Perception, 1989, volume 18, pages 817-826

Parts of visual objects: an experimental test of the minima rule

Myron L Braunstein, Donald D Hoffman, Asad Saidpour
Department of Cognitive Sciences, University of California, Irvine, CA 92717, USA
Received 2 February 1989, in revised form 26 July 1989

Abstract. Three experiments were conducted to test Hoffman and Richards's (1984) hypothesis that, for purposes of visual recognition, the human visual system divides three-dimensional shapes into parts at negative minima of curvature. In the first two experiments, subjects observed a simulated object (surface of revolution) rotating about a vertical axis, followed by a display of four alternative parts. They were asked to select a part that was from the object. Two of the four parts were divided at negative minima of curvature and two at positive maxima. When both a minima part and a maxima part from the object were presented on each trial (experiment 1), most of the correct responses were minima parts (101 versus 55). When only one part from the object—either a minima part or a maxima part—was shown on each trial (experiment 2), accuracy on trials with correct minima parts and correct maxima parts did not differ significantly. However, some subjects indicated that they reversed figure and ground, thereby changing maxima parts into minima parts. In experiment 3, subjects marked apparent part boundaries. 81% of these marks indicated minima parts, 10% of the marks indicated maxima parts, and 9% of the marks were at other positions. These results provide converging evidence, from two different methods, which supports Hoffman and Richards's minima rule.

1 Introduction

How does the human visual system represent the three-dimensional shapes of objects? Almost certainly this question has more than one answer. We interact with objects in diverse ways and, as it is unlikely that a single representation of shape is adequate to this diversity, it is natural to expect that we use different representations for different modes of interaction. If, for instance, one wishes to grasp an object then one would do well to represent the orientation and position of its shape relative to one's body; these relational features of its shape are critical to the task of grasping. If, on the other hand, one wishes to recognize an object then one would do well to represent its shape in a manner independent, as much as possible, from its spatial relationship to one's body; the intrinsic geometry of its shape, not its extrinsic relationship to the viewer, is critical to the task of recognition.

These ideas have circulated for some time (see, eg Marr 1982) and enjoy a fair degree of acceptance. Given this framework, the project for researchers then becomes, in large part, to propose specific representations of shape for specific tasks, to demonstrate the theoretical adequacy of each representation for its associated task, to build simulations, and to conduct psychophysical tests. It is to this last task, the psychophysical testing of a proposed representation, that this paper is devoted. The proposal we investigate is the *minima rule* developed in Hoffman (1983a, 1983b), Hoffman and Richards (1984), Bennett and Hoffman (1987), and Beusmans et al (1987).

The minima rule is a rule for dividing visual shapes into subunits, or 'parts', for the purpose of facilitating visual recognition. The precise statement of the rule is as follows: *Divide a surface into parts at negative minima, along lines of curvature, of the associated principal curvature.* Intuitively, the minima rule divides surfaces into parts at high curvature points within 'dents' or 'valleys' on the surface. Figure 1 shows an example: solid lines are lines of curvature and dashed lines indicate part boundaries as defined by the minima rule. Several more examples are given in Hoffman and Richards (1984).

Briefly, the motivations for this rule are four. First, if two separate objects interpenetrate each other then, at each point along the contour of their intersection, the tangent planes to the surfaces of the two objects almost certainly do not coincide. If one smooths the surface of the composite object in the neighborhood of this contour, then one obtains a contour as defined by the minima rule (Bennett and Hoffman 1987). This contour naturally divides the surface of the composite object into two parts, each part corresponding to one of the two interpenetrated objects. Second, the rule is stated entirely in terms of the intrinsic geometry of surfaces, so that the parts it defines do not depend on the relative positions of viewer and object. This allows the viewer to obtain the same parts from different viewing angles. Third, again because the rule uses only intrinsic geometry, it can be applied equally to familiar and unfamiliar objects. Prior familiarity with an object is not a prerequisite to finding its parts under the minima rule. Fourth, some partitioning rule seems called for since the visual system can recognize objects even from only partial views, and even when components of the object, such as the limbs of an animal, change their spatial relationship.

