Parts and Wholes

Donald D. Hoffman

Vision starts with a shower of photons focused by the lens of the eye onto the retina. Each retina has roughly 125 million photoreceptors, and each photoreceptor signals roughly in proportion to the number of photons it catches (Alexiades & Khanal, 2007). One can think of the photoreceptor mosaic as reporting 125 million numbers, one for the quantity of photons caught at each photoreceptor. This massive array of numbers is the starting point of vision. It has no objects, shapes, parts, colors, textures or motions. The objects we see, and all their visual properties, must be constructed by the visual system out of this bewildering torrent of numbers. In particular, the parts and wholes of visual objects are not given, they must be constructed.

This constructive process has been shaped by natural selection to guide adaptive behavior in our niche. Aeons of random mutations, together with the culling of mutations that are less fit, have led to the constructive processes that create the parts and wholes of objects that we see today. This raises the question: Can we characterize, with mathematical precision, the process by which parts and wholes are constructed?

We construct the visual world to have three spatial dimensions, and we construct whole objects to be compact regions within that three-dimensional (3D) space, each object typically having a well-defined two-dimensional (2D) surface that bounds its compact region.

Most objects are not undifferentiated wholes, but instead are seen as having a natural decomposition into parts. Several mathematically precise rules can account for the parts we see in many objects. Here we explore the minima rule, the short-cut rule, and part salience rules. We begin with the minima rule (Hoffman & Richards, 1984).

Minima Rule: Divide 3D shapes into parts along concave cusps and along negative minima of the principal curvatures, along their associated lines of curvature.

An example of the part boundaries defined by this rule is given in Figure 1. In (a), when the three dots appear to lie on the faces of a single cube, then the lines that bound that cube are concave cusps, and therefore are part boundaries according to the minima rule. In (b) the circular dashed contours are negative minima of the principal curvatures along their associated lines of curvature, and are therefore part boundaries according to the minima rule.

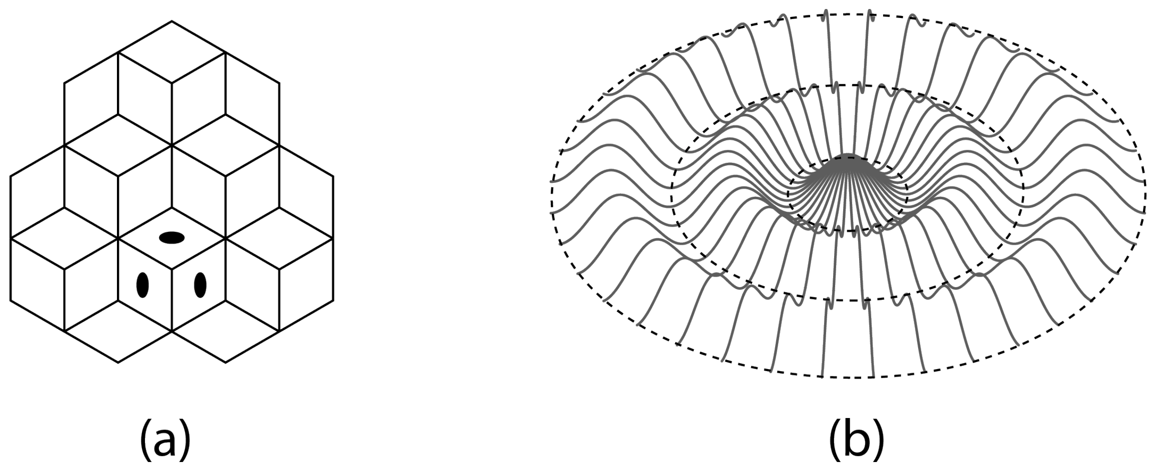

Figure 1. Illustration of part boundaries defined by the minima rule. (a) Part boundaries defined by concave cusps. (b) Part boundaries defined by negative minima of the principal curvatures along their associated lines of curvature.

