to be considered here is phenomenological: we see them when we look at shapes. Figure 2, for instance, presents a cosine surface, which observers almost uniformly see organized into ring-like parts. One part stops and another begins roughly where the dotted circular contours are drawn. But if the figure is turned upside down the organization changes such that each dotted circular contour, which before lay between parts, now lies in the middle of a part. Why the parts change will be explained by the partitioning rule to be proposed shortly; the point of interest here is simply that our visual systems do in fact cut surfaces into parts.

Figure 2. *The cosine surface at first appears to be organized into concentric rings, one ring terminating and the next beginning approximately where the dashed circular contours are drawn. But this organization changes when the figure is turned upside down.*



3. Parts and uniformities of nature

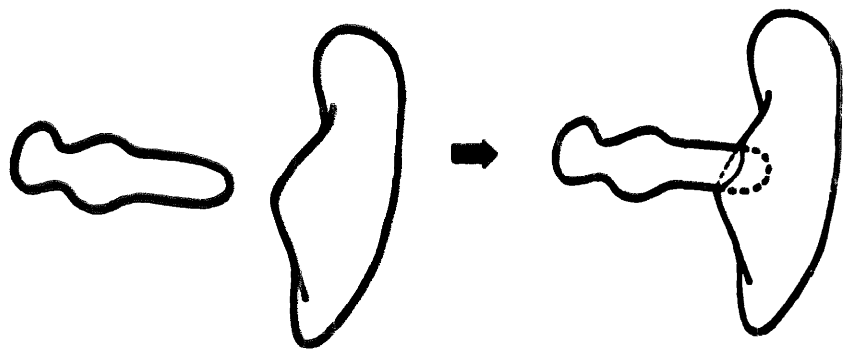
Certainly any proper subset of a surface is a part of that surface. This definition of part, however, is of little use for the task of shape recognition. And although the task of shape recognition constrains the class of suitable part definitions (see Section 5), it by no means forces a unique choice. To avoid an *ad hoc* choice, and to allow a useful correspondence between the world and mental representations of shape, the definition of part should be motivated by a uniformity of nature.[1]

One place not to look for a defining regularity is in the shapes of a part. One could say that all parts are cylinders, or cones, or spheres, or polyhedra, or some combination of these; but this is legislating a definition, not discovering a relevant regularity. And such a definition would have but limited applicability, for certainly not all shapes can be decomposed into just cylinders, cones, spheres, and polyhedra.

If a defining regularity is not to be found in part shapes, then another place

[1]Unearthing an appropriate uniformity is the most creative, and often most difficult, step in devising an explanatory theory for a visual task. Other things being equal, one wants the most general uniformity of nature possible, as this grants the theory and the visual task the broadest possible scope.

Figure 3. *An illustration of the transversality regularity. When any two surfaces interpenetrate at random they always meet in concave discontinuities, as indicated by the dashed contours.*



to look is part intersections. Consider the two three-dimensional blobs depicted in the left of Fig. 3 Certainly these two spatially separated shapes are different parts of this figure. Indeed, each spatially distinct object in a visual scene is a part of that scene. Now if two such separate objects are interpenetrated to form one new composite object, as shown in the right of Fig. 3, then the two objects, which were before separate parts of the visual scene, are surely now prime candidates to be parts of the new composite shape. But can we tell, simply by examining the new composite shape, what the original parts are? That is, is there a way to tell where one part stops and the next part begins on the new composite shape? Fortunately there is a way, one which depends on a regularity in the way two shapes generically intersect. This regularity is called transversality (for a detailed discussion of transversality see Guillemin and Pollack (1974)).

- *Transversality regularity*. When two arbitrarily shaped surfaces are made to interpenetrate they always[2] meet in a contour of concave discontinuity of their tangent planes.

To see this more clearly, observe the silhouette of the composite shape shown in the right of Fig. 3. Notice that this composite silhouette is not smooth at the two points where the silhouette of one of its parts intersects the silhouette of the other part. At these two points the direction of the silhouette's outline (i.e., its tangent direction) changes abruptly, creating a concave cusp (i.e., a cusp which points into the object, not into the

[2]The word *always* is best interpreted "with probability one assuming the surfaces interpenetrate at random".

background) at each of the two points. In fact, such concave discontinuities arise at every point on the surface of the composite shape where the two parts meet. These contours of concave discontinuity of the tangent plane of the composite shape will be the basis for a partitioning rule in the next section. But three observations are in order.

First, though it may sound esoteric, transversality is a familiar part of our everyday experience. A straw in a soft drink forms a circular concave discontinuity where it meets the surface of the drink. So too does a candle in a birthday cake. The tines of a fork in a piece of steak, a cigarette in a mouth, all are examples of this ubiquitous regularity.

Second, transversality does not double as a theory of part growth or part formation (D'Arcy Thompson, 1968). We are not claiming, for example, that a nose was once physically separated from the face and then got attached by interpenetration. We simply note that when two spatially separated shapes are interpenetrated, their intersection is transversal. Later we will see how this regularity underlies the visual definition of separate parts of any composite shape, such as the nose of a face or a limb of a tree, regardless of how the composite shape was created.

Finally, transversality does encompass movable parts. As mentioned earlier, one attraction of parts is that, properly chosen, they make possible a decoupling of configuration and shape in descriptions of articulated objects. But to do this the parts must cut an object at its articulations; a thumb–wrist part on the hand, for instance, would be powerless to capture the various spatial relations that can exist between the thumb and the wrist. Now the parts motivated by transversality will be the movable units, fundamentally because a transversal intersection of two surfaces remains transversal for small pertubations of their positions. This can be appreciated by reviewing Fig. 3. Clearly the intersection of the two surfaces remains a contour of concave discontinuity even as the two surfaces undergo small independent rotations and translations.

4. Partitioning: The minima rule

On the basis of the transversality regularity we can propose a first rule for dividing a surface into parts: divide a surface into parts along all contours of concave discontinuity of the tangent plane. Now this rule cannot help us with the cosine surface because this surface is entirely smooth. The rule must be generalized somewhat, as will be done shortly. But in its present form the rule can provide insight into several well-known perceptual demonstrations.

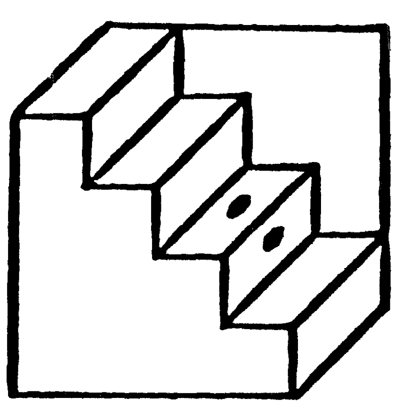
4.1. Blocks world

We begin by considering shapes constructed from polygons. Examine the staircase of Fig. 4. The rule predicts that the natural parts are the steps, and not the faces on the steps. Each step becomes a 'part' because it is bounded by two lines of concave discontinuity in the staircase. (A face is bounded by a concave and a convex discontinuity.) But the rule also makes a less obvious prediction. If the staircase undergoes a perceptual reversal, such that the 'figure' side of the staircase becomes 'ground' and *vice versa*, then the step boundaries must change. This follows because only *concave* discontinuities define step boundaries. And what looks like a concavity from one side of a surface must look like a convexity from the other. Thus, when the staircase reverses, convex and concave discontinuities must reverse roles, leading to new step boundaries. You can test this prediction yourself by looking at the step having a dot on each of its two faces. When the staircase appears to reverse note that the two dots no longer lie on a single step, but lie on two adjacent steps (that is, on two different 'parts').

Similar predictions from the rule can also be confirmed with more complicated demonstrations such as the stacked cubes demonstration shown in Fig. 5. The three dots which at first appear to lie on one cube, lie on three different cubes when the figure reverses.

Still another quite different prediction follows from our simple partitioning rule. If the rule does not define a unique partition of some surface, then the division of that surface into parts should be perceptually ambiguous (unless,

Figure 4. *The Schroder staircase, published by H. Schroder in 1858, shows that part boundaries change when figure and ground reverse. The two dots which at first appear to lie on one step suddenly seem to lie on two adjacent steps when the staircase reverses.*



of course, there are additional rules which can eliminate the ambiguity). An elbow-shaped block provides clear confirmation of this prediction (see Fig. 6). The only concave discontinuity is the vertical line in the crook of the elbow; in consequence, the rule does not define a unique partition of the block. Perceptually, there are three plausible ways to cut the block into parts (also shown in Fig. 6). All three use the contour defined by the partitioning rule, but complete it along different paths.