The minima rule has an implication that we exploit in our psychophysical tests: namely, the parts it defines on a surface depend on the choice of figure and ground. This dependence obtains because the signs of the principal curvatures reverse when one reverses the choice of figure and ground (Hoffman and Richards 1984). Consequently what were negative minima of the principal curvatures become positive maxima, and vice versa, so that the part boundaries, which are identified with the negative minima, must change accordingly. Figure 1 provides an interesting confirmation of this prediction. The dashed contours are the part boundaries that, as predicted by the minima rule, one sees initially. They divide the surface into concentric rings of 'hills'. But if you turn the figure upside down, figure and ground reverse and, as predicted by the rule, the part boundaries appear to move. The dashed contours no longer appear to be part boundaries; instead they appear to lie in the middle of the new parts. This also shows, incidentally, that one's perception of the parts on this surface is not somehow coerced by the presence of the dashed contours.

This dependence of parts on the choice of figure and ground must be taken into account in any psychophysical tests of the minima rule. For if one wishes to determine if subjects perceive the parts predicted by the minima rule, then one must determine also

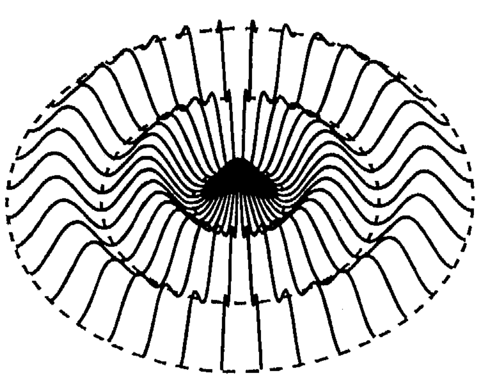

Figure 1. An example of the division of a surface into part boundaries according to the minima rule. Part boundaries on the cosine surface first appear at the positions marked by the dashed lines, but these apparent boundaries change when the figure is turned upside down and figure–ground organization is reversed.

their choices of figure and ground. If a subject can freely reverse the figure-ground assignment, and this is not taken into account, it might appear that positive maxima of curvature, in addition to negative minima of curvature, are used as part boundaries. Subjects are in fact quite adept at reversing figure and ground. Their facility at this has for several years frustrated our attempts to conduct clean tests of the minima rule.

But we have now found stimuli, and an experimental paradigm, which allow us greater control over subjects' choice of figure and ground. The stimuli which we ask subjects to partition are surfaces of revolution presented as structure-from-motion displays. Subjects seem much less likely spontaneously to reverse figure and ground on surfaces presented in motion. Figure-ground reversals are rare for these stimuli, relative to stationary views of the same surfaces. However, subjects can reverse figure and ground on a surface in motion if doing so facilitates their performance in a given task. We will see this in our second experiment.

Recall that a surface of revolution can be generated by taking a plane curve, the 'generating curve', and sweeping it around an axis, as shown in figure 2. Figure 2b shows a generating curve. In figure 2a the axis is to the left of the generating curve. Thus the generated surface of revolution has figure to the left of the generating curve and ground to the right. In figure 2c the opposite obtains—figure is to the right of the generating curve and ground to the left. Because of this figure-ground reversal between the two surfaces of revolution, negative minima of curvature for one are positive maxima of curvature for the other, and vice versa. Therefore the minima rule predicts that different parts should be seen on the two surfaces of revolution, even though both have the same generating contour. The same segment of the generating contour which is a 'part' for one surface should not, according to the minima rule, be a 'part' for the other. Presenting both surfaces in the course of an experiment, then, provides a good control. Each surface of revolution serves as a control for the other surface with the same generating contour. The role of negative minima can therefore be studied independently of other characteristics of the generating curve that might affect the partitioning of the surfaces into parts. This idea is central to the experiments that follow.