To understand the statement of the minima rule, recall that most 2D surfaces embedded in 3D space are orientable, i.e., they admit two distinct fields of surface normals. An exception are unusual surfaces such as the Moebius strip or Klein bottle. For many orientable surfaces, human vision assigns one side of the surface to be “figure,” i.e., the object, and the other side to be “ground,” i.e., the space around the object (Rubin, 1915/1958). We adopt the convention that surface normals point towards the figure side of the surface.

With this convention, concave cusps point toward the figure side of the surface, and convex to its ground side. The concave cusps are part boundaries according to the minima rule. If human vision reverses figure and ground, then concave and convex cusps reverse roles and, according to the minima rule, new part boundaries should be seen. This can be checked in Figure 1(a). At first the three dots appear to be on single cube, and the concave cusps around that cube are its part boundaries. But after viewing the image for a while, figure and ground will appear to reverse and the three dots will suddenly appear to lie on three different cubes. These new cubes are defined by new concave cusps that used to be convex cusps when the three dots were seen to be on a single cube.

Recall that at every smooth point of a 2D surface embedded in 3D space, there are two "principal directions," one direction in which the surface curves most, and an orthogonal direction in which the surface curves least (do Carmo, 1976). Lines of greatest curvature are curves on the surface whose tangents always points in a direction of greatest curvature; lines of least curvature are curves whose tangents always points in a direction of least curvature. In Figure 1(b) the cosine-like curves radiated from the center are lines of curvature. If we adopt the convention that curvature is negative for regions of curves on the surface that are concave, and positive for regions that are convex, then the negative minima of curvature along lines of curvature are indicated by the dashed circular contours in Figure 1 (b). These are the part boundaries defined by the minima rule. One sees parts that look like hills divided from each other by these dashed circular contours, in accord with the minima rule. If one turns the page upside down, then figure and ground will appear to reverse, and the dashed circular contours will no longer be negative minima of curvature. Instead they are positive maxima of curvature, and appear to lie on top of a new set of hills that are defined by the new negative minima of curvature.

The minima rule for part boundaries on the surfaces of 3D objects leads naturally to a minima rule for part boundaries of silhouettes (Hoffman & Richards. 1984).

Minima rule for silhouettes. Divide each silhouette into parts at concave cusps and negative minima of its bounding contour.

This rule is illustrated by the well-known face-goblet illusion shown in Figure 2 (a). It sometimes appears to be a single goblet. But at other times, due to a figure-ground reversal, it appears to be two faces. In Figure 2 (b) are shown the two sets of part boundaries defined by the minima rule for silhouettes. If the bowl is seen as the figure, then the part boundaries are at the points indicated by the short line segments on the left side of the drawing. These part boundaries divide the silhouette into parts that correspond to the lip, bowl, stem and base of the goblet. If the faces are seen as the figure, then the regions of positive curvature become regions of negative curvature, and vice versa. Therefore the parts defined by the minima rule for silhouettes change. The new part boundaries are at the points indicated by the short line segments on the right side of the drawing. These part boundaries divide the silhouette into parts that correspond to the forehead, nose, lips and chin of the faces. Notice that the parts defined by the minima rule correspond to regions of the shape that have single word labels. For instance, if the goblet is figure, then one part defined by the minima rule is called the bowl. This same region, if the face is figure, would require a more complicated phrase to describe it, such as “lower half of the forehead and upper half of the nose.” This suggests that the parts defined by the minima rule are natural units of the visual representation of shape, and are among the units that are named by subsequent linguistic processing in the brain.

