Mereology of Visual Form

Donald D. Hoffman[1]

[1] Department of Cognitive Science, University of California, Irvine 92697 USA

Abstract. Visual forms come in countless varieties, from the simplicity of a sphere, to the geometric complexity of a face, to the fractal complexity of a rugged coast. These varieties have been studied with mathematical tools such as topology, differential geometry and fractal geometry. They have also been examined, largely in the last three decades, in terms of mereology, the study of part-whole relationships. The result is a fascinating body of theoretical and empirical results. In this paper I review these results, and describe a new development that applies them to the problem of learning names for visual forms and their parts.

1 Introduction

From the anatomy and physiology of the retina, we know that the processes of vision begin from a source that is at once rich and impoverished: photon quantum catches at 5 million cones and 120 million rods in each eye [1]. This source is rich in the sheer number of receptors involved, the dynamic range of lighting over which they operate, and the volume of data they can collect over time. This source is impoverished in its language of description. The language can only state how many quanta are caught and by what receptors. It can say nothing about color, texture, shading, motion, depth or objects, all of which are essential to our survival. For this reason we devote precious biological resources—hundreds of millions of neurons in the retina and tens of billions of neurons in the cerebral cortex—to construct richer languages and more adaptive descriptions of the visual world.

A key criterion for these more adaptive descriptions is that they allow us to predict, with economy of effort, future events that can affect our survival. We construct a world of objects and their actions, because carving the world this way lets us quickly learn important predictions. Running toward a rabbit leads to predictably different results than running toward a lion. These are important object-specific properties that cannot be learned in the language of quantum catches.

We carve the world more finely still, dividing objects themselves into parts. Parts aid in the recognition of objects. Parts also allow more refined predictions: If, for instance, one is fighting a conspecific it might be critical to attend to certain parts, such as arms or legs or jaws, and relatively safe to ignore other parts such as ears. Moreover, some parts of shapes are better remembered than others [2–4]. The centrality of parts to human vision can be seen in the following six figures, each of which can be explained, as we will see shortly, by three rules for computing the parts of objects.

In Figure 1 you probably see hill-shaped parts with dashed lines in the valleys between them. But if you turn the figure upside down, you will see a new set of hills, and now the dashed lines lie on top of the new hills [5].

In Figure 2, which of the two half moons on the right looks most similar to the half moon on the left? In controlled experiments almost all subjects say that the bottom looks more similar to the half moon on the left —even though the top half moon, not the bottom, has the same bounding curve as the half moon on the left [6].

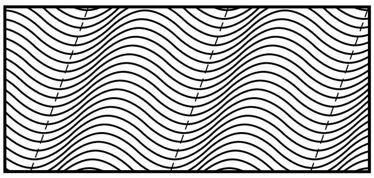


Figure 1.

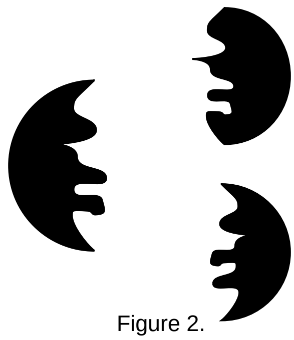
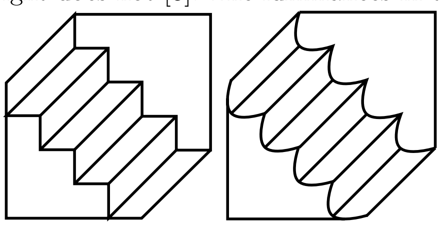


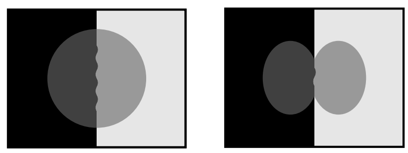
Figure 2.

In Figure 3, most observers say the staircase on the right looks upside down, whereas the one on the left can be seen either as right side up or as upside down [7].

In Figure 4, the display on the left looks transparent, but the one on the right does not [8]. The luminances in the two cases are the same.