Figure 5. *Stacked cubes also show that parts change when figure and ground reverse. Three dots which sometimes lie on one cube will lie on three different cubes when the figure reverses.*

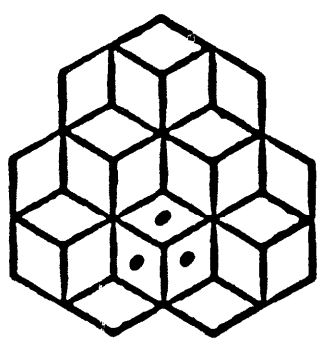
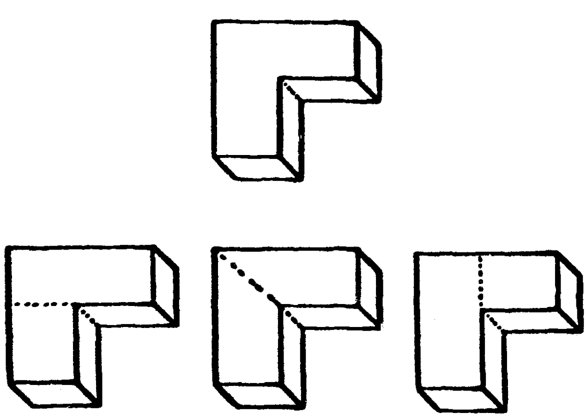


Figure 6. *Elbow-shaped blocks show that a rule partitioning shapes at concave discontinuities is appropriately conservative. The rule does not give a closed contour on the top block, and for good reason. Perceptually, three different partitions seem reasonable, as illustrated by the bottom three blocks.*



4.2. Generalization to smooth surfaces

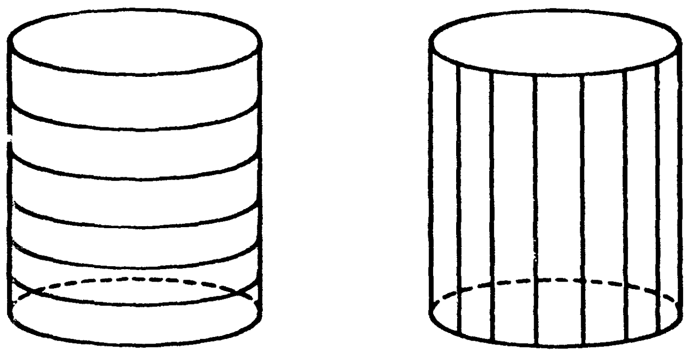
The simple partitioning rule directly motivated by transversality leads to interesting insights into our perception of the parts of polygonal objects. But how can the rule be generalized to handle smooth surfaces, such as the cosine surface? To grasp the generalization, we must briefly digress into the differential geometry of surfaces in order to understand three important concepts: surface normal, principal curvature, and line of curvature. Fortunately, although these concepts are quite technical, they can be understood intuitively.

The surface normal at a point on a surface can be thought of as a unit length needle sticking straight out of (orthogonal to) the surface at that point, much like the spines on a sea urchin. All the surface normals at all points on a surface are together called a field of surface normals. Usually there are two possible fields of surface normals on a surface—either outward pointing or inward pointing. A sphere, for instance, can either have the surface normals all pointing out like spines, or all pointing to its center. Let us adopt the convention that the field of surface normals is always chosen to point into the figure (i.e., into the object). Thus a baseball has inward normals whereas a bubble under water, if the water is considered figure, has outward normals. Reversing the choice of figure and ground on a surface implies a concomitant change in the choice of the field of surface normals. And, as will be discussed shortly, a reversal of the field of surface normals induces a change in sign of each principal curvature at every point on the surface.

It is often important to know not just the surface normal at a point but also how the surface is curving at the point. The Swiss mathematician Leonhard Euler discovered around 1760 that at any point on any surface there is always a direction in which the surface curves least and a second direction, always orthogonal to the first, in which the surface curves most. (Spheres and planes are trivial cases since the surface curvature is identical in all directions at every point.) These two directions at a point are called the principal directions at that point and the corresponding surface curvatures are called the principal curvatures. Now by starting at some point and always moving in the direction of the greatest principal curvature one traces out a line of greatest curvature. By moving instead in the direction of the least principal curvature one traces out a line of least curvature. On a drinking glass the family of lines of greatest curvature is a set of circles around the glass. The lines of least curvature are straight lines running the length of the glass (see Fig. 7).

With these concepts in hand we can extend the partitioning rule to smooth surfaces. Suppose that wherever a surface has a concave discontinuity we smooth the discontinuity somewhat, perhaps by stretching a taut skin over it.

Figure 7. *Lines of curvature are easily depicted on a drinking glass. Lines of greatest curvature are circles. Lines of least curvature are straight lines.*



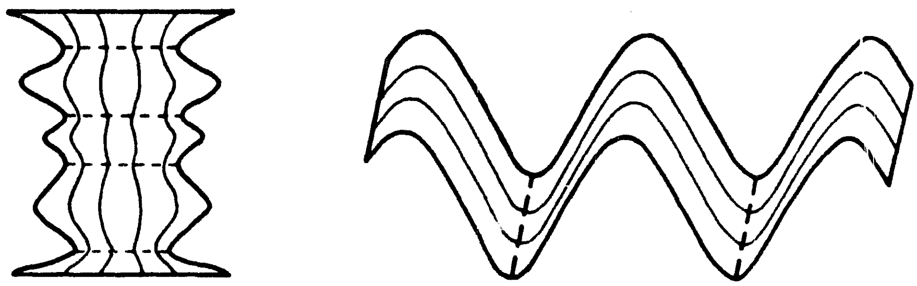
Then a concave discontinuity becomes a contour where, locally, the surface has greatest negative curvature. In consequence we obtain the following generalized partitioning rule for surfaces.

- *Minima rule.* Divide a surface into parts at loci of negative minima of each principal curvature along its associated family of lines of curvature.

The minima rule is applied to two surfaces in Fig. 8. The solid contours indicate members of one family of lines of curvature, and the dotted contours are the part boundaries defined by the minima rule. The bent sheet of paper on the right of Fig. 8 is particularly informative. The lines of curvature shown for this surface are sinusoidal, whereas the family of lines not shown are perfectly straight and thus have zero principal curvature (and no associated minima). In consequence, the product of the two principal curvatures at each point, called the *Gaussian curvature*, is always zero for this surface. Now if the Gaussian curvature is always zero on this surface, then the Gaussian curvature cannot be used to divide the surface into parts. But we see parts on this surface. Therefore whatever rule our visual systems use to partition surfaces cannot be stated entirely in terms of Gaussian curvature. In particular, the visual system cannot be dividing surfaces into parts at loci of zero Gaussian curvature (parabolic points) as has been proposed by Koenderink and van Doorn (1982b).

The minima rule partitions the cosine surface along the circular dotted contours shown in Fig. 2. It also explains why the parts differ when figure and ground are reserved. For when the page is turned upside down the visual system reverses its assignment of figure and ground on the surface (perhaps

Figure 8. *Part boundaries, as defined by the smooth surface partitioning rule, are indicated by dashed lines on several different surfaces. The families of solid lines are the lines of curvature whose minima give rise to the dashed partitioning contour.*



due to a preference for an interpretation which places the object below rather than overhead). When figure and ground reverse so does the field of surface normals, in accordance with the convention mentioned earlier. But simple calculations show that when the normals reverse so too does the sign of the principal curvatures. Consequently minima of the principal curvatures must become maxima and *vice versa*. Since minima of the principal curvatures are used for part boundaries, it follows that these part boundaries must also move. In sum, parts appear to change because the partitioning rule, motivated by the transversality regularity, uses minima of the principal curvatures, and because these minima relocate on the surface when figure and ground reverse. A more rigorous treatment of the partitioning rule is provided in Appendix 1.

5. Parts: Constraints from recognition

The task of visual recognition constrains one's choice of parts and part descriptions. We evaluate the part scheme proposed here against three such constraints—*reliability, versatility*, and *computability*—and then note a non-constraint, *information preservation*.

Reliability. Recognition is fundamentally a process of matching descriptions of what one sees with descriptions already in memory. Imagine the demands on memory and on the matching process if every time one looked at an object one saw different parts. A face, for example, which at one instant appeared to be composed of eyes, ears, a nose, and a mouth, might at a later instant metamorphose into a potpourri of eye–cheek, nose–chin, and mouth–

ear parts—a gruesome and unprofitable transmutation. Since no advantage accrues for allowing such repartitions, in fact since they are uniformly deleterious to the task of recognition, it is reasonable to disallow them and to require that the articulation of a shape into parts be invariant over time and over change in viewing geometry. This is the constraint of reliability (see Marr, 1982; Marr and Nishihara, 1978; Nishihara, 1981; Sutherland, 1968); the parts of a shape should be related reliably to the shape. A similar constraint governs the identification of linguistic units in a speech stream (Liberman *et al*, 1967; Fodor, 1983). Apparently the shortest identifiable unit is the syllable; shorter units like phones are not related reliably to acoustic parameters.