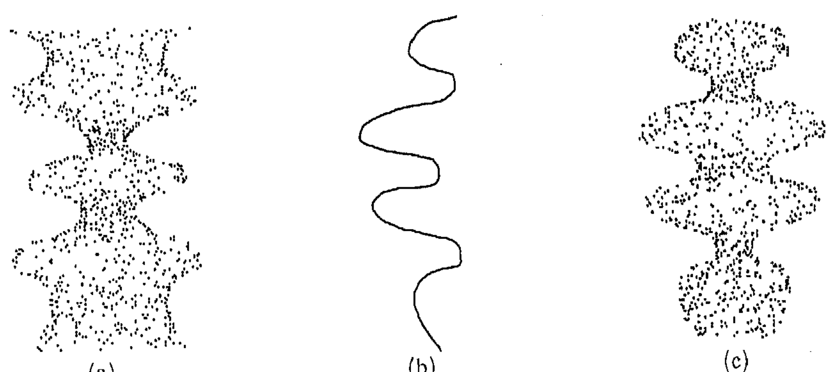

Figure 2. Two surfaces of revolution (a and c) generated from the same 'generating curve" (b). Note that the surface on the left is generated by putting the figure to the left of the curve and the surface on the right is generated by putting the figure to the right of the curve.

2 Experiment 1

In the first experiment we asessed the way subjects partition objects by requiring subjects to recognize parts of a previously seen object. Alternative parts were produced by dividing the objects either at negative minima of curvature (*minima* parts) or at positive maxima of curvature (*maxima* parts). (A third condition, in which the objects were divided at random locations, was used in pilot studies but did not provide useful additional information.) On each trial, subjects had to select from four alternatives a part that was from the object they had just seen. Two of the four alternative parts were from that object, one minima part and one maxima part. Thus there were two 'correct' parts. The other two parts were distractors.

2.1 *Method*

2.1.1 *Subjects.* The subjects were thirteen undergraduates who received extra credit in an introductory psychology or cognitive science class for participating.

2.1.2 *Stimuli.* The stimuli were white dots moving against a black background, simulating points on the surfaces of three-dimensional objects. The volume of each object was defined by a surface of revolution. The generating curves for each surface were computed by means of a b-spline algorithm with the control points for the algorithm selected in the following manner. Twelve general arrangements of control points were used, as shown in figure 3. These arrangements varied in number of control points and in whether the central point(s) in a group were above or below the horizontal line bisecting the top and bottom points. These general constraints were selected to provide sufficient variation among the objects so that the subjects could perform the recognition task. The control points were randomly positioned, within these general constraints. The curve was then repeated four times vertically. Finally, the horizontal coordinates were multiplied by a randomly-positioned sinusoidal function, so that the final curve would not consist of a precisely repeating pattern. Figure 2b provides an example of a curve generated in this manner.

Once a generating curve was selected, both that curve and its mirror image were used to generate surfaces. As noted earlier, the possibility that the curve-generating procedure favored perceptual segmentation of the parts at minima is controlled by the use of each curve and its mirror image to generate surfaces. Thus any point on the curve that defines a negative minimum in one stimulus defines a positive maximum in another, and vice versa.

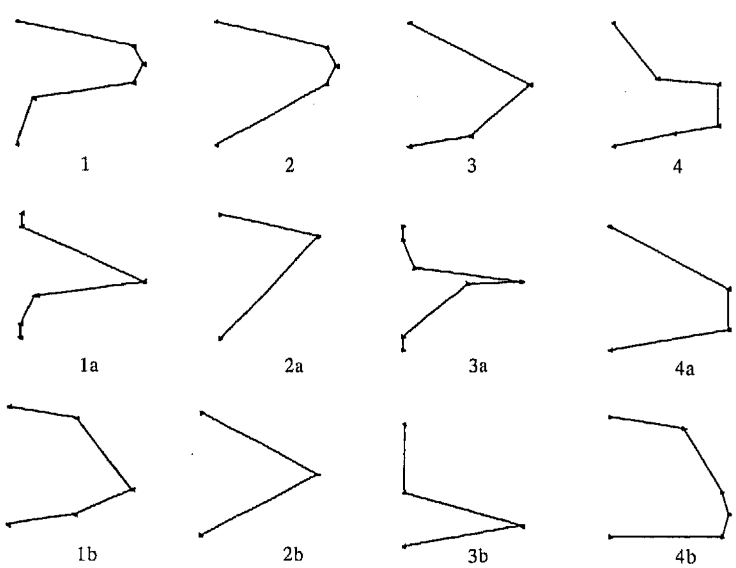

Figure 3. Control points used to generate distinct surfaces. Panels 1–4 were used to generate the eight experimental objects. The control points in panels 1a–4a and 1b–4b were used to generate alternative distractor parts for objects generated with the control points in panels 1–4.