(a) Figure 3. (b)

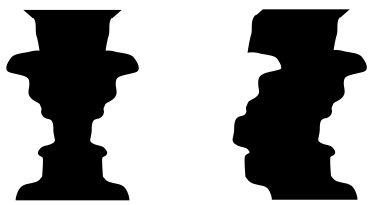


(a) (b)

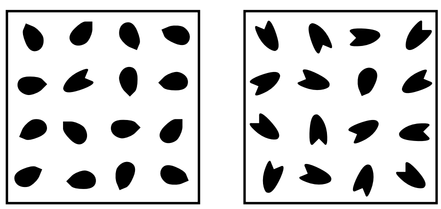
Figure 4.

In Figure 5, the symmetry of the shape on the left is easier to detect than the repetition of the shape on the right [9–12].

In Figure 6, a heart shape pops out among a set of popcorn-shaped distractors, as shown on the left, but not vice versa, as shown on the right [13,14].



(a) Figure 5. (b)



(a) Figure 6. (b)

Figure 6 suggests that we begin to construct parts, or at least boundaries between parts, early in the stream of visual processing. Empirical evidence in support of this suggestion comes from experiments using visual search [13,14] with stimuli like Figure 6. In one experiment conducted by Hulleman, te Winckel and Boselie [13], subjects searched for a heart-shaped target amidst popcorn-shaped distractors, and vice versa. The heart had a concave (i.e., inward pointing) cusp which divided it into two parts, a left and right half. The popcorn had a convex (outward pointing) cusp and no obvious parts. Its convex cusp was chosen to have the same angle as the concave cusp of the heart. The data indicate that subjects search in parallel for the heart targets, but search serially for the popcorn targets. It appears that parts are important enough to the visual system that it devotes sufficient resources to search in parallel for part boundaries. This early and parallel construction of part boundaries explains why parts affect the perception of visual form in the wide variety of ways illustrated in Figures 1 through 6.

2 The Minima Rule

Recent experiments suggest that human vision divides shapes into parts by the coordinated application of three geometric rules: the *minima rule*, the *short-cut rule*, and the *part salience rule*. The rules are as follows:

- **Minima Rule (for 3D Shapes):** *All concave creases and negative minima of the principal curvatures (along their associated lines of curvature) form boundaries between parts* [5,7].
- **Minima Rule (for 2D Silhouettes):** *For any silhouette, all concave cusps and negative minima of curvature of the silhouette's bounding curve are boundaries between parts* [5,7].
- **Short-cut rule:** *Divide silhouettes into parts using the shortest possible cuts. A cut is (1) a straight line which (2) crosses an axis of local symmetry, (3) joins two points on the outline of a silhouette, such that (4) at least one of the two points has negative curvature. Divide 3D shapes into parts using minimal surfaces* [15].
- **Salience rule:** *The salience of a part increases as its protrusion, relative area, and strength of part boundaries increases.* [7].

Together these rules can explain each visual effect illustrated in Figures 1–6. In Figure 1, the straight lines in the valleys are negative minima of the principal curvatures and therefore, according to the minima rule, they are part boundaries. When you turn the illustration upside down this reverses figure and ground, so that negative minima and positive maxima of curvature reverse places. Therefore, according to the minima rule, you should see new part boundaries and new parts [5]. A quick check of the illustration will confirm this prediction.

In Figure 2, the half moon on the top right has the same contour as the half moon on the left. Yet most observers pick the half moon on the bottom right as more similar to the half moon on the left, even though its contour is mirror reversed and two of the minima parts have switched positions. This is explained by the minima rule because this rule carves the bottom half moon into parts at the same points as the half moon on the left, whereas it carves the top half moon into different parts [6]. Apparently shape similarity is computed part by part, not point by point.

In Figure 3 the staircase on the right looks inverted, because the parts defined by the minima rule for the inverted interpretation have more salient part boundaries (sharper cusps) than for the upright interpretation. Other things being equal, human vision prefers that choice of figure and ground which leads to the most salient minima-rule parts [7].

