REPRESENTING SMOOTH PLANE CURVES FOR RECOGNITION: IMPLICATIONS FOR FIGURE-GROUND REVERSAL*

D. D. Hoffman & W. A. Richards

Artificial Intelligence Laboratory and Department of Psychology
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

ABSTRACT

A representation of smooth plane curves for recognition is proposed. The basic representation is a linked list of four primitive shapes, called "codons", which are invariant under rotations, translations and uniform scaling. Psychophysical observations regarding the perception of figure-ground reversals are presented to suggest that a similar representation could be used by the human visual system.

I INTRODUCTION

A vision system sometimes must compute its initial descriptions of a shape without benefit of context. Yet these descriptions should highly constrain the set of possible matches with memory if they are to be useful for recognition. For example, though one cannot reasonably predict the contents of figure 1 prior to seeing it, the shapes are readily recognized. This simple demonstration implies the existence of context-independent rules that provide shape descriptions which can be used to initiate the recognition process. Such rules for smooth plane curves are the subject here.

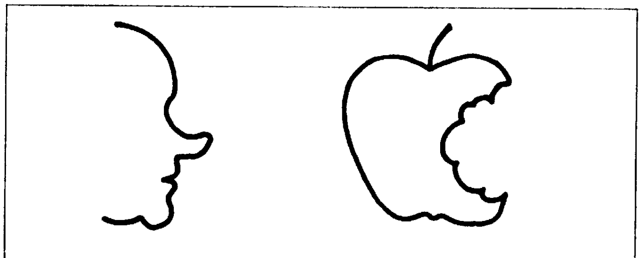

Figure 1. Some shapes recognizable without benefit of context.

*This paper describes research done at the Artificial Intelligence Laboratory and Psychology Department of the Massachusetts Institute of Technology. Support for this research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-80-C-0505 and by NSF and AFOSR under grant 7923110-MCS. D.D. Hoffman was also supported by a Hughes Aircraft Company doctoral fellowship.

**More precisely, the early descriptions should decouple the position, rotation and overall scaling of a shape from the shape itself. This allows the position, rotation, scaling and shape to be made explicit separately, and allows their effects on the recognition process to be disentangled. Thus the goal of rotational invariance for shape descriptions, for example, in no way implies that the rotation of a shape cannot affect its interpretation [1].

To be useful for triggering the recognition process the initial rules should be computable on images, should yield descriptions which are invariant under translations, rotations** and uniform scaling and should provide a first index into a table of shapes in memory. Although a plane curve $\gamma(s) = (x(s), y(s))$ can be specified in many different ways, a description based upon its curvature $\kappa(s)$ is attractive. Such a representation satisfies two of the invariance conditions, namely translation and rotation independence [2]. For any rotation θ and translation (u, v), $\gamma(s)$ is uniquely given by

$$\gamma(s) = \left(\int \cos\phi(s)ds + u, \int \sin\phi(s)ds + v\right),$$

where

$$\phi(s) = \int \kappa(s)ds + \theta$$

However, because curvature is scale dependent, a means for representing $\kappa(s)$ in a scale invariant manner still must be sought. In addition the representation itself should not be a continuous function if it is to serve as an index into memory. Rather, the representation should provide an articulation of $\kappa(s)$ into units which can be described qualitatively. A reasonable approach is to exploit singular points of orders 1 and 0, i.e. maxima, minima and zeroes of curvature [3, 4], since the property of being a singular point is invariant under rotations, translations and uniform scaling.

II MAXIMA, MINIMA AND CURVE "ORIENTATION"

Which points are maxima or minima of curvature depends on the orientation of the curve. Though in general curvature is an unsigned quantity, in the case of plane curves it is possible to assign a sign to the curvature consistently once one of the two possible orientations for the curve is chosen. The orientation is usually specified in figures by an arrow on the curve pointing in the direction in which the curve is to be traversed. By a change in orientation of a plane curve the sign of the curvature changes everywhere along the curve. In particular maxima become minima and vice-versa. The convention adopted here is that figure is to the left and ground to the right as the curve is traversed in an orientation. Thus knowing which side is figure determines the choice of orientation on a curve or, conversely, choosing an orientation determines which side is figure by convention. Minima are then typically associated with concavities of the figure, maxima with convexities (see figure 2). It is possible however for minima to have positive curvature, as in the case of convex closed curves, or maxima to have negative curvature, as when the orientation of the convex closed curve is reversed.

