PERCEPTUAL ORGANIZATION AND THE CURVE PARTITIONING PROBLEM*

M.A. Fischler and R.C. Bolles

SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

I INTRODUCTION

A basic attribute of the human visual system is its ability to group elements of a perceived scene or visual field into meaningful or coherent clusters; in addition to clustering or partitioning, the visual system generally imparts structure and often a semantic interpretation to the data. In spite of the apparent existence proof provided by human vision, the general problem of scene partitioning remains unsolved for computer vision.

Part of the difficulty resides in the fact that it is not clear to what extent semantic knowledge (e.g., recognizing the appearance of a straight line or some letter of the English alphabet), as opposed to generic criteria (e.g., grouping scene elements on the basis of geometric proximity), is employed in examples of human performance. Since, at present, we cannot hope to duplicate human competence in semantic interpretation, it would be desirable to find a task domain in which the influence of semantic knowledge is limited. In such a domain it might be possible to discover the generic criteria employed by the human visual system. One of the main goals of the research effort described in this paper is to find a set of generic rules and models that will permit a machine to duplicate human performance in partitioning planar curves.

II THE PARTITIONING PROBLEM

Even if we are given a problem domain in which explicit semantic cues are missing, to what extent is partitioning dependent on the purpose, vocabulary, data representation, and past experience of the "partitioning instrument," as opposed to being a search for context independent "intrinsic structure" in the data? We argue that rather than having a unique formulation, the partitioning problem must be paramaterized along a number of basic dimensions. In the remainder of this section we enumerate some of these dimensions and discuss their relevance.

A. Intent (Purpose) of the Partitioning Task

In the experiment described in Figure 1, human subjects were presented with the task of partitioning a set of two-dimensional curves with respect to three different objectives: (1) choose a set of contour points that best mark those locations at which curve segments produced by different processes were "glued" together; (2) choose a set of contour points that best allow one to reconstruct the complete curve; (3) choose a set of contour points that would best allow one to distinguish the given curve from others. Each person was given only one of the three task statements. Even though the point selections within a task varied from subject to subject, there was significant overlap and the variations were easily explained in terms of recognized strategies invoked to satisfy the given constraints; however, the points selected in the three tasks were significantly different. Thus, even in the case of data with almost no semantic content, the partitioning problem is NOT a generic task independent of purpose.

B. Partitioning Viewed As An Explanation of Curve Construction

With respect to "process partitioning" (partitioning the curve into segments produced by different processes), a partition can be viewed as an explanation of how the curve was constructed. Explanations have the following attributes which, when assigned different "values," lead to different explanations and thus different partitions: (1) Vocabulary (primitives and relations) — what properties of our data should be represented, and how should these properties be computed? (2) Definition of Noise — in a generic sense, any data set that does not have a "simple (concise)" description is noise. Thus, noise is relative to both the selected descriptive language and an arbitrary level of complexity. The particular choices for vocabulary and the acceptable complexity level determine whether a point is selected as a partition point or considered to be a noise element. (3) Believability — depending on the competence (completeness) of our vocabulary to describe any curve that may be encountered, the selected metric for judging similarity, and the arbitrary threshold we have chosen for believing that a vocabulary term corresponds to some segment of a given curve, partition points will appear, disappear, or shift.

* The research reported herein was supported by the Defense Advanced Research Projects Agency (Contract MDA 903-83-C-0027) and by the National Science Foundation (Contract ECS-7917028).

C. Representation

The form in which the data is presented (i.e., the input representation), as well as the type of data, are critical aspects of the problem definition, and will have a major impact on the decisions made by different approaches to the partitioning task. Some of the key variables are: (1) analog (pictorial) vs digital (quantized) vs analytic description of the curves; (2) single vs multiple "views" (e.g., single vs. multiple quantizations of a given segment); (3) simply-connected (continuous) curves vs self-intersecting curves or curves with "gaps;" (4) for complex situations, is connectivity provided, or must it be established; (5) if a curve possesses attributes (e.g., gray scale, width) other than "shape" that are to serve as partitioning criteria, how are they obtained — by measurement on an actual "image," or as symbolic tags provided as part of the given data set?