In Figure 4 we see transparency on the left but not on the right, even though all the luminances are identical. The reason is that on the right there are minima-rule part boundaries aligned with the luminance boundaries, so we interpret the different grays as different colors of different parts rather than as effects of transparency [8].

In Figure 5 we detect the symmetry on the left more easily than the repetition on the right, an effect noted long ago by Mach [12]. The minima rule explains this because in the symmetric shape the two sides have the same parts, whereas in the repetition shape the two sides have different parts [9]. Again it appears that shapes are compared part by part, not point by point.

In Figure 6 the heart pops out on the left, but the popcorn does not pop out on the right, even though the two have cusps with identical angles [13]. The minima rule explains this because the concave cusp in the heart is a part boundary whereas the convex cusp on the popcorn is not. Minima part boundaries are computed early in the flow of visual processing since parts are critical to the visual representation of shape.

The minima rule makes precise a proposal by Marr and Nishihara [16] that human vision divides shapes into parts at "deep concavities". The minima rule has strong ecological grounding in the principle of transversality from the field of differential topology. This principle guarantees that, except for a set of cases whose total measure is zero, minima part boundaries are formed whenever two separate shapes intersect to form a composite object, or whenever one shape grows out of another [5,7].

3 Other Part Rules

Given the central role of parts in object perception, it is no surprise that several theories of these parts have been proposed. One class of theories claims that human vision uses certain basic shapes as its definition of parts. Proponents of basic-shape theories have studied many alternatives: polyhedra [17–19], generalized cones and cylinders [16,20], geons [21,22], and superquadrics [23]. Of these, the geon theory of Biederman is currently most influential. Geons are a special class of generalized cylinders, and come in 24 varieties. The set of geons is derived from four nonaccidental properties [21,24,25]. These properties are whether (a) the cross section is straight or curved; (b) the cross section remains constant, expands, or expands and contracts; (c) the cross section is symmetrical or asymmetrical; and (d) the axis is straight or curved. These properties are intended to make recognition of geons viewpoint invariant, although some experiments suggests that geon recognition may nevertheless depend on viewpoint [26].

A second class of theories claims that human vision defines parts not by basic shapes but by rules which specify the boundaries between one part and its neighbors. The minima rule is one such theory. Another is the theory of "limbs" and "necks" developed by Siddiqi and Kimia [27], building on their earlier work [28]. They define a limb as "a part-line going through a pair of negative curvature minima with co-circular boundary tangents on (at least) one side of the part-line" ([27], p. 243). Their "part-line" is what I call a "part cut." Two tangents are "co-circular" if and only if they are both tangent to the same circle ([29], p. 829). Siddiqi and Kimia define a neck as "a part-line which is also a local minimum of the diameter of an inscribed circle" ([27], p. 243).

4 Naming Parts

Figure 7a shows a *peen.* After looking at the figure, you know that a peen is the part of a hammer that is shaded gray in Figure 7c. You also are sure that a peen is not the part shaded gray in Figure 7b. This exercise in ostensive definition is easy for us; we discover the meaning of *peen* quickly and without conscious effort. Yet in several respects our performance is striking. When we view the hammer in Figure 7a and guess the meaning of peen, we discard countless possibilities since, as is hinted by Figure 7b, there are countless ways to partition the hammer or any other shape. But despite these countless possibilities we all pick the same part. Deductive logic alone does not compel us to choose a unique part from the figure. Our choice must be constrained by rules in addition to the rules of deductive logic. In principle many different rules can yield a unique choice. But since we all pick the *same* part, it is likely we all use the same rules. What rules then do we use that make us all pick the same part when someone points and names?