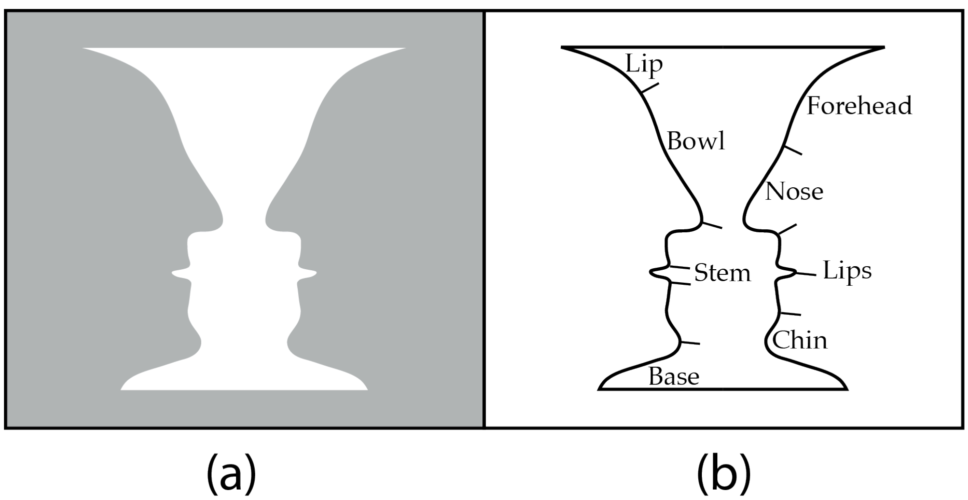

Figure 2. The minima rule for silhouettes, illustrated on the face goblet illusion. (a) The face-goblet illusion. (b) The two sets of part boundaries defined by the minima rule for silhouettes, one for the faces seen as figure and the other for the goblet seen as figure.

The minima rule defines part boundaries, but in some cases does not define a unique part cut. For instance, Figure 3 (a) shows a shape with the part boundaries defined by the minima rule for silhouettes. However these part boundaries could in principle be joined by a part cut either as shown in Figure 3 (b) or (c). Experiments by Singh, Seyranian and Hoffman (1999) demonstrate human vision prefers cuts that are shorter. This "short-cut rule" leads the visual system to prefer the cut in (c) over the cut in (b).

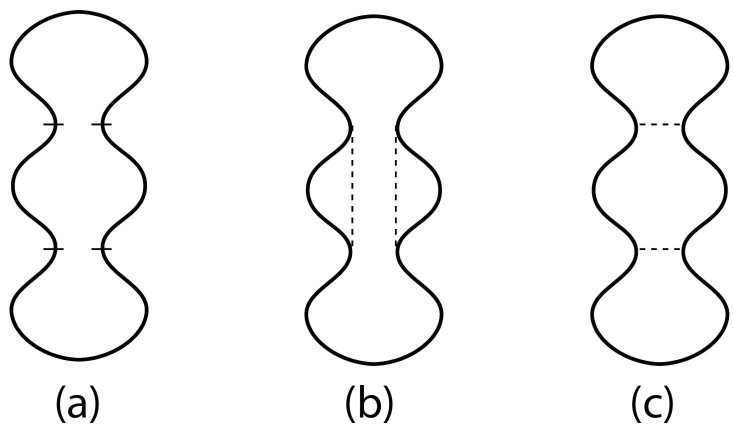

Figure 3. An illustration of the short-cut rule. (a) A silhouette with part boundaries defined by the minima rule for silhouettes. (b) Part cuts that are not preferred by the short-cut rule. (c) Part cuts that are preferred by the short-cut rule.

Part boundaries defined by the minima rule can vary in their perceptual salience (Hoffman & Singh, 1997). In Figure 4 (a) there are two part boundaries, one at the sharp concave cusp in the middle of the shape, and one at the concave cusp directly below it. The angle of the top boundary is more acute than that of the lower boundary, and this makes it perceptually more salient. In Figure 4 (b) there are again two part boundaries, one with high magnitude of curvature, and one directly below it with lower magnitude of curvature. The boundary with higher magnitude of curvature is perceptually more salient. Because curvature is not a scale invariant quantity, curvatures at boundaries must be normalized in a canonical fashion before their salience is compared (see Hoffman & Singh, 1997, for technical details).