D. Evaluation

How do we determine if a given technique or approach to the partitioning problem is successful? How can we compare different techniques? We have already observed that, to the extent that partitioning is a "well-defined" problem at all, it has a large number of alternative formulations and parameterizatlons. Thus, a technique that is dominant under one set of conditions may be inferior under a different parameterization. Never the less, any evaluation procedure must be based on the following considerations: (1) Is there a known "correct" answer (e.g., due to the way the curves were constructed)? (2) Is the problem formulated in such a way that there is a "provably" correct answer? (3) How good is the agreement of the partitioned data with the descriptive vocabulary (models) in which the "explanation" is posed? (4) How good Is the agreement with (generic or "expert") subjective human judgment? (5) What is the trade-off between "false-alarms" and "misses" in the placement of partition points. To the extent that it is not possible to ensure a perfect answer (in the placement of the partition points), there is no way to avoid such a trade-off. Even If the the relative weighting between these two types of errors is not made explicit, it is inherent in any decision procedure — including the use of subjective human judgment.

In spite of all of the previous discussion in this section, it might still be argued that if we take the union of all partition points obtained for all reasonable definitions and parameterizations of the partition problem, we would still end up with a "small" set of partition points for any given curve, and further, there may be a generic procedure for obtaining this covering set. While a full discussion of this possibility is is not feasible here, we can construct a counterexample to the unqualified conjecture based on selecting a very high ratio of the cost of a miss to a false-alarm In selecting the partition points. A (weak) refutation can also be based on the observation that if a generic covering set of partition points exists, then there should be a relatively consistent way of ordering all the points on a given curve as to their being acceptable partition points; the experiment presented in Figure 1 indicates that, in general, such a consistent ordering does not exist.

III PRINCIPLES OF EFFECTIVE (ROBUST) MODEL-BASED INTERPRETATION

What underlies our choice of partitioning criteria? We assert that any competent partitioning technique will incorporate the following principles.

A. Stability

The "principle of stability," is the assertion that any valid perceptual decision should be stable under at least small perturbations of both the imaging conditions and the decision algorithm parameters. This generalization of the assumption of "general position" also subsumes the assertion (often presented as an assumption) that most of a scene must be describable in terms of continuous variables if meaningful interpretation is to be possible.

It is interesting to observe that many of the constructs in mathematics (e.g., the derivative) are based on the concepts of convergence and limit, also subsumed under the stability principle. Attempts to measure the digital counterparts of the mathematical concepts have traditionally employed window type "operators" that are not based on a limiting process; it should come as no surprise that such attempts have not been very effective.

In practice, if we perturb the various imaging and decision parameters, we observe relatively stable decision regions separated by obviously unstable intervals (e.g., the two distinct percepts produced by a Necker cube). The stable regions represent alternative hypotheses that generally cannot be resolved without recourse to either additional and more restrictive assumptions, or semantic (domain-specific) knowledge.

B. Complete, Concise, and Complexity Limited Explanation

The decision-making process in image interpretation, i.e. matching image derived data to a priori models, not only must be stable, but must also explain all the structure observable in the data. Equally important, the explanation must satisfy specific criteria for believability and complexity. Believability is largely a matter of offering the simplest possible description of the data and, in addition, explaining any deviation of the data from the models (vocabulary) used In the description. Even the simplest description, however, trust also be of limited complexity; otherwise or it will not be understandable and thus not believable.

By making the foregoing principles explicit, we can directly invoke them (as demonstrated In the following section) to formulate effective algorithms for perceptual organization.

IV INSTANTIATION OF THE THEORY: SPECIFIC TECHNIQUES FOR CURVE PARTITIONING

In this section we present two effective new algorithms for curve partitioning. In each case, we first describe the the algorithm, and later indicate how it was motivated and constrained by the principles just presented. In both algorithms, the key ideas are: (1) to view each point, or segment of a curve, from as many perspectives as possible, retaining only those partition points receiving the highest level of multiple confirmation; and (2) inhibiting the further selection of partition points when the density of points already selected exceeds a preselected or computed limit.