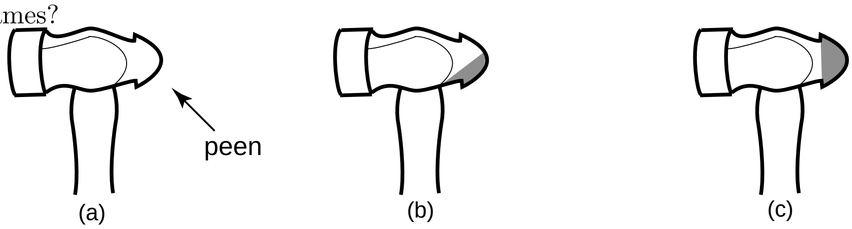


Figure 7.

We propose the following hypothesis:

> **Minima-Part Bias.** When human subjects learn, by ostensive definition, names for parts of objects, they are biased to select parts defined by the minima and short-cut rules.

The motivation for this hypothesis is that the units of representation computed by the visual system and delivered to higher cognitive processes are natural candidates to be named by the language system. I propose that parts defined by the minima rule are such units. As an example, in Figure 8a is a curve, which is ambiguous as to which side is figure and which is ground. If we take the ground to be on the left, as in Figure 8b, then the minima rule gives us the part boundaries that are indicated by the short line segments. As you can see in Figure 8b, each part boundary lies in a region of the curve that is concave with respect to the figure (or, equivalently, convex with respect to the ground), and each part boundary passes through the point with highest magnitude of curvature within its region. If we take the ground to be on the right, as in Figure 8c, then the minima rule gives us a completely different set of part boundaries, indicated by the new set of short line segments. The reason is that switching figure and ground also switches what is convex and concave, and the minima rule says to use only the concave cusps and concave minima of curvature as part boundaries.

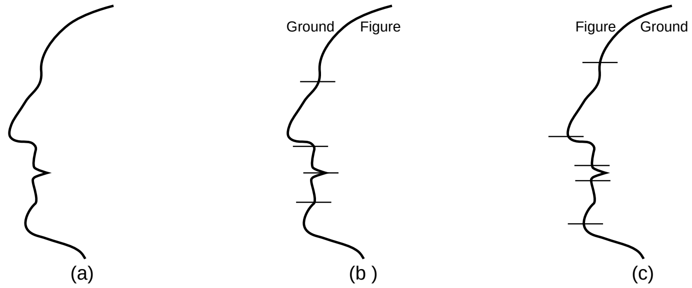


Figure 8.

We can use the curve in Figure 8a to create the well-known face goblet illusion, as shown in Figure 9a.

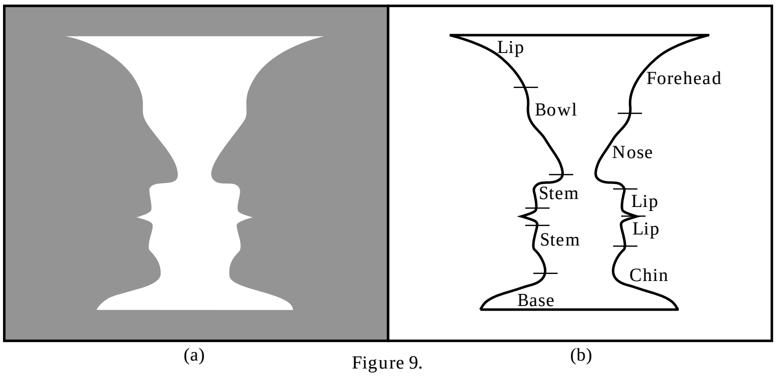


Figure 9.

One can see this either as a goblet in the middle or two faces on the sides. If we see the faces as figure, then the minima rule gives the part boundaries shown on the right side of Figure 9b. These part boundaries divide the face, from top to bottom, into a forehead, nose, upper lip, lower lip, and chin, as labeled in the figure. If instead we see the goblet as figure, then the minima rule gives the part boundaries shown on the left side of Figure 9b. These part boundaries divide the goblet, from top to bottom, into a lip, bowl, stem (with three parts), and base. Note that the parts defined by the minima rule are the parts we name. Other parts are not named. For instance, take the bowl part of the goblet and ask what we would call this same part on the face (where it is not a part defined by the minima rule). We would say it is "the lower part of the forehead and the upper part of the nose." This is not a name but a complex description of an unnamed part of the face. Thus the parts that we name on the face and the goblet are precisely the parts derived from the minima rule.