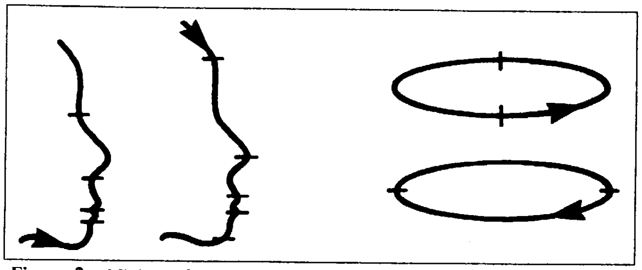

Figure 2. Minima of curvature (indicated by slashes). Arrows indicate curve orientation.

III SEGMENTATION

Maxima, minima and zeroes of curvature are all candidate points for partitioning a curve into units in a manner invariant under rotations, translations and uniform scaling. To choose among them we require that the units should reflect natural parts of shapes [5, 6]. Fortunately, when 3-D parts are joined to create complex objects concavities will generally be created in the silhouette. Segmentation of the image at concavities therefore immediately encodes in a straightforward manner an important property of the natural world that is not captured by maximas or zeroes of curvature. This is our *general position* argument for segmentation at minima of curvature.*

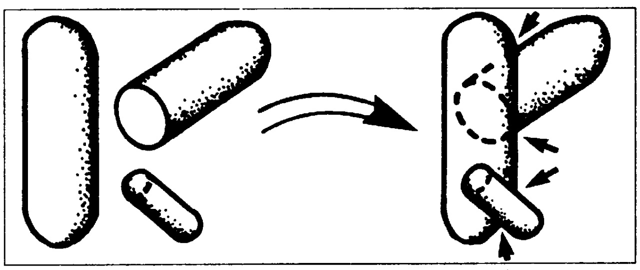

Figure 3. Joining parts generally produces concavities in the resulting silhouette.

IV DESCRIPTION OF PARTS: CONTOUR CODONS

Minima of curvature are used to break a curve into segments, whereas maxima and zeroes are used to describe the shape of each segment. There are four basic types of segments, which we call "contour codons". Furthermore, only certain codon joins (pairwise connections) are allowable.

First, all curve segments contain zero, one or two points of zero curvature. (This assumes that when $\kappa(s) = 0$, $\kappa'(s) \neq 0$). Segments with no zeroes are called type 0 codons, those with two zeroes are called type 2 codons. If a segment has exactly one zero, the zero may be encountered either before (type 1^-) or after (type 1^+) the maximum point of the segment when traversing the curve in the chosen orientation. Thus there are four basic codon types (figure 4).

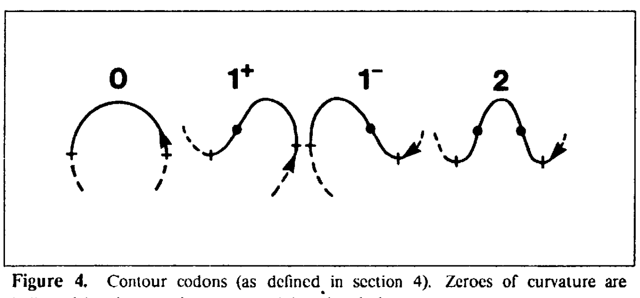

Figure 4. Contour codons (as defined in section 4). Zeroes of curvature are indicated by dots on the curves, minima by slashes.

If desired, codons can be further described by noting the positions of the maximum and any zeroes of curvature within the segment, normalized by the total arc length of the segment. This should be done qualitatively at first and then more quantitatively as is necessary. For example, label the first minimum encountered in traversing a segment the "tail" of the segment, and the other minimum label the "head". Then the position of the maximum can be given crudely as much closer to the head, much closer to the tail, or approximately in the middle. This gives the "skew" of the curve. Zeroes can similarly be described as closer to the maximum point, closer to the head/tail, or approximately in the middle between the maximum and the head or tail.

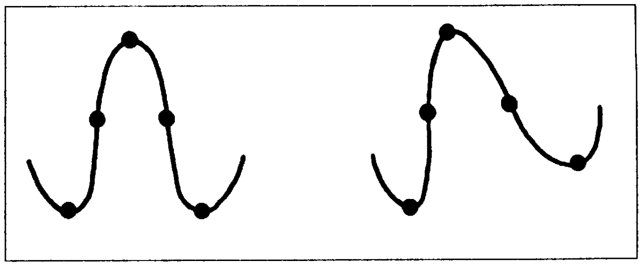

Figure 5. Curve segments with identical singularities but different shapes.