A. Curve Partitioning Based on Detecting Local Discontinuity

In this sub-section we present a new approach to the problem of finding points of discontinuity ("critical points") on a curve. Our criterion for success is whether we can match the performance of human subjects given the same task (e.g., see Figure 1). The importance of this problem from the standpoint of the psychology of human vision dates back to the work of Attneave [1954]. However, it has long been recognized as a very difficult problem, and no satisfactory computer algorithm currently exists for this purpose. An excellent discussion of the problem may be found In in Davis [1977]; other pertinent references include Rosenfeld [1975], Freeman [1977], Kruse [1978], and Pavlidis [1980]. Results and observations akin and complementary to those presented here can be found In Hoffman [1982] and in Witkin [1983].

Most approaches equate the search for critical points with looking for points of high curvature. Although this intuition seems to be correct, it is incomplete as stated (I.e., it does not explicitly take into account "explanation" complexity); further, the methods proposed for measuring curvature are often inadequate in their selection of stability criteria.

We have developed an algorithm for locating critical points that Invokes a model related to, but distinct from, the mathematical concept of curvature. The algorithm labels each point on a curve as belonging to one of three categories: (a) a point in a smooth interval, (b) a critical point, or (c) a point in a noisy interval. To make this choice, the algorithm analyzes the deviations of the curve from a chord or "stick" that is iteratively advanced along the curve (this will be done for a variety of lengths, which is analogous to analyzing the curve at different resolutions) . If the curve stays close to the chord, points in the interval spanned by the chord will be labeled as belonging to a smooth section. If the curve makes a single excursion away from the chord, the point in the interval that is farthest from the chord will be labeled a critical point (actually, for each placement of the chord, an accumulator associated with the farthest point will be incremented by the distance between the point and the chord). If the curve makes two or more excursions, points in the interval will be labeled as noise points.

We should note here that "noisy" intervals at low resolution (large chord length) will have many critical points at higher resolution (small chord length). The distance from a chord that defines a significant excursion is a function of the expected noise along the curve and the length of the chord.

At each resolution (i.e., stick size), the algorithm orders the critical points according to the values in their accumulators and selects the best ones first. To avoid setting an arbitrary "goodness" threshold for distinguishing critical from ordinary points, we use a complexity criterion. To halt the selection process, we stop when the points being suggested are too close to those selected previously at the given resolution. In our experiments we define "too close" as being within a quarter of the stick length used to suggest the point.

After the critical points have been selected at the coarsest resolution, the algorithm is applied at higher resolutions to locate additional critical points that are outside the regions dominated by previously selected points. Figure 2 shows the critical points found along several curves. (We note that this critical point detection procedure does not locate inflection points or smooth transitions between segments, such as the transition from an arc of a circle to a line tangent to the circle.)

The above algorithm appears to be very effective, especially for finding obvious partition points and in not making "ugly" mistakes (i.e., choosing partition points at locations that none of our human subjects would pick). Its ability to find good partition points is based on evaluating each point on the curve from multiple viewpoints (placements of the stick) — a direct application of the principle of stability. Requiring that the partition points remain stable under changes In resolution (i.e., small changes in stick length) did not appear to be effective and was not employed; in fact, stick length was altered by a significant amount in each iteration, and partition points found at these different scales of resolution were not expected to support each other, but were assumed to be due to distinct phenomena.

The avoidance of ugly mistakes was due to our method of limiting the number of partition points that could be selected at any level of resolution, or in any neighborhood of a selected point (i.e., limiting the explanation complexity). One concept we invoked here, related to that of complete explanation, was that the detection procedure could not be trusted to provide an adequate explanation when more than a single critical point was in Its field of view, and in such a situation, any decision was deferred to later iterations at higher levels of resolution (i.e., shorter stick lengths).

Finally, in accord with our previous discussion, the algorithm has two free parameters that provide control over its definition of noise (i.e., variations too small or too close together to be of interest), and its willingness to miss a good partition point so as to be sure it does not select a bad one.

B. Curve Partitioning Based on Detecting Process Homogenity

To match human performance in partitioning a curve, by recognizing those locations at which one generating process terminates and another begins, is orders of magnitude more difficult than partitioning based on local discontinuity analysis. As noted earlier, a critical aspect of such performance is the size and effectiveness of the vocabulary (of a priori models) employed. Explicitly providing a general purpose vocabulary to the machine would entail an unreasonably large amount of work — we hypothesize that the only effective way of allowing a machine to acquire such knowledge is to provide it with a learning capability.