A total of 2000 dots were used to define each surface. The dots were spaced at approximately equal vertical intervals and randomly positioned horizontally within the boundaries defined by the generating curve. The height of the projected surface on the CRT was 13.8 cm. The maximum width was 8.0 cm. The entire pattern was rotated continuously about a vertical axis and shown in orthographic projection. As dots on the back surface were not displayed, an average of 1000 dots were visible.

The part displays consisted of four parts arranged in a square, as shown in figure 4. The parts were generated by dividing objects either at negative minima of curvature or at positive maxima of curvature. To prevent recognition of parts on the basis of accidental features in the random-dot patterns, different selections of random dots were used in the full object and in the parts.

Four boundary curves were used to generate the objects for the experimental trials. By using each curve and its mirror image, we generated eight objects. These are shown in figure 5. Two different part displays were generated for each of the eight objects.

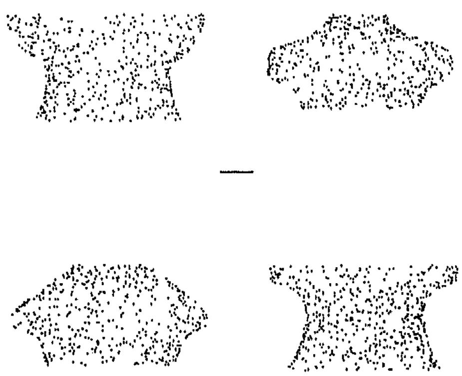

Figure 4. A part display, showing the initial position of the cursor.

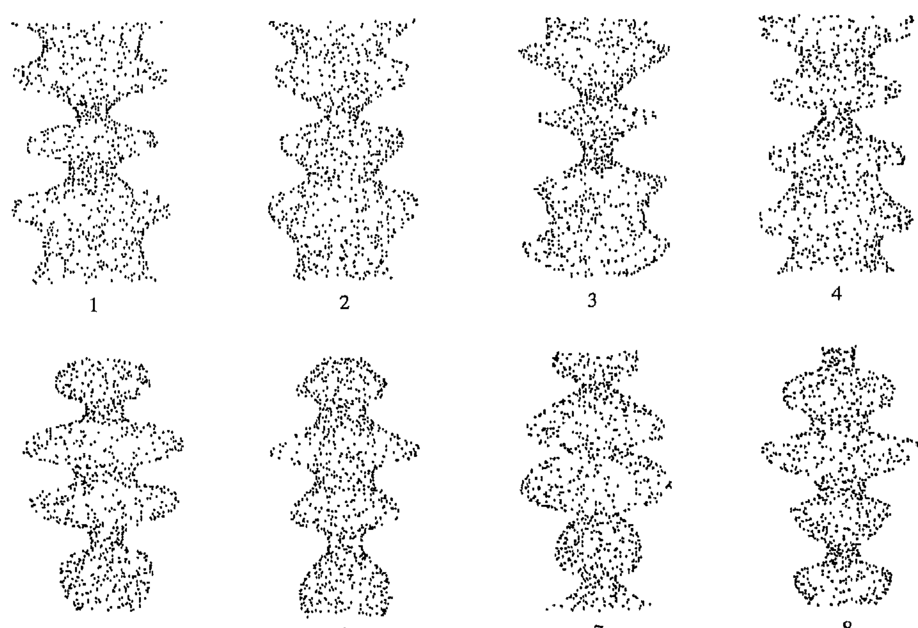

Figure 5. The eight objects used in experiments 1–3. The two objects in each column are based on the same boundary curve.

Each part display consisted of one minima part and one maxima part randomly selected from the object (the correct responses) and one minima part and one maxima part randomly selected from two other objects not used in the experiment (the distractor parts). The four alternative parts were randomly arranged in the part display. Four additional objects, based on two other boundary curves, were generated for practice trials.