The outline shape of an object can play an important role in its recognition across depth rotations [30]. To test the minima-part bias with such outline shapes, Rodriguez, Nilson, Singh and Hoffman generated five random silhouettes each having five minima parts [31]. On each trial the observer saw one silhouette with an arrow pointing to it. Orientations of the silhouettes were changed from trial to trial. The observer was instructed that the arrow pointed to "a dax" on the shape. The syntax of count nouns was used in these instructions to direct the observer's attention to shape rather than substance interpretations of "a dax" [32–35]. On each trial the observer was given three choices, displayed in random order, for the dax: (1) a minima part, (2) a maxima part, and (3) a convex part cut at inflections. The arrow was placed so that it pointed towards the inflection, in order to minimize possible biases due to the position of the arrow. They found that minima parts were chosen as the dax about 75% of the time, far more frequently than the maxima or inflections.

To further control for possible biases of the arrow position, a second experiment eliminated the arrow. Instead observers were instructed that the shape had a "dax" near its top. On each trial the silhouette was rotated so that the 3 parts of interest (minima, maxima, inflections) were near the top, but never precisely vertical. They again found that minima parts were chosen about 75% of the time.

The minima-part bias makes a striking prediction. Suppose a subject can see a shape undergo reversals of figure and ground. Then the parts of that shape that the subject will name should also change each time figure and ground reverse. The reason is that the minima rule defines part boundaries only at concave regions of an object. Thus when figure and ground reverse so also do concave and convex, so that subjects should see a new set of part boundaries. This prediction was illustrated in Figures 8 and 9. Rodriquez et al. tested this prediction of the minima-part bias, using a simple method to induce a figure-ground reversal: global reversal of contrast together with enclosure [31]. Subjects viewed a shape, with an arrow pointing toward a "dax" on the shape, and picked the part that looked most natural to be the "dax". On a different trial they saw precisely the same shape, with the arrow pointing in exactly the same way, but with reversed contrast. This reversed contrast induced subjects to reverse figure and ground. In this case the minima-part bias predicted that they would pick a different part, one with the new minima of curvature for its part boundaries, as the most natural "dax". The shapes used were five random curves each having three or four minima parts. Observers chose between minima, maxima, and inflection options as before. The results were again as predicted by the minima-rule bias.

Rodriguez et al. also devised at test in which the predictions of the geon theory differ dramatically from those of the minima-rule bias [31]. They used a two-alternative forced-choice paradigm in which subjects saw two objects side by side and were told that one of the objects had a dax on it. Subjects had to choose which of the two objects had the dax on it. For most pairs of objects that were shown to the subjects, the minima rule and geon theory predicted opposite choices. Each of the objects was composed of two shapes. The first, the base shape, was either an elongated box or a cylinder. The second was one of twelve shapes, six of which were geons and six of which were nongeons. The geons were (1) a curved box or cylinder, (2) a tapered box or cylinder, and (3) a curved and tapered box or cylinder. The six nongeons were created from the geons by smoothly changing the cross section from square to circular or vice versa as it sweeps along the axis. Each of these twelve shapes was attached to it base shape in one of two ways: (1) with a minima part boundary at the attachment or (2) with no minima boundary at the attachment. This led to a total of 18 composite objects. Thus there were four types of objects, defined by the type and attachment of the second shape. These were (1) geons with minima, (2) geons without minima, (3) nongeons with minima, and (4) nongeons without minima. We can label these, respectively, +G+M, +G-M, -G+M, and -G-M. In a two-alternative forced-choice paradigm there are $\binom{4}{2} = 6$ ways of presenting pairs of these objects. These ways are listed below, together with the predictions of the geon theory and the minima rule as to which of the two objects is most likely to be chosen as the one having the dax.