For our purposes in this investigation, we chose a problem in which the relevant vocabulary was extremely limited: the curves to be partitioned are composed exclusively of straight lines and arcs of circles. Our goal here was to develop a procedure for locating critical points along a curve in such a way that the segments between the critical points would be satisfactorily modeled by either a straight-line segment or a circular arc. Relevant work addressing this problem has been done by Montanari [1970], Ramer [1972], Pavlidis [1974], Liao [1981], and Lowe [1982].

Our approach is to analyze several "views" of a curve, construct a list of possible critical points, and then select the optimum points between which models from our vocabulary can be fitted. For our experiments we quantized an analytic curve at several positions and orientations (with respect to a pixel grid), then attempted to recover the original model.

For each view (quantization) of the curve we locate occurrences of lines and arcs, marking their ends as prospective partition points. This is accomplished by randomly selecting small seed segments from the curve, fitting to them a line or arc, examining the fit, and then extending as far as possible those models that exhibit a good fit. After a large number of seeds have been explored in the different views of the curve, the histogram (frequency count as a function of path length) of beginnings and endings is used to suggest critical points (in order of their frequency of occurrence). Each new critical point, considered for inclusion in the explanation of how the curve is constructed, introduces two new segments which are compared to both our line and circle models. If one or both of the segments have acceptable fits, the corresponding curve segments are marked as explained. Otherwise, the segments are left to be explained by additional critical points and the partitions they imply. The addition of critical points continues until the complete curve is explained.

While admittedly operating in a relatively simple environment, the above algorithm exhibits excellent performance. This is true even in the difficult case of finding partition points along the smooth interface between a straight line and a circle to which the line is tangent.

Both basic principles, stability and complete explanation, are deeply embedded in this algorithm. Retaining only those partition points which persist under different "viewpoints" was motivated by the principle of stability. Our technique for evaluating the fit of the segment of a curve between two partition points, to both the line and circle models, requires that the deviations from an acceptable model have the characteristics of "white" (random) noise; this is an instantiation of the principle of complete explanation, and is based on our previous work presented in Bolles [1982].

V DISCUSSION

We can summarize our key points as follows:

(1) The partition problem does not have a unique definition, but is parameterized with respect to such items as purpose, data representation, trade-off between different error types (false-alarms vs misses), etc.

(2) Psychologically acceptable partitions are associated with an implied explanation that must satisfy criteria for accuracy, complexity, and believability. These criteria can be formulated in terms of a set of principles, which, in turn, can guide the construction of effective partitioning algorithms (i.e., they provide necessary conditions).

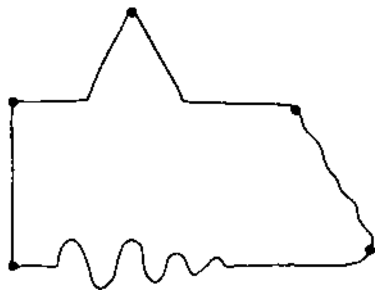
One implication contained in these observations is that a purely mathematical definition of "intrinsic structure" (i.e., a definition justified solely by appeal to mathematical criteria or principles) cannot, by itself, be sufficiently selective to serve as a basis for duplicating human performance in the partitioning task; generic partitioning (i.e., partitioning in the absence of semantic content) is based on psychological "laws" and physiological mechanisms, as well as on correlations embedded in the data.

REFERENCES

A complete list of references and an extended version of this paper are available from the authors.

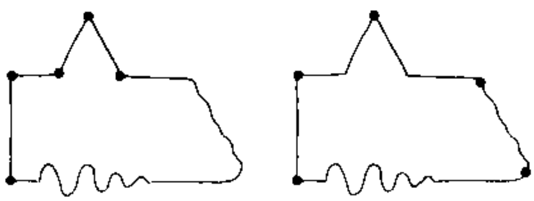
TASK 1: Select *AT MOST* 5 points to describe this line drawing so that you will be able to reconstruct it as well as possible 10 years from now, given just the sequence of selected points.

Since five points were sufficient to form an approximate convex hull of the figure, virtually everyone did so, selecting the 5 points shown below.