2.1.3 *Apparatus.* The displays were presented on an IMI 455 vector graphics display scope with 4096 by 4096 resolution. Subjects viewed the display monocularly in a dimly lit room from a distance of 1.0 m through a tube arrangement that limited the field of view to an area within the borders of the display scope. The height of each object was 7.9 deg visual angle. The maximum width was 4.6 deg. During the part displays, a cursor was present on the screen that could be moved to positions below each of the four parts by means of a two-dimensional force joystick. A response button was located adjacent to the joystick.

2.1.4 *Procedure.* The subjects were told that they would see a rotating object which they were to study carefully. They were told that four parts would then be presented. Their task was to find a part that was from the object, to move the cursor under the part by means of the joystick, and to press the response button when they had done so. On each trial the object was presented for 20 s. After an interstimulus interval of 1 s, the parts were presented and remained visible until the subject pressed the button. The first four trials were practice trials with objects that were not used in the experimental trials. There were sixteen experimental trials—eight objects each with two different sets of alternative responses. The sixteen experimental trials were randomized separately for each subject.

2.2 *Results*

Of 208 responses (thirteen subjects responding on sixteen trials), 156 were correct and 52 were incorrect. Of the correct responses, 101 were minima selections and 55 were maxima selections. Of the incorrect responses, 30 were minima selections and 22 were maxima selections. An analysis of variance (ANOVA) for the two part types, minima and maxima, and eight objects was conducted with the frequency of correct responses as the dependent variable. The main effect of part type was significant ($F_{1,12} = 8.03$, $p < 0.05$). The main effect of object was not significant ($F_{7,84} = 1.98$) but there was a significant interaction of part type with object ($F_{7,84} = 2.77$, $p < 0.05$). For all objects except object 4, there were more correct choices of minima parts than maxima parts. Object 4 showed a reversal in this trend, accounting for the significant interaction.

These results indicate that when a task can be performed with either parts broken at negative minima or parts broken at positive maxima, there is a clear tendency for the parts broken at negative minima to be used. This tendency varied with subjects, with proportions of correct responses that were minima parts of, in descending order, 0.85, 0.82, 0.81, 0.77, 0.75, 0.73, 0.73, 0.62, 0.60, 0.57, 0.42, 0.38, and 0.33 for the thirteen subjects.

3 Experiment 2

In experiment 2 we examined performance on a task that could not be successfully accomplished on each trial by dividing the objects into parts at negative minima for later recognition. In this experiment, there was only one correct part among the four alternative parts presented on each trial. This was a minima part on half the trials and a maxima part on the other half. If subjects responded with a minima part on each trial they could be correct on at most 50% of the trials.

3.1 *Method*

The stimuli, apparatus, and procedure were identical to those used in experiment 1, except that only one part on each trial was from the object just presented. The correct part was a minima part on half of the trials and a maxima part on the remaining trials. Seventeen naive subjects were drawn from the same population as for experiment 1.

3.2 *Results*

The proportion of correct responses was 0.40, as compared to 0.75 in experiment 1. The probability of a correct response by chance was 0.25, compared to 0.50 in experiment 1. Accuracy can be compared directly in the two experiments by converting the proportions of correct responses to d' values (Swets 1964). The d' values for experiments 1 and 2 were 0.95 and 0.51 respectively.

An ANOVA on the frequency of correct responses for the two part types and eight objects showed a significant main effect for object ($F_{7,112} = 2.63$, $p < 0.05$). The proportions of correct responses for objects 1-8 were 0.38, 0.44, 0.38, 0.68, 0.29, 0.32, 0.41, and 0.26 respectively. A posteriori comparisons (Tukey's HSD test) showed the proportion of correct responses to object 4 to be significantly greater ($p < 0.05$) than to any other object. Additional significant differences were found between the proportion of correct responses for object 2 versus objects 5 and 8, and object 7 versus object 8.

The proportions of correct responses on the trials with correct minima parts and correct maxima parts were 0.46 and 0.34 respectively. This difference was not significant ($F_{1,16} = 3.95$) nor was the interaction of object with part type ($F_{7,112} = 1.92$).