The minima rule satisfies this reliability constraint because it uses only surface properties, such as extrema of the principal curvatures, which are independent (up to a change in sign) of the coordinate system chosen to parametrize the surface (Do Carmo, 1976). Therefore the part boundaries do not change when the viewing geometry changes. (The part boundaries do change when figure and ground reverse, however.)

Versatility. Not all possible schemes for defining parts of surfaces are sufficiently versatile to handle the infinite variety in shape that objects can exhibit. Other things being equal, if one of two partitioning schemes is more versatile than another, in the sense that the class of objects in its scope properly contains the class of objects in the scope of the other scheme, the more versatile scheme is to be preferred. A partitioning scheme which can be applied to any shape whatsoever is most preferable, again other things being equal. This versatility constraint can help choose between two major classes of partitioning schemes: boundary-based and primitive-based. A *boundary-based* approach defines parts by their contours of intersection, not by their shapes. A *primitive-based* approach defines parts by their shapes, not by their contours of intersection (or other geometric invariants, such as singular points).

Shape primitives currently being discussed in the shape representation literature include spheres (Badler and Bajcsy, 1978; O'Rourke and Badler, 1979), generalized cylinders (Binford, 1971; Brooks *et al.*, 1979; Marr and Nishihara, 1978; Soroka, 1979), and polyhedra (Baumgart, 1972; Clowes, 1971; Guzman, 1969; Huffman, 1971; Mackworth, 1973; Waltz, 1975), to name a few (see Ballard and Brown, 1982). The point of interest here is that, for all the interesting work and conceptual advances it has fostered, the primitive-based approach has quite limited versatility. Generalized cylinders, for instance, do justice to animal limbs, but are clearly inappropriate for faces, cars, shoes, ... the list continues. A similar criticism can be levelled against each proposed shape primitive, or any conjunction of shape primitives. Perhaps a large enough conjunction of primitives could handle most shapes we do in fact encounter, but the resulting proposal would more resemble a restaurant menu than a theory of shape representation.

A boundary-based scheme on the other hand, if its rules use only the geometry (differential or global) of surfaces, can apply to any object whose bounding surface is amenable to the tools of differential geometry—a not too severe restriction.[3] Boundary rules simply tell one where to draw contours on a surface, as if with a felt marker. A boundary-based scheme, then, is to be preferred over a primitive-based scheme because of its greater versatility.

The advantage of a boundary-based scheme over a primitive-based scheme can also be put this way: using a boundary-based scheme one can locate the parts of an object without having any idea of what the parts look like. This is not possible with the primitive-based scheme. Of course one will want descriptions of the parts one finds using a boundary-based scheme, and one may (or may not) be forced to a menu of shapes at this point. Regardless, a menu of part shapes is not necessary for the task of locating parts. In fact a menu-driven approach restricts the class of shapes for which parts can be located. The minima rule, because it is boundary-based and uses only the differential geometry of surfaces, satisfies the versatility constraint—all geometric surfaces are within its scope.[4]

Computability. The partitioning scheme should in principle be computable using only information available in retinal images. Otherwise it is surely worthless. This is the constraint of *computability*. Computability is not to be confused with efficiency. Efficiency measures how quickly and inexpensively something can be computed, and is a dubious criterion because it depends not only on the task, but also on the available hardware and algorithms. Computability, on the other hand, states simply that the scheme must in principle be realizable, that it use only information available from images.

We have not yet discussed whether our parts are computable from retinal

[3]Shapes outside the purview of traditional geometric tools might well be represented by fractal-based schemes (Mandelbrot, 1982; Pentland 1983). Candidate shapes are trees, shrubs, clouds—in short, objects with highly crenulate or ill-defined surfaces.

[4]One must, however, discover the appropriate scales for a natural surface (Hoffman, 1983a, b; Witkin, 1983). The locations of the part boundaries depend, in general, on the scale of resolution at which the surface is examined. In consequence an object will not receive a single partitioning based on the minima rule, but will instead receive a nested hierarchy of partitions, with parts lower in the hierarchy being much smaller than parts higher in the hierarchy. For instance, at one level in the hierarchy for a face one part might be a nose. At the next lower level one might find a wart on the nose. The issue of scale is quite difficult and beyond the scope of this paper.

images (but see Appendix 2). And indeed, since minima of curvature are third derivative entities, and since taking derivatives exaggerates noise, one might legitimately question whether our part boundaries are computable. This concern for computability brings up an important distinction noted by Marr and Poggio (1977), the distinction between theory and algorithm. A theory in vision states what is being computed and why; an algorithm tells how. Our partitioning rule is a theoretical statement of what the part boundaries should be, and the preliminary discussion is intended to say why. The rule is not intended to double as an algorithm so the question of computability is still open. Some recent results by Yuille (1983) are encouraging though. He has found that directional zero-crossings in the shading of a surface are often located on or very near extrema of one of the principal curvatures along its associated lines of curvature. So it might be possible to read the part boundaries directly from the pattern of shading in an image, avoiding the noise problems associated with taking derivatives (see also Koenderink and van Doorn, 1980, 1982a). It is also possible to determine the presence of part boundaries directly from occluding contours in an image (see Appendix 2).

Information preservation: A non-constraint. Not just any constraints will do. The constraints must follow from the visual task; otherwise the constraints may be irrelevant and the resulting part definitions and part descriptions inappropriate. Because the task of recognition involves classification, namely the assignment of an individual to a class or a token to a type, not all the information available about the object is required. Indeed, in contrast to some possible needs for machine vision (Brady, 1982b, 1982c), we stress that a description of a shape for recognition need not be information preserving, for the goal is not to reconstruct the image. Rather it is to make explicit just what is key to the recognition process. Thus, what is critical is the form of the representation, what it makes explicit, how well it is tailored to the needs of recognition. Raw depth maps preserve all shape information of the visible surfaces, but no one proposes them as representations for recognition because they are simply not tailored for the task.

6. Projection and parts

We have now discussed how 'parts' of shapes may be defined in the three-dimensional world. However the eye sees only a two-dimensional projection. How then can parts be inferred from images? Again, we proceed by seeking a regularity of nature. As was noted earlier, the design of the visual system exploits regularities of nature in two ways: they underlie the mental categories used to represent the world and they license inferences from impoverished visual data to descriptions of the world. The role of transversality in the design of the mental category 'part' of shape is an example of the first case. In this section we study an example of the second case. We find that lawful properties of the singularities of the retinal projection permit an inference from retinal images to three-dimensional part boundaries. For simplicity we restrict attention to the problem of inferring part boundaries from silhouettes.

Consider first a discontinuous part boundary (i.e., having infinite negative curvature) on a surface embedded in three dimensions (Fig. 3). Such a contour, when imaged on the retina, induces a concave discontinuity in the resulting silhouette (notice the concave cusps in the silhouette of Fig. 3). Smooth part boundaries defined by the minima partitioning rule can also provide image cusps, as shown in the profiles of Fig. 1. It would be convenient to infer the presence of smooth and discontinuous part boundaries in three dimensions from concave discontinuities in the two-dimensional silhouette, but unfortunately other surface events can give rise to these discontinuities as well. A torus (doughnut), for instance, can have two concave discontinuities in its silhouette which do not fall at part boundaries defined by the minima rule (see Fig. 9).

Fortunately, it is rare that a concave discontinuity in the silhouette of an object does not indicate a part boundary, and when it does not this can be detected from the image data. So one can, in general, correctly infer the presence or absence of part boundaries from these concave discontinuities. The proof of this useful result (which is banished to Appendix 2) exploits regularities of the singularities of smooth maps between two-dimensional manifolds. We have seen how a regularity of nature underlies a mental category, *viz.*, 'part' of shape; here we see that another regularity (e.g., a singularity regularity) licenses an inference from the retinal image to an instance of this category.