CASE	OBJECT 1	OBJECT 2	GEON	MINIMA
1	+G+M	+G-M	same	object 1
2	+G+M	-G+M	object 1	same
3	+G+M	-G-M	object 1	object 1
4	+G-M	-G+M	object 1	object 2
5	+G-M	-G-M	object 1	same
6	-G+M	-G-M	same	object 1

As you can see from this table, in five of the six cases the minima rule and the geon theory make different predictions. Only in the third case do their predictions agree. Rodriquez et al. used all cases in a two-alternative forced-choice paradigm to test which theory correctly predicts observers' choices. They found that where the predictions disagreed, each subject chose overwhelmingly in accord with the minima bias and not in accord with the geon theory.

In a final experiment, Rodriguez et al. found that subjects also use the minima-part bias in generalizing the names of parts. On each trial they showed a subject an object with a single attached part, and told the subject that the object had a "dax" on it. Then they showed the subject that same object, but with the part transformed either by translation, scaling, or both translation and scaling. On some trials a minima part was transformed, on others a maxima part, and on others an inflection part. Subjects had to decide if this new object also had a "dax" on it. Subjects did generalize the part name to transformed minima parts, but not to transformed maxima or inflection parts. This indicates that the minima-rule bias guides both the initial attachment of names to parts, and the generalization of part names to new parts.

4 Conclusion

Visual form is not available at the retina, but must be constructed by tens of billions of neurons in the visual system. The description of visual form requires the visual system to carve the visual world into objects, and to carve these objects even more finely into parts. Human vision apparently does this by constructing minima-rule part boundaries early in the course of visual processing. These boundaries, together with geometric rules for constructing part cuts which use these boundaries, leads to an articulation of objects into parts. The potential shapes of these parts are countless, and are not limited to a preordained set such as geons or generalized cylinders. These parts are among the fundamental units of visual description that are delivered to higher cognitive processes, including language. As a result, in language acquisition humans employ the minima-part bias when learning the names of parts by ostensive definition. The minima-part bias has so far only been tested in adults. It will be of interest to see if young children also employ this bias when learning names for parts.

References

1. Tyler, C.W.: Analysis of Human Receptor Density. In: Lakshminarayanan, V. (ed.): Basic and Clinical Applications of Vision Science. Kluwer Academic, Norwell, Massachusetts (1997) 63–71
2. Bower, G.H., Glass, A.L.: Structural Units and the Redintegrative Power of Picture Fragments. J. Exper. Psych. **2** (1976) 456–466
3. Braunstein, M.L., Hoffman, D.D., Saidpour, A.: Parts of Visual Objects: An Experimental Test of the Minima Rule. Perception **18** (1989) 817–826
4. Palmer, S.E.: Hierarchical Structure in Perceptual Representation. Cognit. Psychol. **9** (1977) 441–474
5. Hoffman, D.D., Richards, W.A.: Parts of Recognition. Cognition **18** (1984) 65–96
6. Hoffman, D.D.: The Interpretation of Visual Illusions. Sci. Am. **249** 6 (1983) 154–162
7. Hoffman, D.D., Singh, M.: Salience of Visual Parts. Cognition **63** (1997) 29–78
8. Singh, M., Hoffman, D.D.: Part Boundaries Alter the Perception of Transparency. Psych. Sci. **9** (1998) 370–378
9. Baylis, G.C., Driver, J.: Obligatory Edge Assignment in Vision: The Role of Figure and Part Segmentation in Symmetry Detection. J. Exp. Psych.: Hum. Percept. Perform. **21** (1995) 1323–1342
10. Bruce, V., Morgan, M.J.: Violations of Symmetry and Repetitions in Visual Patterns. Perception **4** (1975) 239–249
11. Hoffman, D.D.: Visual Intelligence: How We Create What We See. W.W. Norton, New York (1998)
12. Mach, E.: The Analysis of Sensations, and the Relation of the Physical to the Psychical. Dover, New York (1885/1959) (Translated by Williams, C.M.)
13. Hulleman, J., te Winkel, W., Boselie, F.: Concavities as Basic Features in Visual Search: Evidence from Search Asymmetries. Percept. Psychophys. **62** (2000) 162–174
14. Wolfe, J.M., Bennett, S.C.: Preattentive Object Files: Shapeless Bundles of Basic Features. Vis. Res. **37** (1997) 24–43