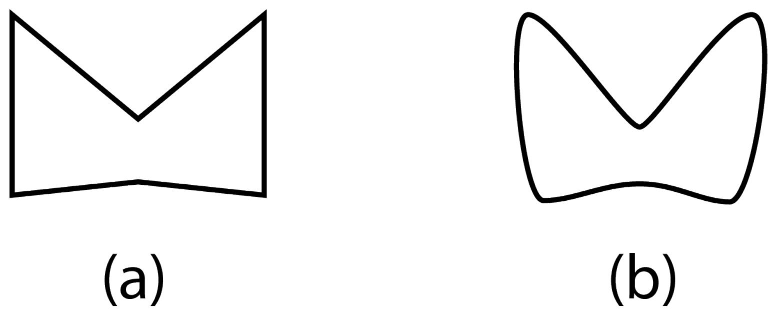

Figure 4. Salience of part boundaries. (a) The concave cusp on top has a more acute angle than the concave cusp directly below it, and is therefore more salient. (b) The negative minima boundary on top has greater (normalized) curvature than the one directly below it, and therefore is more salient.

The salience of a part boundary influences the perception of figure and ground (Hoffman & Singh, 1997). In Figure 5(a) the standard face-goblet illusion has been altered so that the part boundaries associated with the goblet are more salient. In Figure 5 (b) the part boundaries associated with the faces are more salient. When subjects were shown these images for 250 ms, they were about 3 times as likely to see Figure 5 (a) as a goblet than Figure 5 (b).

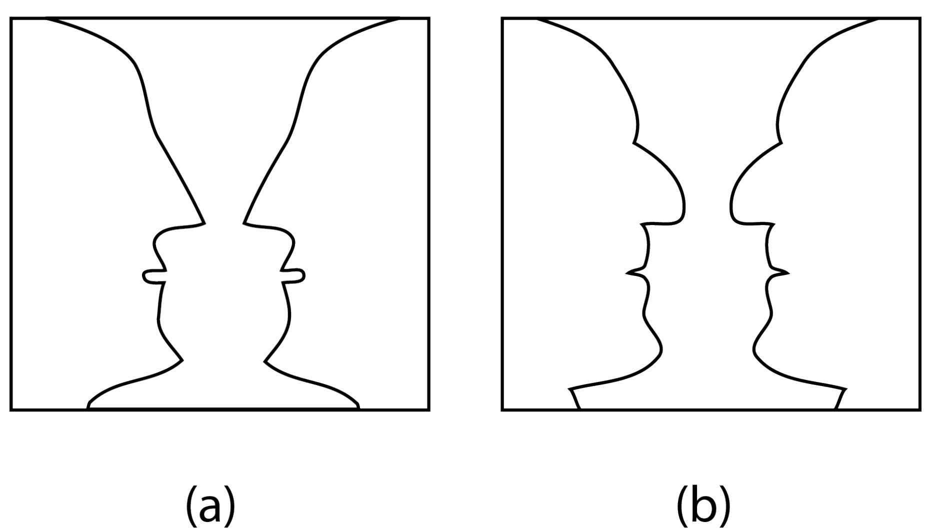

Figure 5. Salience of part boundaries affects perception of figure and ground. In (a) the boundaries associated with the goblet are more salient, and in (b) the boundaries associated with the faces are more salient. Human subjects are more likely to see (a) as being a goblet and (b) as being two faces.

The salience of a part is influenced by the salience of its boundaries (Hoffman & Singh, 1997). It is also influenced by the protrusion of the part—which is roughly how far the part sticks out—and by the volume of the part relative to the volume of the whole object. These informal descriptions can be given precise geometric definitions (Hoffman & Singh, 1997).

If parts defined by the minima rule correspond to natural units of the visual representation of shape, then one would expect that such parts would affect judgments of shape similarity. In Figure 6, which of the two half moons on the right looks most similar to the single half moon on the left? In a controlled experiment, subjects found the half moon on the bottom to be more similar (Hoffman 1983a, 1983b). Note that the bounding curve of the half moon on top is point-for-point identical to the half moon on the left. Indeed, they fit like puzzle pieces. The half moon on the bottom has been figure-ground reversed from the half moon on the left, and the position of two parts has been switched. If our visual judgments of shape similarity were based on a point-for-point comparison, then the half moon on the top would be judged more similar. However, human vision appears to judge shape similarity part by part, rather than point by point.