Figure 9. *A torus can have concave discontinuities (inducated by the arrows) which do not correspond to part boundaries.*

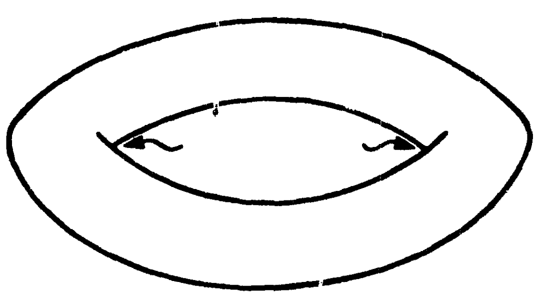
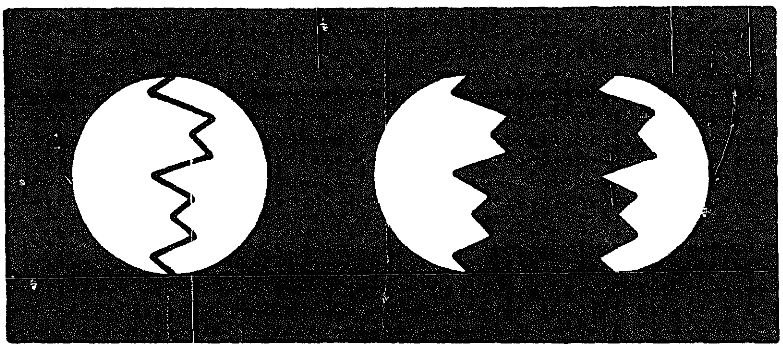


Figure 10. *A reversing figure, similar to Attneave (1974), appears either as an alternating chain of tall and short mountains or as a chain of tall mountains with twin peaks.*



The singularity regularity, together with transversality, motivates a first partitioning rule for plane curves: *Divide a plane curve into parts at concave cusps.* Here the word *concave* means concave with respect to the silhouette (figure) side of the plane curve. A concavity in the figure is, of course, a convexity in the ground.

This simple partitioning rule can explain some interesting perceptual effects. In Fig. 10, for instance, the same wiggly contour can look either like valleys in a mountain range or, for the reversed figure–ground assignment, like large, twin-peaked mountains. The contour is carved into parts differently when figure and ground reverse because the partitioning rule uses only concave cusps for part boundaries. And what is a concave cusp if one side of the contour is figure must become a convex cusp when the other side is figure, and *vice versa*. There is an obvious parallel between this example and the reversible staircase discussed earlier.

6.1. Geometry of plane curves

Before generalizing the rule to smooth contours we must briefly review two concepts from the differential geometry of place curves: principal normal and curvature. The principal normal at a point on a curve can be thought of as a unit length needle sticking straight out of (orthogonal to) the curve at that point, much like a tooth on a comb. All the principal normals at all points on a curve together form a field of principal normals. Usually there are two possible fields of principal normals—either leftward pointing or rightward pointing. Let us adopt the convention that the field of principal normals is always chosen to point into the figure side of the curve. Reversing the choice of figure and ground on a curve implies a concomitant change in the choice of the field of principal normals.

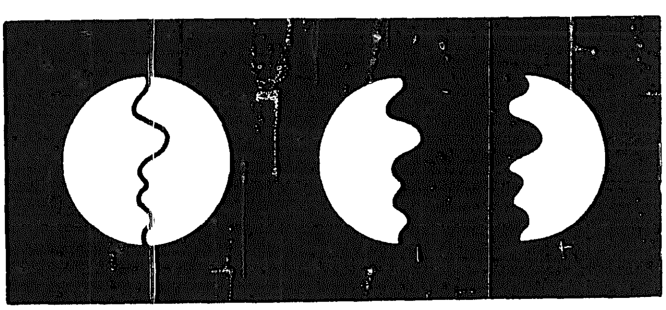
Curvature is a well-known concept. Straight lines have no curvature, circles have constant curvature, and smaller circles have higher curvature than larger circles. What is important to note is that, because of the convention forcing the principal normals to point into the figure, concave portions of a smooth curve have negative curvature and convex portions have positive curvature.

6.2. Parts of smooth curves

It is an easy matter now to generalize the partitioning rule. Suppose that wherever a curve has a concave cusp we smooth the curve a bit. Then a concave cusp becomes a point of negative curvature having, locally, the greatest absolute value of curvature. This leads to the following generalized partitioning rule: *Divide a plane curve into parts at negative minima of curvature.*[5]

Several more perceptual effects can be explained using this generalized partitioning rule. A good example is the reversing figure devised by Attneave (see Fig. 11). He found that by simply scribbling a line through a circle and separating the two halves one can create two very different looking contours. As Attneave (1971) points out, the appearance of the contour depends upon

Figure 11. *Attneave's reversing figure, constructed by scribbling a line down a circle. The apparent shape of a contour depends on which side is perceived as figure.*



[5]Transversality directly motivates using concave cusps as part boundaries. Only by smoothing do we include minima as well (both in the case of silhouette curves and in the case of part boundaries in three dimensions). Since the magnitude of the curvature at minima decreases with increased smoothing, it is useful to introduce the notion of the strength or goodness of a part boundary. The strength of a part boundary is higher the more negative the curvature of the minimum. Positive minima have the least strength, and deserve to be considered separately from the negative minima, a possibility suggested to us by Shimon Ullman.

which side is taken to be part of the figure, and does not depend upon any prior familiarity with the contour.

Now we can explain why the two halves of Attneave's circle look so different. For when figure and ground reverse, the field of principal normals also reverses in accordance with the convention. And when the principal normals reverse, the curvature at every point on the curve must change sign. In particular, minima of curvature must become maxima and *vice versa*. This repositioning of the minima of curvature leads to a new partitioning of the curve by the partitioning rule. In short, the curve looks different because it is organized into fundamentally different units or chunks. Note that if we chose to define part boundaries by inflections (see Hollerbach, 1975; Marr, 1977), or by both maxima and minima of curvature (see Duda and Hart, 1973), or by all tangent and curvature discontinuities (Binford, 1981), then the chunks would not change when figure and ground reverse.

A clear example of two very different chunkings for one curve can be seen in the famous face–goblet illusion published by Turton in 1819. If a face is taken to be figure, then the minima of curvature divide the curve into chunks corresponding to a forehead, nose, upper lip, lower lip, and chin. If instead the goblet is taken to be figure then the minima reposition, dividing the curve into new chunks corresponding to a base, a couple of parts of the stem, a bowl, and a lip on the bowl. It is probably no accident that the parts defined by minima are often easily assigned verbal labels.

Demonstrations have been devised which, like the face–goblet illusion, allow more than one interpretation of a single contour but which, unlike the face–goblet illusion, do not involve a figure–ground reversal. Two popular examples are the rabbit–duck and hawk–goose illusions (see Fig. 13). Because these illusions do not involve a figure–ground reversal, and because in consequence the minima of curvature never change position, the partitioning rule

Figure 12. *The reversing goblet can be seen as a goblet or a pair of facial profiles (adapted from Turton, 1819). Defining part boundaries by minima of curvature divides the face into a forehead, nose, upper lip, lower lip, and chin. Minima divide the goblet into a base, a couple parts of the stem, a bowl, and a lip on the bowl.*

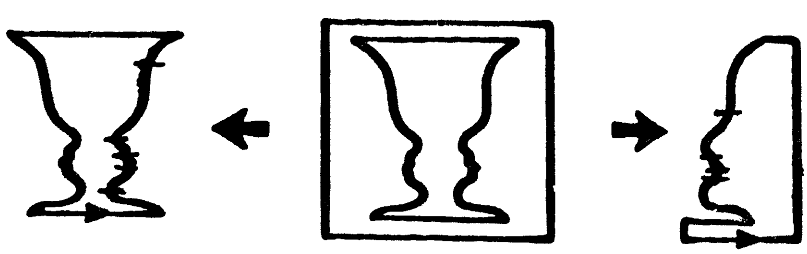
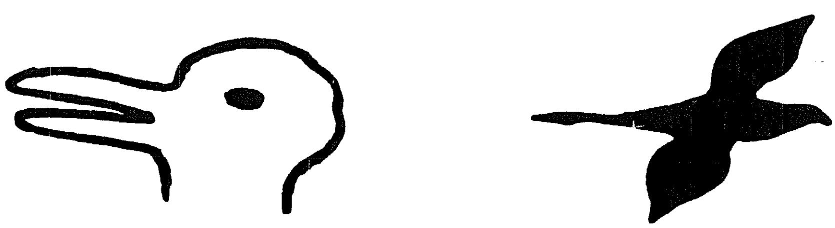


Figure 13. *Some ambiguous shapes do not involve a reversal of figure and ground. Consequently, the part boundaries defined by minima of curvature do not move when these figures change interpretations. In this illustration, for instance, a rabbit's ear turns into a duck's bill without moving, and a hawk's head turns into a goose's tail, again without moving.*



