A SEMANTICS-BASED DECISION THEORY REGION ANALYZER

by

Yoram Yakimovsky and Jerome A. Feldman
Artificial Intelligence Laboratory
Stanford University
Stanford, California

ABSTRACT

The problem of breaking an image into meaningful regions is considered. Bayesian decision theory is seen to provide a mechanism for including problem dependent (semantic) information in a general system. Some results are presented which make the computation feasible. A programming system based on these ideas and their application to road scenes is described.

INTRODUCTION

The problem of segmentation, breaking a complex image into sections, is a central problem in machine perception. The analogous problem arises in the analysis of speech and, for,that matter, in any problem of overwhelming size. We will concentrate on the image segmentation problem, but most of the ideas are of wider applicability. The main idea is to apply (Bayesian) decision theory techniques and the use of problem-dependent information (semantics) to attack the image segmentation problem.

The segmentation problem for T.V. images is as follows: given a picture of some scene, we have a rectangle composed of some "00 x 300 points and for each such point some information on the light intensity and perhaps color. For any further processing 60000 points is far too many, so, depending on the perception task that we have in mind, the image should be segmented into regions. Each of these regions should be meaningful in the problem domain and the relevant information needed for the specific task should be easily obtainable.

There has been a great deal of work on segmenting images and a certain limited success has been achieved. 21 Some biological and meteorological images can be effectively segmented using known techniques. However, for images like those arising in road scenes or presented to assembly-line robots, the existing algorithms do not suffice. A major problem is that the existing algorithms use absolute criteria such as intensity difference, boundary strength, etc. to form regions,2,3 But the criteria for what Is a "region" will surely vary with context. Certain shades of green, yellow and brown might be merged into a single region of grass in a scene, yet distinguishing the same set of colors might be crucial for region separation in another scene or even in another part of the same scene (assume for instance that a yellow car occludes part of the grass). Another critical consideration is the goal of the perception. For Bome problems separating the green grass from yellow grass will be essential, In others completely confusing.

The importance of goal direction and context-dependence for effective problem solving is now well understood In artificial intelligence and scene analysis is just another example. One can certainly write a special purpose region analyzer for a fixed class of images and it will work better than any general algorithm. This, in fact, has been done successfully by Brice and Fennema[3] and Harlow and Eisenbeis[13] and Is sometimes just the right thing to do. The obvious difficulty with this ad-hoc approach is that it requires a lot of work to build or modify each individual program.

In this paper we present *a* theoretical framework for a general system incorporating context dependence in a region analyzer. The theory has been developed Into a computer program for region analysis and a number of experiments on real pictures have been performed. There is an enormous amount of work remaining to be done, but we feel that a promising start has been made.

Before describing the system in more detail, we must make one additional point of clarification. It is a tenent of artificial intelligence research that any information that can be brought to bear will be helpful in a given task. This is especially true in machine perception, but our current efforts do not attempt to exploit it fully. Region analysis is assumed to be a preliminary (relatively fast) partitioning of an image before further processing. For this reason, we have made no attempt to include semantic features like three-dimensional shape analysis in the region analyzer. We are still studying the capabilities of our semantic structure. As more experience becomes available we will be able to determine which Information should be used in the segmentation process and which should be left for higher level processing.

1. THEORY

The underlying theory of our system is Bayesian decision theory. The ideas are quite beautiful and powerful, but have not received as much attention as they should from artificial intelligence workers. Even a brief description would be beyond the scope of this paper: "Elementary Decision Theory" and "Decision Analysis - Introductory lecures on choice under uncertainty" are good introductions and "Optimal Statistical Decisions" and 'Mathematical Statistics a Decision Analysis Approach" are advanced texts. 4,20,5,8 The two central ideas are the use of a utility function to measure the value of various alternatives and an optimality theorem. This theorem shows that any adequate (admissible) strategy is equivalent to a strategy of maximizing expected utility for some choices of utility function and probabilities. The theory provides a complete world view (like, e.g. logic) and has been applied to many management problems. The difficulty in practice is that it is difficult to select the utility and probability functions and to actually carry out the computations. We describe below how we attack these problems.

For region analysis, we can define the utility to be the probability that the analysis is correct, contingent upon two factors: the (a priori) context knowledge about the picture domain, and the values of measurements on this particular image, i.e. we are to maximize:

P(global-tnterpretatton | context.values of measurements)

An interpretation will divide the image into regions and attach a type label (meaning) to each region. One choice of evaluating the overall interpretation would be obtained by considering each region independently. If for a given partition of the image Into regions we have (R(i) i=I>nI regions then the interpretation assigns label IHT(i) to region R(i). The values of INT(i) will be sky, grass, road etc. depending on context and goals. Then assuming independence we will want to maximize the expression:

$$\prod_{1} \text{PfR(i) is INT(i) | context,values of measurements on R(i)t}$$

over all partitions of the image into regions and assignments of labels to regions. This is quite conventional so far and is, in fact, too simple for our purposes. We want to account for two additional considerations. First we must use the model to get a good segmentation of the image into regions. For example, we might want to merge green, yellow and brown patches to create the whole area that we call grass. Secondly, we want to use additional semantic constraints e.g. the grass is below the sky, to influence the total probability of an analysis of the scene.

In an attempt to enrich the semantic structure to support more of the problem knowledge and to provide for control mechanism on the region growing algorithm, the semantic structure was allowed to have also "first order structure". In addition to the properties of each individual region wc have, for each pair of adjacent regions of some interpretations, expected relative properties and some expected features of their common boundary line. For instance if we have two adjacent regions, one which is named "sky" and the other "hill" then we expect that the sky is above the hill, is a brighter blue than the hills, and that the boundary is usually a more or less horizontal smooth line. The relative properties are usually more significant than the absolute properties since they are less sensitive to variation between pictures. This semantic model is too limited to describe all that is known of a scene, but many classes of scenes can be Segmented properly with first order methods.