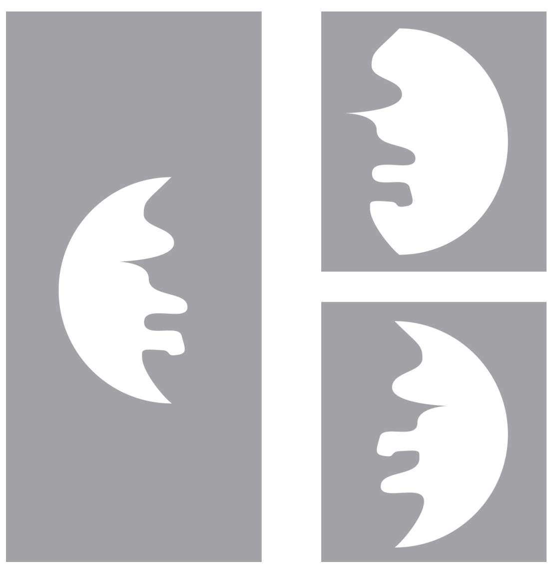

Figure 6. A demonstration that minima parts influence judgments of shape similarity.

Our perceptions of symmetry and repetition provide another demonstration that human vision analyzes shapes part by part rather than point by point. Mach (1885) noted that it is often easier to detect symmetry in a shape than to detect repetition. For instance, in Figure 7 the repetition in (a) is difficult to detect whereas the symmetry in (b) is easy to detect. What is surprising about this from a mathematical point of view is that symmetry involves repetition *plus* mirror reversal. So the number of mathematical transformations required to create symmetry is more than that required to create repetition.

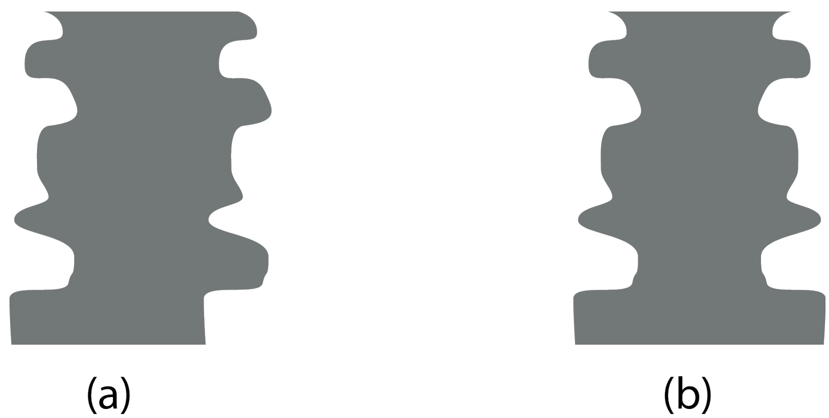

Figure 7. Detecting symmetry and repetition. In (a) the repetition is difficult to detect. In (b) the symmetry is easy to detect.

Baylis and Driver (1994, 1995a, 1995b) found that repetition can be made more easy to detect than symmetry if one reverses figure and ground. For instance, in Figure 8 the symmetry in (a) is not easily detected whereas in (b) the repetition is easily detected. The difference between Figure 8 and 7 is how figure and ground are seen across the curves. In Figure 7 the perception of figure and ground makes the two sides of the symmetric figure have the same minima-rule parts (mirror reversed), making symmetry easier to detect. In Figure 8 the perception of figure and ground makes the two sides of the repetition figure have the same minima-rule parts, making repetition easier to detect.