The lack of a significant difference between accuracy on trials on which the correct part was a minima part and trials on which the correct part was a maxima part might be interpreted as a failure to support the hypothesis that objects are divided at negative minima for recognition. There is another interpretation, however. The recognition task in this experiment could be performed successfully by dividing objects at negative minima on only half the trials, and this may have forced subjects to use other strategies. Indeed, such strategies were reported during debriefing. Eight of the seventeen subjects reported deliberately trying to reverse figure and ground while viewing the objects. Figure-ground reversal would change positive maxima to negative minima and vice versa, and could allow subjects to perform the task by dividing parts at negative minima even when the experimental manipulation divided the parts at positive maxima. A figure-ground reversal strategy would have been successful on only half the trials, but may have blurred the distinction between negative minima and positive maxima trials.

4 Experiment 3

In experiment 3 we used a different measure of the division of objects into parts: subjects were asked to mark the part boundaries that they perceived. Our second aim in experiment 3 was to examine the possibility that perceived part boundaries occurred at narrow points on objects, rather than at negative minima of curvature. In the randomly-generated objects used in the first two experiments, the narrowest points on the objects tended to occur at negative minima of curvature. In experiment 3, an additional object was included that was not randomly generated but was designed specifically to separate these two variables. For this control object, the narrowest points did not occur at negative minima of curvature. Instead, they occurred along the arcs of circles, where there is no change in curvature. The control object is shown in figure 6.

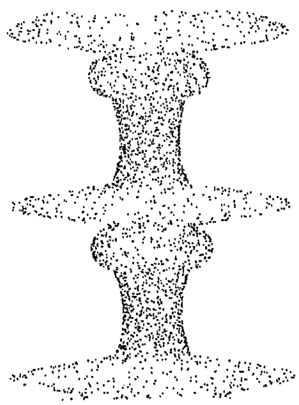

Figure 6. The control object used in experiment 3, with narrow points occurring along the arcs of circles.

4.1 *Method*

4.1.1 *Subjects.* The subjects were eight naive volunteers from the same population as in experiments 1 and 2.

4.1.2 *Stimuli.* The same eight objects were used as in experiments 1 and 2, but the size of the objects was increased to 17.3 cm in height and 10.0 cm in width (9.9 deg by 5.7 deg). (This was done to make those objects closer in size to the control object. The new height encompassed 2000 plotting positions vertically, and was the largest the original objects could be made without increasing the number of dots or allowing gaps in the display.) The control object contained arcs of circles which were placed between two negative minima of curvature such that the narrowest points on the objects were along these arcs. To generate sufficient arc length, this object was made somewhat larger than the randomly-generated objects (23.4 cm high by 17 cm wide or 13.3 deg by 9.7 deg). This new height encompassed 2700 plotting positions and 2700 dots were used in the control displays.

4.1.3 *Apparatus.* The apparatus was the same as in experiments 1 and 2.

4.1.4 *Procedure.* The subjects were instructed to move a horizontal red line up and down each object by means of the joystick and to press the response button to mark the boundaries between parts of the object. They were told that they should mark between three and ten parts on each object. These limits were introduced to provide general guidance in the task, in lieu of providing a definition for the term 'part'. The intent was to avoid extreme responses, such as considering each complete object to be a part or each dot to be a part.

When the subject pressed the response button a horizontal green line appeared at the marked position. The subject could remove a previous mark by repositioning the red line at that location and pressing the response button again. This deleted the green line at that position. The subject continued until satisfied that a green line was present at each part boundary, and indicated this by moving the red line below the object and pressing the button.

The experiment began with one practice trial, followed by the eight randomly-generated objects. This was followed by two presentations of the control object, once upright and once inverted.

4.2 *Results*

The responses for the eight randomly-generated objects were analyzed in several ways, all of which gave generally the same picture. In one type of analysis, the responses were categorized as minima boundary selections, maxima boundary selections, and other selections, based on whether the part was marked within 5% of the vertical height of the object from a negative minimum or a positive maximum. The proportions of minima, maxima, and other responses were 0.81, 0.10, and 0.09. The responses were also analyzed according to the mean vertical distance of each mark from the nearest negative minima and from the nearest positive maxima. These means were 3.8 and 15.0 mm (0.22 and 0.86 deg) respectively. (The means should be equal if the subjects were responding randomly.) As a more precise measure, distances were measured along the curve. The mean distances from the nearest negative minima and nearest positive maxima, respectively, measured this way were 4.5 and 20.2 mm (0.26 and 1.16 deg) respectively. A two-way ANOVA (object by minima versus maxima) showed a significant main effect for minima versus maxima ($F_{1,7} = 47.3$, $p < 0.01$). The main effect of object ($F_{7,49} = 2.3$) and the interaction ($F_{7,49} = 3.0$) were not significant.