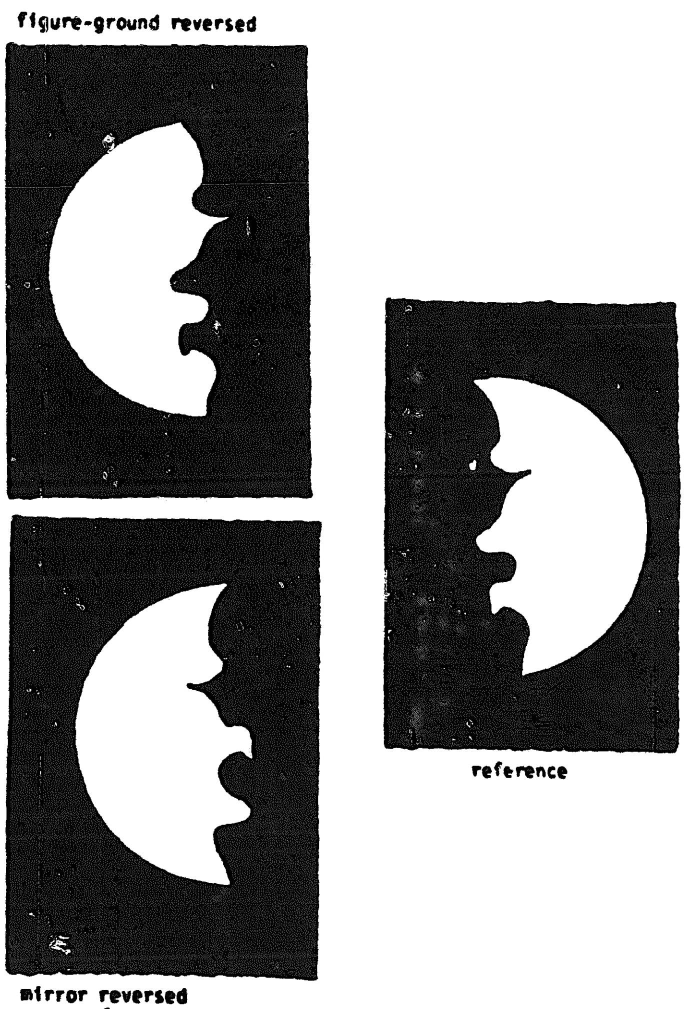
must predict that the part boundaries are identical for both interpretations of each of these contours. This prediction is easily confirmed. What is an ear on the rabbit, for instance, becomes an upper bill on the duck.

If the minima rule for partitioning is really used by our visual systems, one should expect it to predict some judgments of shape similarity. One case in which its prediction is counterintuitive can be seen in Fig. 14. Look briefly at the single half-moon on the right of the figure. Then look quickly at the two half-moons on the left and decide which seems more similar to the first (go ahead). In an experiment performed on several similar figures, we found that nearly all subjects chose the bottom half-moon as more similar. Yet if you look again you will find that the bounding contour for the top half-moon is identical to that of the right half-moon, only figure–ground reversed. The bounding contour of the bottom half-moon, however, has been mirror reversed, and two parts defined by minima of curvature have been swapped. Why does the bottom one still look more similar? The minima rule gives a simple answer. The bottom contour, which is not figure–ground reversed from the original contour, has the same part boundaries. The top contour, which is figure–ground reversed from the original, has entirely different part boundaries.

7. Holes: A second type of part

The minima rule for partitioning surfaces is motivated by a fact about generic intersections of surfaces: surfaces intersect transversally. As Fig. 3 illustrates, this implies that if two surfaces are interpenetrated and left together to form a composite object then the contour of their intersection is a contour of

Figure 14. *A demonstration that some judgments of shape similarity can be predicted by the minima partitioning rule. In a quick look, the bottom left half-moon appears more similar to the right half-moon than does the top left one. However the bounding contour of the top left half-moon is identical to that of the right half-moon, whereas the bounding contour of the bottom left half-moon has been mirror reversed and has had two parts interchanged.*



concave discontinuity on the composite surface. Now suppose instead that after the two surfaces are interpenetrated one surface is pulled out of the other, leaving behind a depression, and then discarded. The depression created in this manner has just as much motivation for being a 'part' on the basis of transversality as the parts we have discussed up to this point.

As can be seen by examining the right side of Fig. 3, the contour that divides one part from the other on the composite object is precisely the same contour that will delimit the depression created by pulling out the penetrating part. But whereas in the case of the composite object this contour is a contour of *concave* discontinuity, in the case of the depression this contour is a contour of *convex* discontinuity. And smoothing this contour leads to positive extrema of a principal curvature for the case of a depression. We are led to conclude that a shape can have at least two kinds of parts—'positive parts' which are bounded by negative extrema of a principal curvature, and 'negative parts' (holes) bounded by positive extrema of a principal curvature.

This result presents us with the task of finding a set of rules that determine when to use positive extrema or negative extrema as part boundaries. We do not have these rules yet, but here is an example of what such rules might look like. If a contour of negative extrema of a principal curvature is not a closed contour, and if it is immediately surrounded (i.e., no intervening extrema) by a closed contour of positive extrema of a principal curvature, then take the contour of positive extrema as the boundary of a (negative) part.

Note in any case that what we will not have are single parts bounded by both negative and positive extrema of a principal curvature.

8. Perception and induction

Inferences and regularities of nature have cropped up many times in the theory and discussions presented here. It is useful to explore their significance more fully.

Perceptual systems inform the perceiver about properties of the world she needs to know. The need might be to avoid being eaten, to find what is edible, to avoid unceremonious collisions, or whatever. The relevant knowledge might be the three-dimensional layout of the immediate surrounds, or that ahead lies a tree loaded with fruit, or that crouched in the tree is an unfriendly feline whose perceptual systems are also at work reporting the edible properties of the world. Regardless of the details, what makes the perceptual task tricky is that the data available to a sensorium invariably underdetermine the properties of the world that need to be known. That is, in general there are infinitely many states of the world which are consistent with the available sense data. Perhaps the best known example is that although the world is three-dimensional, and we perceive it as such, each retina is only two-dimensional. Since the mapping from the world to the retina is many-to-one, the possible states of the world consistent with a retinal image, or any series of retinal images, are many. The upshot of all this is that knowledge of the world is inferred. Inference lies at the heart of perception (Fodor and Pylyshyn, 1981; Gregory, 1970; Helmholtz, 1962; Hoffman, 1983b; Marr, 1982).

An inference, reduced to essentials, is simply a list of premises and a conclusion. An inference is said to be *deductively valid* if and only if the conclusion is logically guaranteed to be true given that the premises are true. So, for example, the following inference, which has three premises and one conclusion, is deductively valid: "A mapping from 3-D to 2-D is many-to-one. The world is 3-D. A retinal image is 2-D. Therefore a mapping from the world to a retinal image is many-to-one." An inference is said to be *inductively strong* if and only if it is unlikely that the conclusion is false while its premises are true, and it is not deductively valid (see Skyrms, 1975).[6] So the following inference is inductively strong: "The retinal disparities across my visual field are highly irregular. Therefore whatever I am looking at is not flat." Though this inference is inductively strong, it can prove false, as is in fact the case whenever one views a random dot stereogram.

In perceptual inferences the sensory data play the role of the premises, and the assertions about the state of the world are the conclusions. Since the state of the world is not logically entailed by the sensory data, perceptual inferences are not of the deductive variety—therefore they are inductive.

This is not good news. Whereas deductive inference is well understood, inductive inference is almost not understood at all. Induction involves a morass of unresolved issues, such as projectibility (Goodman, 1955), abduction (Levi, 1980; Peirce, 1931), and simplicity metrics (Fodor, 1975). These problems, though beyond the scope of this paper, apply with unmitigated force to perceptual inferences and are thus of interest to students of perception (Nicod, 1968).

But, despite these difficulties, consider the following question: If the premises of perceptual inferences are the sensory data and the conclusion is an assertion about the state of the world, what is the evidential relation between perceptual premises and conclusions? Or to put it differently, how is it possible that perceptual interpretations of sensory data bear a nonarbitrary (and

[6]The distinction between deductively valid and inductively strong inferences is not mere pedantry; the distinction has important consequences for perception, but is often misunderstood. Gregory (1970, p. 160), for instance, realizes the distinction is important for theories of perception, but then claims that "Inductions are generalizations of instances." This is but partly true. Inductive inferences may proceed from general premises to general conclusions, from general premises to particular conclusions, as well as from particular premises to general conclusions (Skyrms, 1975). The distinction between inductive and deductive inferences lies in the evidential relation between premises and conclusions.

even useful) relation to the state of the world? Or to put it still differently, why are perceptual inferences inductively strong?

Surely the answer must be, at least in part, that since the conclusion of a perceptual inference is a statement about the world, such an inference can be inductively strong only if it is motivated by laws, regularities, or uniformities of nature. To see this in a more familiar context, consider the following inductively strong inference about the world: "If I release this egg, it will fall". The inference here is inductively strong because it is motivated by a law of nature—gravity. Skeptics, if there are any, will end up with egg on their feet.

Laws, regularities, and uniformities in the world, then, are crucial for the construction of perceptual inferences which have respectable inductive strength. Only by exploiting the uniformities of nature can a perceptual system overcome the paucity of its sensory data and come to useful conclusions about the state of the world.