As shown in figure 5, two segments can have identically placed maxima and zeroes, identical curvatures at the maxima and minima, and yet appear quite different. The difference is the behavior of the curvature between the singular points. This behavior can be described in an appropriately invariant manner by the integral of curvature between each of the singular points:

$$\int_a^b \kappa(s)ds = \theta(b) - \theta(a),$$

where $\theta(s)$ is the angle of the tangent at $\gamma(s)$ given by $\theta(s) = \tan^{-1}(y'(s)/x'(s))$. A representation which notes the integral of curvature between the singular points will give different descriptions for curves A and B in figure 5.

There are restrictions on how codons may be joined at minima. Define a *codon join* by the operation $a \circ b$, $a, b \in \{0, 1^-, 1^+, 2\}$ indicating that the head of a is smoothly connected to the tail of b. Note that in general $a \circ b \neq b \circ a$ and hence the codon sequence is critical. Not all conceivable codon joins are possible (figure 6).

The fact that not all conceivable codon joins are allowable suggests that the codon representation may be amenable to error correction techniques. Consider, for example, the codon string $\ldots c_{j-1}\, c_j\, c_{j+1} \ldots$. If all codon joins were allowable then c_j could take any one of the four values $\{0\ 1^+\ 1^-\ 2\}$ regardless of the values of c_{j-1} and c_{j+1}. Thus the value of c_j would be independent of its context. Using figure 6, however, one can show that in actuality the context of a codon restricts its range of possible values to two on average. One can also show that in one third of the contexts c_j is actually uniquely determined.

*When general position is violated special rules may be needed to partition and describe the resulting image contour [5]. This will not be considered here.

	0	1^-	1^+	2
0	+	+	+	+
1^-	+	—	+	—
1^+	+	+	—	+
2	+	+	—	+

Figure 6. Table of allowable codon joins. Rows and columns are labelled by codon type, with the intended codon join sequence at each table entry being (row type, column type). Legal joins are indicated by a +, others by a —.

V RELATION TO PERCEPTION

The representation of plane curves for recognition proposed here can explain the well known observation that a curve can look very different depending on which side is perceived as figure and which as ground [7]. (See figure 7a). The explanation is that a curve looks different because its representations under the two possible orientations are completely different. Since the positions of the minima of curvature are not invariant under a change in orientation (direction of traversal) of a curve, the parts of a curve as specified in its representation can be quite different for the two orientations. (See figure 7b). If one chose to define parts by zeroes of curvature [8], or by minima *and* maxima [3, 4, 10], the parts would not differ under a change in orientation of a curve.

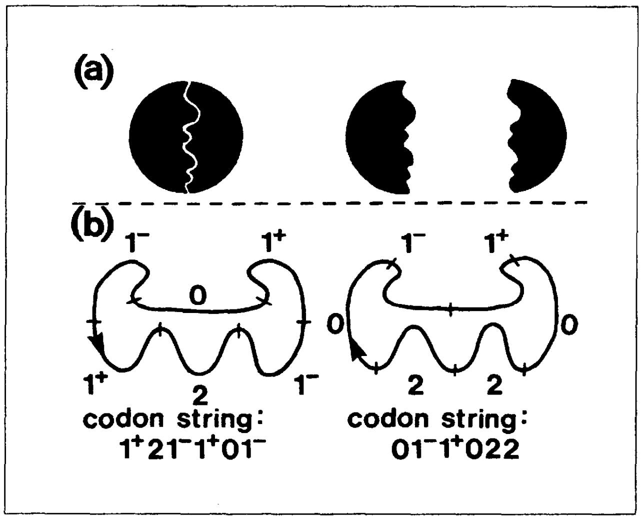

Figure 7. (a) Figure-ground reversal makes bounding curve appear different. (b) Different codon string descriptions are assigned to a curve in its two orientations.

The perceptual significance of minima can also be demonstrated by a simple modification of Rubin's ambiguous face–goblet illusion [9]. If the locations of minima are indicated by occlusions, then the perception of the curve is biased toward one of its two orientations and either the face or goblet impression becomes more apparent (figure 8). However, when this figure is viewed at a distance so the added lines are not visible, then the classical instability returns. Neither highlighting the zeroes nor highlighting the maxima has comparable effect because they do not correspond to natural points for segmentation.