We also measured the distances between part boundaries marked by the subjects and the nearest parabolic curve (curve of zero Gaussian curvature). The mean distance measured along the curve was 9.0 mm (0.52 deg). This is twice the mean distance to the nearest negative minima, indicating that subjects were not dividing objects into parts at loci of zero Gaussian curvature. (The role of parabolic curves in shape perception is discussed in Koenderink and van Doorn 1982.)

For the two control object trials, no subject marked a part boundary at a narrow point that was not at a negative minimum of curvature. The mean distance along the curve between a mark and a negative minimum was 8.5 mm (0.49 deg). The mean distance between a mark and a positive maximum was 63.9 mm (3.66 deg).

5 Discussion

We have presented evidence from two very different procedures which supports the hypothesis that subjects divide objects into parts perceptually at negative minima of curvature. In one procedure, subjects could perform a recognition task using either parts divided at negative minima or parts divided at positive maxima; 77% of the subjects responded more frequently with minima parts. In another procedure, subjects were asked to mark part boundaries, and 81% of the boundaries marked were at or near negative minima.

When a task could not be performed on all trials on the basis of parts divided at negative minima, as in experiment 2, there was a general degradation in performance. Subjects did not perform significantly better on trials on which there was a correct minima response, but appeared to be trying other strategies, including perceptually reversing figure and ground to change maxima into minima.

In randomly-generated surfaces of revolution, negative minima are likely to be at narrow points. Responses to the control stimuli in experiment 3, in which an object was presented that contained narrow points that were not at negative minima, demonstrated that subjects were choosing negative minima rather than narrow points as part boundaries.

However, the use of surfaces of revolution as stimuli may limit the generality of the results in other ways. For this special class of surfaces, negative minima of the principal curvatures along lines of curvature also happen to be extrema of other differential geometric quantities, eg Gaussian curvature and mean curvature. So the present study, by itself, does not rule out the possibility that another partitioning rule, stated in terms of these other quantities, better captures the partitioning strategies of subjects. Because the locations of negative minima and positive maxima depend upon figure-ground relationships, the present results provide evidence against any partitioning strategies which use quantities that are invariant under reversals of figure and ground, such as extrema of Gaussian curvature. However, other quantities may not be eliminated so easily. Further studies, with a greater diversity of shapes, could perhaps distinguish between these candidates and the minima rule.

Biederman (1987) has proposed a theory of human image understanding in which object recognition is based on parts that can be described by means of a modest set of generalized cone components. According to this theory, objects are segmented at regions of deep concavity. This segmentation hypothesis is consistent with the minima rule. The present results are consistent with Biederman's findings that contour deletions at regions of concavity in drawings of familiar objects impair recognition more than contour deletions at other locations.

In Beusmans et al (1987) the minima rule is augmented to deal with parts that are intrusions into a surface (in addition to the protrusions captured by the unaugmented minima rule). Further studies are needed to test the psychological validity of the proposed augmentations.

Acknowledgements. This research was supported by Office of Naval Research Contract N00014-88-K-0354 and National Science Foundation Grants BNS-8607530, BNS-8819565, and IRI-8700924.

References

Bennett B M, Hoffman D D, 1987 "Shape decompositions for visual shape recognition: The role of transversality" in *Image Understanding* Ed. W Richards (Norwood, NJ: Ablex) pp 215-256

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

Hoffman D D, 1983a *Representing Shapes for Visual Recognition* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA

Hoffman D D, 1983b "The interpretation of visual illusions" *Scientific American* **249**(6) 154-162

Hoffman D D, Richards W, 1984 "Parts of recognition" *Cognition* **18** 65-96

Koenderink J, Doorn A van, "The shape of smooth objects and the way contours end" *Perception* **11** 129-137

Marr D, 1982 *Vision* (San Francisco, CA: W H Freeman)

Swets J A, 1964 *Signal Detection and Recognition by Human Observers* (New York: John Wiley)