If this is the case, it has an obvious implication for perceptual research: identifying the regularities in nature which motivate a particular perceptual inference is not only a good thing to do, but a *sine qua non* for explanatory theories of perception.[7] An explanatory theory must state not only the premises and conclusion of a particular perceptual inference, but also the lawful properties of the world which license the move from the former to the latter. Without all three of these ingredients a proposed theory is incomplete.

---

[7]At least two conditions need to be true of a regularity, such as rigidity, for it to be useful: (1) It should in fact be a regularity. If there were not rigid objects in the world, rigidity would be useless. (2) It should allow inductively strong inferences from images to the world, by making the 'deception probability', to be defined shortly, very close to zero. For instance, let w (world) stand for the following assertion about four points in the world: "are in rigid motion in 3-D". Let i (image) stand for the following assertion about the retinal images of the same four points: "have 2-D positions and motions consistent with being the projections of rigid motion in 3-D". Then what is the probability of w given i? The existence of rigid objects does not in itself make this conditional probability high. Using Bayes' theorem we find that $P(w|i) = P(w) \cdot P(i|w)/[P(w) \cdot P(i|w) + P(-w) \cdot P(i|-w)]$. Since the numerator and the first term of the denominator are identical, this conditional probability is near one only if $P(w) \cdot P(i|w) \gg P(-w) \cdot P(i|-w)$. And since $P(-w)$, though unknown is certainly much greater than zero, $P(w|i)$ is near one only if $P(i|-w)$—let's call this the 'deception probability'—is near zero. Only if the deception probability is near zero can the inference from the image to the world be inductively strong. A major goal of 'structure from motion' proofs (Bobick, 1983; Hoffman and Flinchbaugh, 1982; Longuet-Higgins and Prazony, 1981; Richards *et al.*, 1983; Ullman, 1979) is to determine under what conditions this deception probability is near zero. Using an assumption of rigidity, for instance, Ullman has found that with three views of three points the deception probability is one, but with three views of four points it is near zero.

9. Conclusion

The design of the visual system exploits regularities of nature in two ways: they underlie the mental categories used to represent the world and they license inferences from incomplete visual data to useful descriptions of the world. Both uses of regularities underlie the solution to a problem in shape recognition. Transversality underlies the mental category 'part' of shape; singularities of projection underlie the inference from images to parts in the world.

The partitioning rules presented in this paper are attractive because (1) they satisfy several constraints imposed by the task of shape recognition, (2) they are motivated by a regularity of nature, (3) the resulting partitions look plausible, and (4) the rules explain and unify several well-known visual illusions.

Remaining, however, is a long list of questions to be answered before a comprehensive, explanatory theory of shape recognition is forthcoming. A partial list includes the following. How are the partitioning contours on surfaces to be recovered from two-dimensional images? How should the surface parts be described? All we have so far is a rule for cutting out parts. But what qualitative and metrical descriptions should be applied to the resulting parts? Can the answer to this question be motivated by appeal to uniformities and regularities in the world? What spatial relations need to be computed between parts? Although the part definitions don't depend upon the viewing geometry, is it possible or even necessary that the predicates of spatial relations do (Rock, 1974; Yin, 1970)? How is the shape memory organized? What is the first index into this memory?

The task of vision is to infer useful descriptions of the world from changing patterns of light falling on the eye. The descriptions can be reliable only to the extent that the inferential processes which build them exploit regularities in the visual world, regularities such as rigidity and transversality. The discovery of such regularities, and the mathematical investigation of their power in guiding particular visual inferences, are promising directions for the researcher seeking to understand human vision.

Appendix 1

Surface partitioning in detail

This appendix applies the surface partitioning rule to a particular class of surfaces: surfaces of revolution. The intent is to convey a more rigorous understanding of the rule and the partitions it yields. Since this section is quite mathematical, some readers might prefer to look at the results in Fig. 16 and skip the rest.

Notation. Tensor notation is adopted in this section because it allows concise expression of surface concepts, (see Dodson and Poston, 1979; Hoffman, 1983a; Lipschutz, 1969). A vector in $\Re^3$ is $\mathbf{x} = (x^1, x^2, x^3)$. A point in the parameter plane is (u^1, u^2). A surface patch is $\mathbf{x} = \mathbf{x}(u^1, u^2) = (x^1(u^1, u^2), x^2(u^1, u^2), x^3(u^1, u^2))$. Partial deriviatives a denoted by subscripts:

$$\mathbf{x}_1 = \frac{\partial \mathbf{x}}{\partial u^1}, \mathbf{x}_2 = \frac{\partial \mathbf{x}}{\partial u^2}, \mathbf{x}_{12} = \frac{\partial^2 \mathbf{x}}{\partial u^1 \partial u_2}, \text{etc.}$$

A tangent vector is $d\mathbf{x} = \mathbf{x}_1 du^1 + \mathbf{x}_2 du^2 = \mathbf{x}_i du^i$ where the Einstein summation convention is used. The first fundamental form is

$$| = d\mathbf{x} \cdot d\mathbf{x} = \mathbf{x}_i \cdot \mathbf{x}_j du^i du^j = g_{ij} du^i du^j$$

where the g_{ij} are the first fundamental coefficients and $i, j = 1, 2$.

The differential of the normal vector is the vector $d\mathbf{N} = \mathbf{N}_i du^i$ and the second fundamental form is

$$|| = d^2\mathbf{x} \cdot \mathbf{N} = \mathbf{x}_{ij} \cdot \mathbf{N} du^i du^j = b_{ij} du^i du^j$$

where the b_{ij} are the second fundamental coefficients and $i, j = 1, 2$.

A plane passing through a surface S orthogonal to the tangent plane of S at some point P and in a direction $du^i{:}du^j$ with respect to the tangent plane intersects the surface in a curve whose curvature at P is the *normal curvature* of S at P in the direction $du^i{:}du^j$. The normal curvature in a direction $du^i{:}du^j$ is $k_n = ||/|$. The two perpendicular directions for which the values of k_n take on maximum and minimum values are called the *principal directions*, and the corresponding curvatures, k_1 and k_2, are called the *principal curvatures*. The *Gaussian curvature* at P is $K = k_1 k_2$. A *line of curvature* is a curve on a surface whose tangent at each point is along a principal direction.

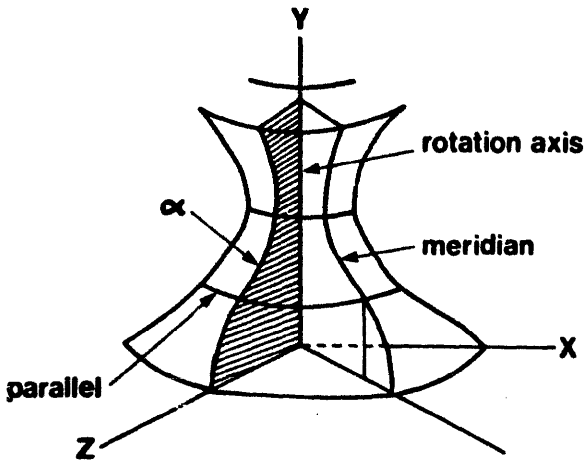
Partitions of a surface of revolution. A surface of revolution is a set $S \subset \Re^3$ obtained by rotating a regular plane curve α about an axis in the plane which does not meet the curve. Let the x^1x^3 plane be the plane of α and the x^3 axis the rotation axis. Let

$$\alpha(u^1) = (x(u^1), z(u^1)),\ a < u^1 < b,\ z(u^1) > 0.$$

Let u^2 be the rotation angle about the x^3 axis. Then we obtain a map

$$\mathbf{x}(u^1, u^2) = (x(u^1)\cos(u^2), x(u^1)\sin(u^2), z(u^1))$$

Figure 15. *Surface of revolution.*



from the open set $U = \{(u^1, u^2) \in \Re^2; 0 < u^2 < 2\pi, a < u^1 < b\}$ into S (Fig. 15). The curve α is called the *generating curve* of S, and the x^3 axis is the *rotation axis* of S. The circles swept out by the points of α are called the *parallels* of S, and the various placements of α on S are called the *meridians* of S.