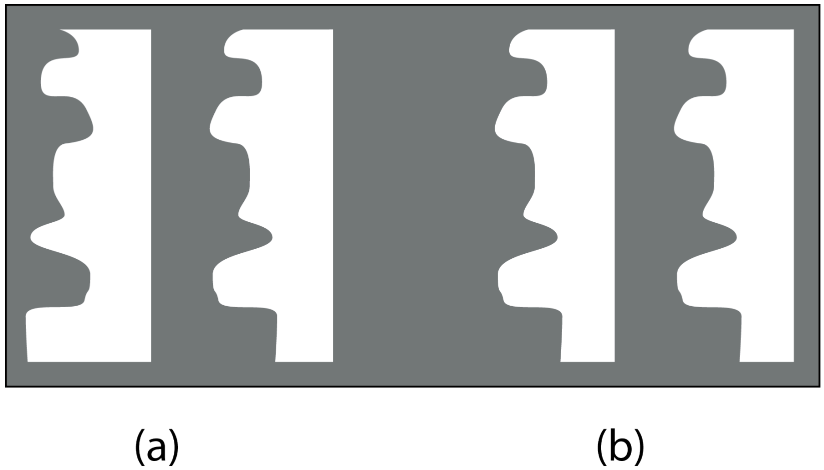

Figure 8. Symmetry and repetition revisited. In (a) the symmetry is not easily detected and in (b) the repetition is easily detected.

Part boundaries defined by the minima rule for silhouettes appear to be computed quickly and early in the stream of visual processing. Figure 9 illustrates this with a display similar to that used by Hulleman, te Winkel & Boselie (2000). In Figure 9 (a) the object with a part boundary pops out among convex distracters. In Figure 9 (b) the convex object does not pop out among objects with part boundaries. Visual search experiments by Xu and Singh (2002) indicate that parsing objects at negative minima of curvature occurs obligatorily.

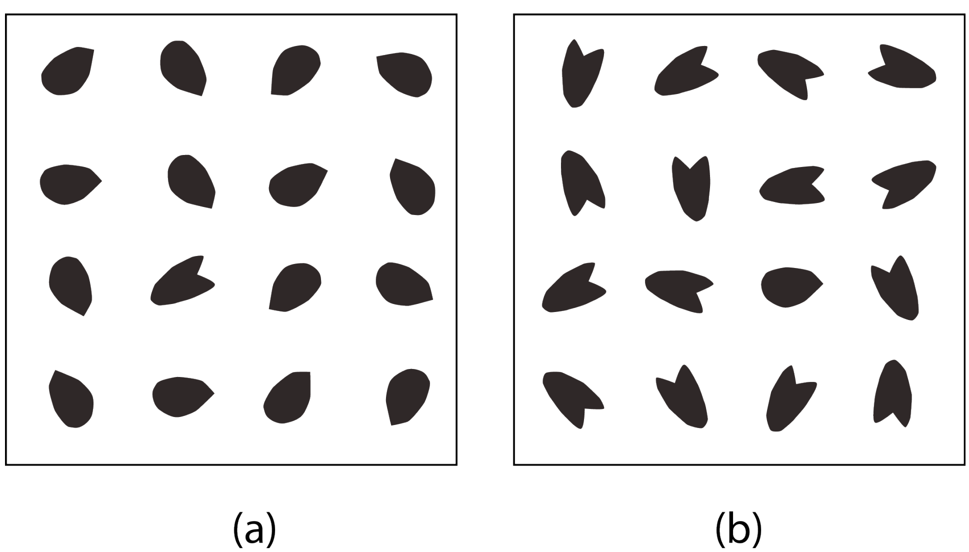

Figure 9. Demonstration similar to stimuli from Hulleman et al. (2000) showing that part boundaries are computed quickly and early in visual processing. In (a) the object with a part boundary pops out. In (b) the convex object does not pop out.

Parts defined by the minima rule affect the perception of transparency (Singh & Hoffman, 1998). Figure 10 (a) shows a standard example of transparency. In Figure 10 (b) minima part boundaries have been added at precisely the location where luminance changes, and this results in a greatly reduced perception of transparency. If the part boundary is not aligned with the luminance change, as in Figure 10 (c), then the perception of transparency returns. The loss of transparency in Figure 10 (b) appears to have two contributions, one due to parts and one due to genericity. Different parts can have different properties, and the visual system is therefore inclined to attribute the change in luminance in Figure 10 (b) to a difference in part properties rather than to transparency. If the cusps in Figure 10 (b) are made convex rather than concave, there is again a reduction in perceived transparency, but not as much as for the part boundaries. This indicates that the nongeneric alignment of cusps (concave or convex) also inclines the visual system to reject transparency.