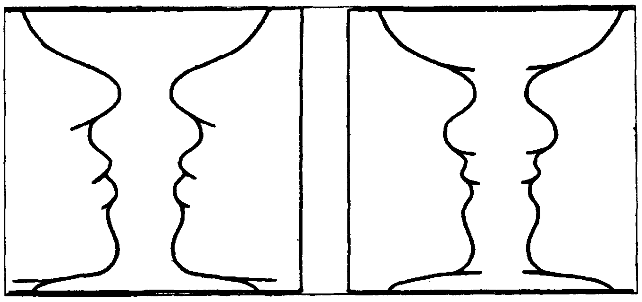

Figure 8. Rubin's face–goblet illusion segmented at minima to reduce the instability.

VI LOCK-KEY AND MIRROR REVERSAL TRANSFORMS

The codon representation is designed to decouple the shape of a contour from its disposition in space and its overall size. Consequently the shape description is invariant under rotation, translation and uniform scaling of the contour. However, as demonstrated in figure 7b, the shape of a contour is not invariant when the direction of traversal along the contour is reversed. It is not difficult to convince oneself also that the shape description is not invariant under a mirror reversal of the contour. The question naturally arises, Are there simple rules that define how the codon string of a contour is transformed when the contour undergoes a mirror reversal or a reversal in direction of traversal (change in orientation)?

In the case of a mirror reversal the rule is quite simple. The mirror transform of a codon string is obtained by reversing the direction in which the string is read (right to left rather than left to right) and reversing the sign attached to each type 1 codon. Thus the mirror transform of $< 1^+ \, 2 \, 1^- \, 1^+ \, 0 \, 1^- >$ is $< 1^+ \, 0 \, 1^- \, 1^+ \, 2 \, 1^- >$. This rule can also be used to find symmetries within a single contour. If, for example, one half of a codon string is found to be the mirror transform of the remainder of the string, a necessary condition for the curve to be symmetric has been found. Note that this applies to skew symmetry as long as zeroes of curvature are not made to appear or disappear by the skew.

When the sense of traversal of a curve is reversed the codon string transformation rule, called the lock-key transform, is unique but apparently not simple. It is perhaps most easily specified as a map from pairs of concatenated codons to codon singletons. The codon doublets which map to each codon singleton are $\{(00,\ 01^+,\ 1^-0,\ 1^-1^+) \mapsto 0\}$, $\{(01^-,\ 02) \mapsto 1^-\}$, $\{(1^+0,\ 20) \mapsto 1^+\}$ and $\{(1^+1^-,\ 1^+2,\ 21^-,\ 22) \mapsto 2\}$. This lock-key mapping can be used, for example, to transform each of the codon strings of figure 7b, into the other.

VII SUMMARY

An approach to the representation of plane curves for recognition has been sketched. It is suggested that minima of curvature can be used to break a curve into parts, and that maxima and zeroes can be used to describe the parts. This approach explains why a curve can appear quite different when figure and ground are reversed. Extensions of the approach to piecewise-smooth curves are presented in Hoffman and Richards [11]. Extensions to surfaces are desirable.

ACKNOWLEDGEMENTS

We thank J. Rubin, J. Scheuhammer, S. True, K. Stevens and A. Bobick for comments on previous drafts.

REFERENCES

[1] Rock, I.: The perception of disoriented figures, *Sci. Am.*, **230**, 1, 78-85 (1974).

[2] Do Carmo, M.: *Differential geometry of curves and surfaces*, Prentice-Hall, New Jersey (1976).

[3] Attneave, F.: Some informational aspects of visual perception, *Psychol. Rev.*, **61**, 183-193 (1954).

[4] Duda, R., Hart, P.: *Pattern classification and scene analysis*, Wiley, New York, 337-341 (1973).

[5] Marr, D.: Analysis of occluding contour, *Proc. R. Soc. Lond.* B. **197**, 441-475 (1977).

[6] Marr, D.: *Vision: a computational investigation into the human representation and processing of visual information*, Freeman, San Francisco (1982).

[7] Attneave, F.: Multistability in perception,*Sci. Am.*, **225**, 6, 63-71 (1971).

[8] Hollerbach, J.: *Hierarchical shape description of objects by selection and modification of prototypes*, Massachusetts Institute of Technology AI-TR-346, 47-51 (1975).

[9] Rubin E.: Figure and ground, In D.C. Beardslee & M. Wertheimer (Eds.), *Readings in perception*, D. Van Nostrand, New York (1958).

[10] Brady M, 1982 "Parts description and acquisition using vision" *Proceedings of the Society of Photo-optical and Instrumentation Engineers*

[11] Hoffman D, Richards W, 1982 "Representing plane curves for visual recognition" *MIT AI Memo* 630