Let $\cos(u^2)$ be abbreviated as c and $\sin(u^2)$ as s. Then $\mathbf{x}_1 = (x_1c, x_1s, z_1)$ and $\mathbf{x}_2 = (-xs, xc, 0)$. The first fundamental coefficients are then

$$g_{ij} = \mathbf{x}_i \cdot \mathbf{x}_j = \begin{pmatrix} x_1^2 + z_1^2 & 0 \\ 0 & x^2 \end{pmatrix}.$$

The surface normal is

$$\mathbf{N} = \frac{\mathbf{x}_1 \times \mathbf{x}_2}{|\mathbf{x}_1 \times \mathbf{x}_2|} = \frac{(z_1c, z_1s, -x_1)}{\sqrt{z_1^2 + x_1^2}}.$$

If we let u be arc length along α then $\sqrt{z_1^2 + x_1^2} = 1 = g_{11}$ and

$$\mathbf{N} = (z_1c, z_1s, -x_1).$$

The second fundamental coefficients are

$$b_{ij} = \mathbf{x}_{ij} \cdot \mathbf{N} = \begin{pmatrix} x_{11}z_1 - x_1z_{11} & 0 \\ 0 & -xz_1 \end{pmatrix}.$$

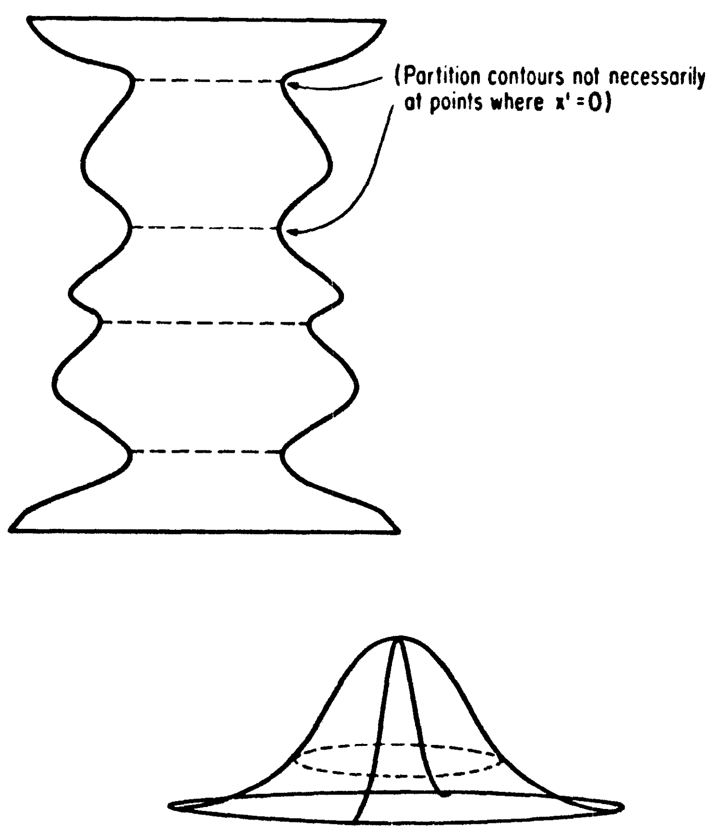
Since $g_{12} = b_{12} = 0$ the principal curvatures of a surface of revolution are

$$k_1 = b_{11}/g_{11} = x_{11}z_1 - x_1z_{11}$$
$$k_2 = b_{22}/g_{22} = -z_1/x.$$

The expression for k_1 is identical to the expression for the curvature along α. In fact the meridians (the various positions of α on S) are lines of curvature, as are the parallels. The curvature along the meridians is given by the expression for k_1 and the curvature along the parallel is given by the expression for k_2. The expression for k_2 is simply the curvature of a circle of radius x multiplied by the cosine of the angle that the tangent to α makes with the axis of rotation.

Observe that the expressions for k_1 and k_2 depend only upon the parameter u^1, not u^2. In particular, since k_2 is independent of u^2 there are no extrema or inflections of the normal curvature along the parallels. The parallels are circles. Consequently no segmentation contours arise from the lines of curvature associated with k_2. Only the minima of k_1 along the meridians are used for segmentation. Fig. 16 shows several surfaces of revolution with the

Figure 16. *Partitions on surfaces of revolution.*



minima of curvature along the meridians marked. The resulting segmentation contours appear quite natural to human observers.

As a surface of revolution is flattened along one axis, the partitioning contours which are at first circles become, in general, more elliptical and bow slightly up or down.

Appendix 2

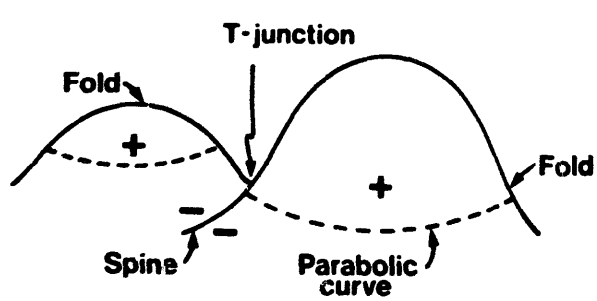
Inferring part boundaries from image singularities

In general, a concave discontinuity in a silhouette indicates a part boundary (as defined by the minima rule) on the imaged surface. This appendix makes this statement more precise and then examines a special case.

Only two types of singularity can arise in the projection from the world to the retina (Whitney, 1955). These two types are *folds* and *spines* (see Fig. 17). Intuitively, folds are the contours on a surface where the viewer's line of sight would just graze the surface, and a spine separates the visible portion of a fold from the invisible. A contour on the retina corresponding to a fold on a surface is called an *outline* (Koenderink and van Doorn, 1976, 1982b). A *termination* is a point on the retina corresponding to a spine on a surface. A *T-junction* (see Fig. 17) occurs where two outlines cut each other.

We wish to determine the conditions in which a T-junction indicates the presence of a part boundary. Two results are useful here. First, the sign of curvature of a point on an outline (projection of a fold) is the sign of the Gaussian curvature at the corresponding surface point (Koenderink and van Doorn, 1976, 1982b). Convex portions of the outline indicate positive Gaussian curvature, concave portions indicate negative Gaussian curvature, and inflections indicate zero Gaussian curvature. Second, the spine always occurs

Figure 17. *Singularities of the retinal projection.*



at a point of negative Gaussian curvature. That is, the visible portion of a fold always ends in a segment whose projected image is concave (Koenderink and van Doorn, 1982b).

The scheme of the proof is the following. Suppose that the folds on both sides of a T-junction have convex regions, as shown in Fig. 17. Then the sign of the Gaussian curvature is positive, and in fact both principal curvatures are positive, in these two regions. Now the presence of a spine indicates that these regions of positive Gaussian curvature are separated by a region of negative Gaussian curvature. This implies that the principal curvature associated with one family of lines of curvature is negative in this region. But then the principal curvature along this family of lines of curvature must go from positive to negative and back to positive as the lines of curvature go from one hill into the valley and back up the other hill. If this is true, then in the generic case the principal curvature will go through a negative minimum somewhere in the valley—and we have a part boundary.

There are two cases to consider. In the first the loci where one principal curvature goes from positive to negative (parabolic curves) surround each hill. In the second case the parabolic curve surrounds the valley between the two hills. We consider only the first case.

In the first case there are two ways that the lines of curvature entering the valley from one parabolic curve might fail to connect smoothly with lines of curvature entering the valley from the other parabolic curve: they might intersect orthogonally or not at all. If they intersect orthogonally then the two principal curvatures must both be negative, and the Gaussian curvature, which is the product of the two principal curvatures, must be positive. But the valley between the parabolic contours has negative Gaussian curvature, a contradiction.

If the lines of curvature fail to intersect then there must be a singularity in the lines of curvature somewhere in the region having negative Gaussian curvature. However, "The net of lines of curvature may have singular properties at umbilical points, and at them only." (Hilbert and Cohn-Vossen, 1952, p. 187). Umbilical points, points where the two principal curvatures are equal, can only occur in regions of positive Gaussian curvature—again a contradiction. (Here we assume the surface is smooth. A singularity could also occur if the surface were not smooth at one point in the valley. But in the generic case part boundaries would still occur.)

This proof requires that the two folds of a T-junction each have a convex region. The two folds of T-junctions on a torus do not satisfy this condition—they are always concave. Thus it is a simple matter to determine from an image when a T-junction warrants the inference of a part boundary.

The proof outlined here is a special case. A general proof is needed which specifies when a concave cusp in a silhouette indicates the presence of a part boundary or two different objects. The more general proof would not use the relation between spine points and Gaussian curvature. The proof might run roughly as follows: a concave cusp is a double point in the projection. A line connecting the two points on the surface which project to the cusp necessarily lies outside the surface between the two points. But then the surface is not convex everywhere between these two points. Consequently there is a concave discontinuity (part boundary) between the points or the Gaussian curvature must go negative. If the Gaussian curvature goes from positive (convex) to negative and then back to positive (convex), one of the principal curvatures must also. But this implies it has a negative minimum, in the general case, and so we have a smooth part boundary.

References

Attneave, F. (1974) Multistability in perception, *Scient. Am.*, *225*, 63–71.

Badler, N. and Bajcsy, R. (1978) Three-dimensional representations for computer graphics and computer vision. *Comp. Graph.*, *12*, 153–160.