15. Singh, M., Seyranian, G., Hoffman, D.D.: Parsing Silhouettes: The Short-Cut Rule. Percept. Psychophys. **61** (1999) 636–660
16. Marr, D., Nishihara, H.K.: Representation and Recognition of Three-Dimensional Shapes. Proc. Roy. Soc. Lond. B, **200** (1978) 269–294
17. Roberts, L.G.: Machine Perception of Three-Dimensional Solids. In: Tippett, J.T. et al. (eds.): Optical and Electrooptical Information Processing. MIT Press, Cambridge, Massachusetts (1965) 211–277
18. Waltz, D.: Generating Semantic Descriptions from Drawings of Scenes with Shadows. In: Winston, P. (ed.): The Psychology of Computer Vision. McGraw-Hill, New York (1975) 19–91
19. Winston, P.A.: Learning Structural Descriptions from Examples. In: Winston, P. (ed.): The Psychology of Computer Vision. McGraw-Hill, New York (1975) 157–209
20. Brooks, R.A.: Symbolic Reasoning among 3-D Models and 2-D Images. Artific. Intell. **17** (1981) 205–244
21. Biederman, I.: Recognition-By-Components: A Theory of Human Image Understanding. Psych. Rev. **94** (1987) 115–147
22. Biederman, I.: Higher Level Vision. In: Osherson, D., Kosslyn, S., Hollerbach, J. (eds.) An Invitation to Cognitive Science, Volume 2. MIT Press, Cambridge, Massachusetts (1990) 41–72
23. Pentland, A.P.: Perceptual Organization and the Representation of Natural Form. Artific. Intell. **28** (1986) 293–331
24. Lowe, D.: Perceptual Organization and Visual Recognition. Kluwer, Amsterdam (1985)
25. Witkin, A.P., Tenenbaum, J.M.: On the Role of Structure in Vision. In: Beck, J., Hope, B., Rosenfeld, A. (eds.) Human and Machine Vision. Academic Press, New York (1983) 481–543
26. Tarr, M.J., Williams, P., Hayward, W.G. Gauthier, I.: Three-Dimensional Object Recognition Is Viewpoint Dependent. Nature Neurosci. **1** (1998) 275–277
27. Siddiqi, K., Kimia, B.B.: Parts of Visual Form: Computational Aspects. IEEE PAMI **17** (1995) 239–251
28. Kimia, B.B., Tannenbaum, A.R., Zucker, S.W.: Entropy Scale-Space. Plenum Press, New York (1991) 333–344
29. Parent, P., Zucker, S.W.: Trace Inference, Curvature Consistency, and Curve Detection. IEEE PAMI **11** (1989) 823–839
30. Hayward, W.G., Tarr, M.J., Corderoy, A.K.: Recognizing Silhouettes and Shaded Images Across Depth Rotation. Perception **28** (1999) 1197–1215
31. Rodriguez, T., Nilson, C., Singh, M., Hoffman, D.D.: Word and Part. (under review).
32. Imai, M., Gentner, D.: A Cross-Linguistic Study of Early Word Meaning: Universal Ontology and Linguistic Influence. Cognition **62** (1997) 169–200
33. Landau, B., Smith, L.B., Jones, S.: Syntactic Context and the Shape Bias in Children's and Adults' Lexical Learning. J. Mem. Lang. **31** (1992) 807–825
34. Soja, N.N., Carey, S., Spelke, E.S.: Ontological Categories Guide Young Children's Inductions of Word Meanings: Object Terms and Substance Terms. Cognition **38** (1990) 179–211
35. Subrahmanyam, K., Landau, B., Gelman, R. Shape, Material, and Syntax: Interacting Forces in Children's Learning in Novel Words for Objects and Substances. Lang. Cognit. Proc. **14** (1999) 249–281