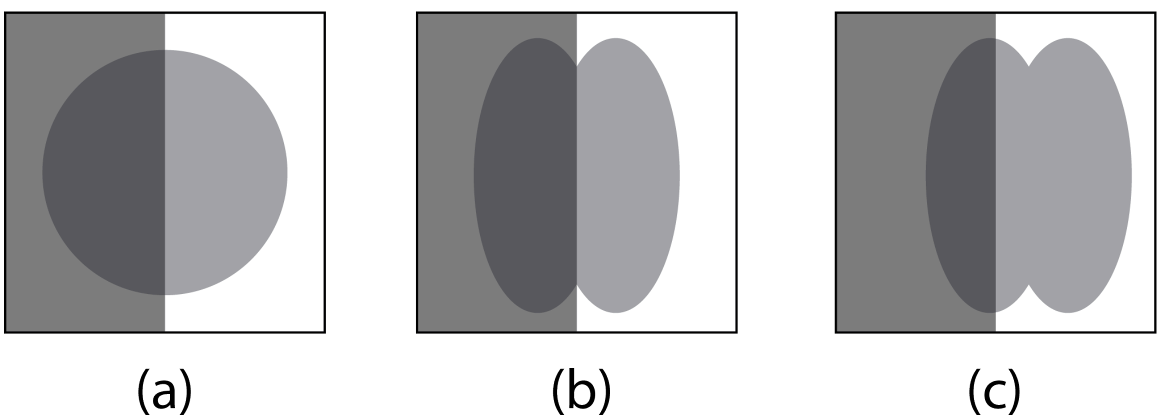

Figure 10. Minima part boundaries and perceived transparency. The transparency seen in (a) disappears in (b) where a part boundary aligns with the luminance change. Transparency is seen again in (c) where the part boundary is not aligned with the luminance change.

Parts defined by the minima rule affect recognition memory. Braunstein, Hoffman and Saidpour (1989) showed surfaces of revolution to human observers. After each surface was presented, observers were shown four separate parts and had to choose which part was in the surface just seen. Two of the parts were distracters, but two were actual parts of the surface, one defined by negative minima of curvature and one defined by positive maxima of curvature. Observers were twice as likely to choose the minima part rather than the maxima part, suggesting that these parts are natural units for recognition memory.

Minima parts affect visual attention. Singh and Scholl found that attention can shift more quickly within a part than across parts, and that greater salience of a part boundary makes the shift across parts even slower (Scholl, 2000; Singh and Scholl, 2000).

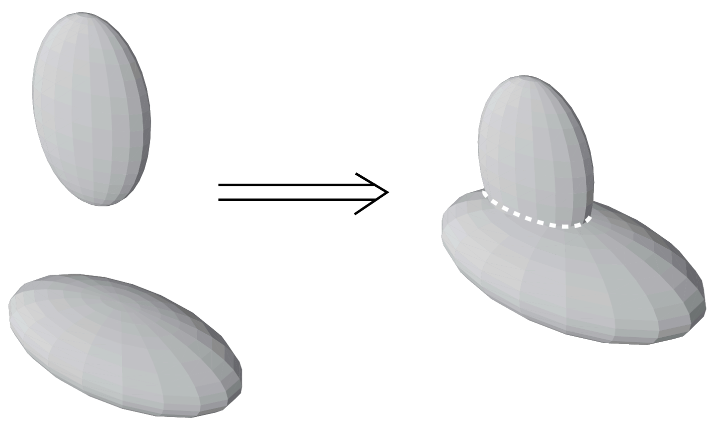

Figure 11. Transversal intersection and concave cusp.

Why does the visual system use concave cusps and negative minima of curvature to define part boundaries? Consider two separate objects, as shown on the left in Figure 11. If those two objects are joined to form a single object, as shown on the right in Figure 11, then the two original objects are good candidates for natural parts of the new object. In the generic case, the two objects intersect transversally (Guillemin & Pollack, 1974), leading to the concave cusp indicated by dashed contours. If this cusp is smoothed, it leaves a negative minimum of curvature. Thus genericity and transversality appear to be the mathematical principals underlying the minima rule.