Ballard, D. and Brown, C. (1982) *Computer Vision*, Englewood Cliffs, N.J., Prentice-Hall.

Baumgart, B. (1972) *Winged edge polyhedron representation*. STAN-CS-320, AIM-179, Stanford AI Lab.

Binford, T. (1971) Visual perception by computer. IEEE Conf. Syst. Cont., Miami.

Binford, T. (1981) Inferring surfaces from images. *Art. Intell.*, *17*, 205–244.

Bobick, A. (1983) A hybrid approach to structure-from-motion. Association for Computing Machinery Workshop on Motion: Representation and Perception.

Brady, J.M. (1982a) Parts description and acquisition using vision. *Proc. Soc. Photo-opt. Instrument. Eng.*

Brady, J.M. (1982b) Criteria for representations of shape. In A. Rosenfeld and J. Beck (eds.), *Human and Machine Vision*.

Brady, J.M. (1982c) Describing visible surfaces. In A. Hanson and E. Riseman (eds.), *Computer Vision Systems*.

Brooks, R., Greiner Russell, and Binford, T. (1979) The ACRONYM model based vision system. *Proc. Int. Joint Conf. Art. Intell.*, *6*, 105–113.

Clowes M. (1971) On seeing things. *Art. Intell.*, *2*, 79–116.

Dennett, D. (1978) Intentional systems. In *Brainstorms*. Montgomery, VT, Bradford.

Do Carmo, M. (1976) *Differential Geometry of Curves and Surfaces*. Englewood Cliffs, NJ, Prentice-Hall.

Dodson, D. and Poston, T. (1977) *Tensor Geometry*. London, Pitman.

Duda, R. and Hart, P. (1973) *Pattern Classification and Scene Analysis*. New York, Wiley.

Fodor, J. (1975) *The Language of Thought*, Cambridge, MA, Harvard University Press.

Fodor, J. (1983) *The Modularity of Mind*, Cambridge, MA, MIT Press.

Fodor, J. and Pylyshyn, Z. (1981) How direct is visual perception?: Some reflections on Gibson's "Ecological Approach". *Cog.*, *9*, 139–196.

Goodman, N. (1955) *Fact, Fiction and Forecast*. Cambridge, MA, Harvard University Press.

Gregory, R. (1970) *The Intelligent Eye*. New York, McGraw-Hill.

Guillemin, V. and Pollack, A. (1974) *Differential Topology*. Englewood Cliffs, NJ, Prentice-Hall.

Guzman, A. (1969) Decomposition of a visual scene into three-dimensional bodies. In A. Grasseli (ed.), *Automatic Interpretation and Classification of Images*, New York, Academic Press.

Helmholtz, H. (1962) *Treatise on Physiological Optics, Volume 3*, Dover reprint.

Hilbert, D. and Cohn-Vossen, S. (1952) *Geometry and the Imagination*, New York, Chelsea.

Hoffman, D. (1983a) Representing Shapes for Visual Recognition. MIT Ph.D. thesis.

Hoffman, D. (1983b) The interpretation of visual illusions, *Scient. Am.*, *249*, 154–162.

Hoffman D. and Flinchbaugh, B. (1982) The interpretation of biological motivation, *Biol. Cybernet.*, *42*, 195–204.

Hoffman, D. and Richards, W. (1982) Representing smooth plane curves for visual recognition: Implications for figure-ground reversal. *Proc. Am. Ass. Art. Intell.*, 5–8.

Hollerbach, J. (1975) *Hierarchical Shape Description of Objects by Selection and Modification of Prototypes*, MIT AI-TR-346.

Huffman, D. (1971) Impossible objects as nonsense sentences. *Mach. Intell.*, *6*.

Koenderink, J. and van Doorn, A. (1976) The singularities of the visual mapping. *Biol. Cybernet.*, *24*, 51–59.

Koenderink, J. and van Doorn, A. (1979) The internal representation of solid shape with respect to vision. *Biol. Cybernet.*, *32*, 211–216.

Koenderink, J. and van Doorn, A. (1980) Photometric invariants related to solid shape. *Optica Acta*, *7*, 981–996.

Koenderink, J. and van Doorn, A. (1982a) Perception of solid shape and spatial lay-out through photometric invariants. In R. Trappl (ed.), *Cybernetics and Systems Research*, Amsterdam, North-Holland.

Koenderink, J. and van Doorn, A. (1982b) The shape of smooth objects and the way contours end. *Perception*, *11*, 129–137.

Levi, I. (1980) *The Enterprise of Knowledge*, Cambridge, MA, MIT Press.

Liberman, A., Cooper, F., Shankweiler, D., and Studdert-Kennedy, M. (1967) The perception of the speech code. *Psychol. Rev.*, *74*, 431–461.

Lipschutz, M. (1969) *Differential Geometry*, (Schaum's Outline), New York, McGraw-Hill.

Longuet-Higgins, H.C. and Prazdny, K. (1981) The interpretation of a moving retinal image. *Proc. R. Soc. Lond.*, *B208*, 385–397.

Mackworth, A. (1973) Interpreting pictures of polyhedral scenes. *Art. Intell.*, *4*, 121–137.

Mandelbrot, B. (1982) *The Fractal Geometry of Nature*. San Francisco, Freeman.

Marr, D. (1977) Analysis of occluding countour. *Proc. R. Soc. Lond.*, *B197*, 441–475.

Marr, D. (1982) *Vision*. San Francisco, Freeman.

Marr, D. and Nishihara, H.K. (1978) Representation and recognition of the spatial organization of three-dimensional shapes. *Proc. R. Soc. Lond.*, *B200*, 269–294.

Marr, D. and Poggio, T. (1977) From understanding computation to understanding neural circuitry. *Neurosci. Res. Prog. Bull.*, *15*, 470–488.

Nicod, J. (1968) *Geometry and Induction*, Berkeley, University of California Press.

Nishihara, H.K. (1981) Intensity, visible-surface, and volumetric representations. *Art. Intell.*, *17*, 265–284.

O'Rourke, J. and Badler, N. (1979) Decomposition of three-dimensional objects into spheres. *IEEE Trans. Pattern Anal. Mach. Intell.*, *1*.

Pentland, A. (1983) Fractal-based description. *Proc. Int. Joint Conf. Art. Intell.*

Peirce, C. (1931) *Collected Papers*, Cambridge, MA, Harvard University Press.

Richards, W., Rubin, J.M. and Hoffman, D.D. (1983) Equation counting and the interpreation of sensory data. *Perception*, *11*, 557–576, and MIT AI Memo 618 (1981).

Rock, I. (1974) The perception of disoriented figures. *Scient. Am. 230*, 78–85.

Skyrms, B. (1975) *Choice and Chance*. Belmont, Wadsworth Publishing Co.

Soroka, B. (1979) Generalized cylinders from parallel slices. *Proc. Pattern Recognition and Image processing*, 421–426.

Spivak, M. (1970) *Differential Geometry, Volume 2*, Berkeley, Publish or Perish.

Sutherland, N.S. (1968) Outlines of a theory of visual pattern recognition in animals and man. *Proc. R. Soc. Lond., B171*, 297-317.

Thompson, D'Arcy (1968) *On Growth and Form*, Cambridge, University of Cambridge Press.

Turton, W. (1819) *A Conchological Dictionary of the British Islands*, (frontispiece), printed for John Booth, London. [This early reference was kindly pointed out to us by J.F.W. McOmie.]

Ullman, S. (1979) *The Interpretation of Visual Motion*. Cambridge, MA, MIT Press.

Waltz, D. (1975) Understanding line drawings of scenes with shadows. In P. Winston (ed.), *The Psychology of Computer Vision*, New York, McGraw-Hill.

Whitney, H. (1955) On singularities of mappings of Euclidean spaces. I. Mappings of the plane into the plane. *Ann. Math.*, *62*, 374–410.

Witkin, A.P. (1983) Scale-space filtering. *Proc. Int. Joint Conf. Artificial Intelligence.*

Yin, R. (1970) Face recognition by brain-injured patients: A dissociable ability? *Neuropsychologia*, *8*, 395–402.

Yuille, A. (1983) *Scaling theorems for zero-crossings*. MIT AI Memo 722.

Résumé

Les auteurs suggèrent que le système visuel pour la reconnaissance des objets, décompose les formes en éléments et qu'il utilise pour cela une règle définissant les frontères de ces éléments plutôt que leurs formes. Cette règle exploite une régularité de la nature: la transversalité. Les éléments, leurs descriptions et leurs relations spatiales fournissent un premier index dans la mémoire des formes. On peut avec cette règle rendre compte de plusieurs illusions visuelles. Les auteurs insistent sur le rôle de l'inférence inductive et concluent en indiquant les problèmes non résolus.