TASK 2: Assume that a friend of yours is going to be asked to recognize this line drawing on the basis of the information you supply him about it. He will be presented with a set of drawings, one of which will be a rotated and scaled version of this curve. You are only allowed to provided him with *A SEQUENCE OF AT MOST 5 POINTS.* Mark the points you would select.

Since 5 points were not enough to outline all the key features of the figure, the subjects had to decide what to leave out. They seemed to adopt one of two general strategies: (a) use the limited number of points to describe one distinct feature well (illustrated by the selection on the left), or (b) use the points to outline the basic shape of the figure (shown on the right).



TASK 3: This line drawing was constructed by piecing together segments produced by different processes. Please indicate where you think the junctions between segments occur *AND VERY BRIEFLY DESCRIBE EACH SEGMENT.* Use as few points as possible, but no more than 5.

The constraint of being limited to 5 points forced the subjects to consider the whole curve and develop a consistent, global explanation. The basic strategy seemed to be a recursive one in which they first partitioned the curve into 2 segments by placing a breakpoint at position 1 and another one at either position 2 or position 3 to separate the smooth curves from the sharp corners. Then they used the remaining points to subdivide these segments according to a vocabulary they selected that included such things as triangles, rectangles, and sinusoids. For example, almost everyone placed breakpoints at positions 3 and 4 and described the enclosed segment as part of a triangle. Similarly the segment between positions 1 and 5 was generally described as a decaying sinusoid. It is interesting to note that in task 1 the subjects consistently placed a point close to position 5 but always farther to the right, because they were trying to approximate a convex hull. The different purposes led to different placements.

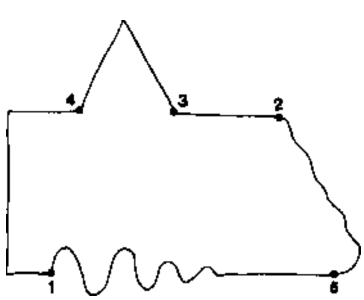


FIGURE 1 EXPERIMENTS IN WHICH HUMAN SUBJECTS WERE ASKED TO SEGMENT A CURVE

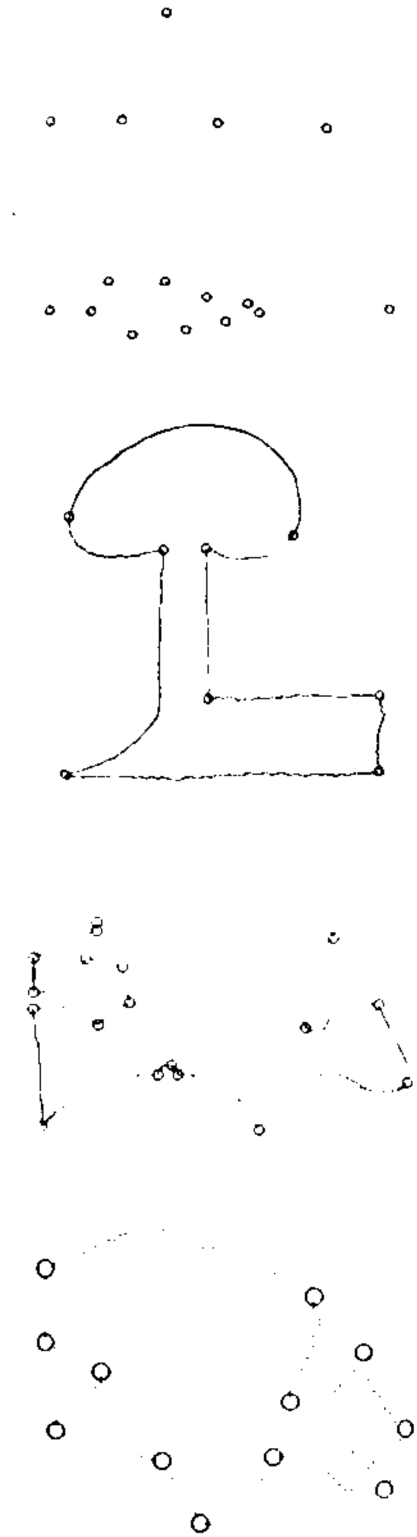


FIGURE 2 LOCAL DISCONTINUITY PARTITIONING