It remains an open problem to precisely characterize all the qualitatively different kinds of parts that can result from applying the minima and short-cut rules to 3D shapes. Prior theories of shape perception have used various primitive shapes as parts, including polyhedra (Roberts, 1965; Waltz, 1975; Winston, 1975), generalized cones and cylinders (Binford, 1971; Brooks, 1981; Marr and Nishihara, 1978), superquadrics (Pentland, 1986) and geons (Biederman, 1987). Although there is substantial evidence for the importance of parts in the visual perception of shapes, there is no experimental evidence that any of these particular primitives play a role in the human visual perception of parts. The set of geons, for instance, is too restrictive. Geons can have truncated or pointed tips, but cannot have rounded tips, even though the distinction between rounded, pointed and truncated tips is a nonaccidental property (Binford, 1981; Lowe, 1987; Witkin & Tenenbaum, 1983) that survives projection and is easily seen by human vision. No geon can have a cross section that changes shape as it sweeps along the geon's axis, but such changing cross sections pose no problem for human vision.

In summary, human vision appears to decompose shapes into parts using the minima and short-cut rules, both for 3D shapes and for 2D silhouettes. The resulting parts appear to be computed quickly and early in visual processing, to affect our perceptions of shape similarity, to influence our recognition memory for shapes, to interact with our perceptions of symmetry and repetition, to alter our perceptions of transparency, and to influence our deployment of visual attention. Parts are a powerful component in the analysis and representation of visual shapes.

References

Alexiades, V., & Khanal, H. (2007). Multiphoton response of retinal rod photoreceptors. Sixth Mississippi State Conference on Differential Equations and Computational Simulations, *Electronic Journal of Differential Equations*, Conference 15, 1-9.

Baylis, G.C., & Driver, J. (1994). Parallel computation of symmetry but not repetition in single visual objects. *Visual Cognition, 1,* 377-400.

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

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

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

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

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

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

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

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

Guillemin, V. & Pollack, A. 1974. *Differential topology*. Englewood Cliffs, New Jersey: Prentice-Hall.

Hoffman, D.D. (1983a). *Representing shapes for visual recognition*. PhD Thesis, MIT.

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

Hoffman, D. D. (1998). *Visual Intelligence: How we create what we see.* New York: Norton.

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

Hoffman, D. D. & Singh, M. (1997). Salience of visual parts. *Cognition, 63*, 29-78.

Hulleman, J., te Winkel, W., & Boselie, F. (2000). Concavities as basic features in visual search: Evidence from search asymmetries. *Perception & Psychophysics*, 62, 162-174.

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

Mach, E. 1885/1959. *The analysis of sensations, and the relation of the physical to the psychical*. New York: Dover. (Translated by C.M. Williams.)

Marr, D., & Nishihara, H.K. (1978). Representation and recognition of three-dimensional shapes. *Proceedings of the Royal Society of London*, *Series B*, 200, 269–294.

Pentland, A.P. (1986). Perceptual organization and the representation of natural form. *Artificial Intelligence*, 28, 293–331.

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

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

Scholl, B. (2001). Objects and attention: The state of the art. *Cognition*, 80, 1-46.

Singh, M. & Hoffman, D. D. (1998). Part boundaries alter the perception of transparency. *Psychological Science, 9*, 370-378.

Singh, M, & Scholl, B. (2000). Using attentional cueing to explore part structure. Poster presented at the *Annual Symposium of Object Perception and Memory*, New Orleans, Louisiana (November 2000).

Singh, M., Seyraninan, G. D., & Hoffman (1999). Parsing silhouettes: The short-cut rule. *Perception and Psychophysics, 61,* 636-660.

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

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

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

Press.

Xu, Y. & Singh, M. (2002). Early computation of part structure: Evidence from visual search. *Perception & Psychophysics*, 64, 1039-1054.