Recall that we want to get a partition of the input and interpretation for the segments (regions) and boundaries so as to maximize the likelihood of having the right interpretation. Let R(i) be the i-th region, B(i,j) the boundary between region R(i) and R(j) (if it exists) and the label of R(i) be INT(i). Then with our first order assumption the expression that we want to maximize is:

$$\text{Pfglobel interpretation | values of measurementil} = \prod \text{?l P[R(i) is INT(i) j values of measurements on R(i)}\}$$

$$\prod \text{x'fl PfB(l.j) is between INT(i) and INT(j) | B(i,j)'s measurements]}$$

$$\prod_{B(i.j)} \qquad [1]$$

The use of [1] represents more than just our belief that properties of individual regions and boundaries will suffice for our semantics. It also entails an assumption that the probability can be factored into the product above. This amounts to assuming that the probabilities of interpretations of each region (boundary) are dependent on the local properties of the individual region (boundary) and are independent of all other measurements. Thus boundary B(i,j) which is the boundary between R(i) and R(j) is evaluated as boundary between INT(i) and IKT(j), where R(i) is labeled INT(i) and R(j) is labeled INT(j).

For example, if INT(i) is "sky" and INT(j) is "hill", the evaluation will include factors involving the expected direction, smoothness, etc. of a boundary between sky and hill. TheBe factors are assumed to be independent of the particular color etc. of the sky and hill. If the independence assumption seems to be unreasonable, consider the following argument:

$$\text{PfinterpretationI - values of measurement, } \textit{context)} = \frac{\text{Pfvalues of measurements(interpretation, context"! x Pjinterpretatlon | context Rvalues of measurements [context)}}{\text{Pfvalues of measurements(interpretation, context"!}} \qquad [2]$$

Now

Pfvalues of measurements on R(i) I R(i) is INT(i), context!

and

Pfvalues of measurements on B(i,j} | R(i) is INT(i) and R(j) is INTfj).context]

are plausibly considered independent of each other. A similar argument can be Used for the factorization of the other two terms in the expression on the right of [I]. Putting these terms together give us back what we have in [1]. Prom all picture models which were described in the literature the model described in reference ly is most similar to ours in the basic methodology aspects.

For a given utility function (like [11) there are standard techniques in decision theory for finding the maximum utility. Unfortunately, the general techniques are too slow and much of our effort has gone into developing algorithms for efficiently computing an approximately optimal partition. The region growing algorithm starts with many small regions and on each iteration merges two adjacent regions (regions with a common boundary). The two basic decisions are which pair of regions to merge on each iteration and when to stop the algorithm. These two decisions can be controlled directly by the limited probabilistic semantic world model that we have-

In general, on each iteration of region growing the pair of regions whose common boundary is the weakest in the current image partition will be merged. Hence the control of the region growing algorithm is by evaluation of the boundary strength. We will show how our semantic representation can be used directly to compute the goundary strenth.

The second task of the semantics is to produce the stopping criterion for the region grower. In our case since we want to maximize

Pfinterpretation | measurements' values, context)

the optimal partition will be the one with that interpretation which maximizes this likelihood estimate over all partitions and all possible interpretations of partitions which do not allow faLse boundaries (boundaries between two regions which are interpreted as parts of the same final object).

In order to have an effective way to determine that probability we need a relatively fast way to obtain or approximate for a given partition the optimal interpretation and its value.

In the next section, we will describe relatively fast methods for computing upper and lower bounds on the optimal value of the probability of a given partition.

The bounds on the value of the global interpretation will be used as follows: The algorithm will collapse regions, and generate a sequence of image partitions. For each partition generated, the bounds on the possible value of the best interpretation will be evaluated. Then when the collapsing has been carried too far (as observed by strong decline of the possible Interpretation value) the system will hack-up to the most promising partitions observed while growing the regions (as indicated by the lower and upper bounds estimates of the quality of the partition observed). Next we will search for the best interpretation for the partitions observed whose bounds were high enough to make it possible that they are the best partition observed. The current algorithm will simply choose the best of these, but more sophisticated procedures can be used if necessary.

Althernatively we can use the procedure which assigns meaning to all regions directly as the stopping criterion. That is, if the best global assignment found for the given image partition does not interpret any boundary as an erroneous boundary (e.g. a boundary between regions of the same interpretation) we stop merging (see discussion below on lower bound estimates for the image partition for more details).

2. PROGRAM ORGANIZATION

The program has four basic sections: 1) the initialization, 2) initial assignment and boundary strength evaluation, j) heuristics to evaluate a partition by approximating its optimal assignment, and *4)* a limited interactive learning system.

The first step is to take a vector of values of measurements at a set of sample points (in our system usually 500-1000). The local measurements currently indicate only the dominant gray level around the sample point or if the image is observed through several color filters the dominant gray level through each of the color filters.

We employ a preliminary region-merging for initialization. The idea is to use a very crude algorithm on the reduced problem. The simple merging algorithms considered were:

1) Take an image point and grow a region around it consisting of all image points which can be connected to the starting point by a path of points which satisfies the following conditions: a) each adjacent pair of points along the path are adjacent geometrically, b) the jump in the value of the measurements vector between two adjacent points along the path is less than some threshold. 2^ Note that some modifications are needed to treat gradual but strong changes.

2) This initialization is an extension of 1. It initializes as in 1 and then collapses, independent of order, all boundaries with strength less than some threshold. The advantage here over method 1 is the option to use more sophisticated boundary strength evaluation.3

Our initialization method utilizes a sampling technique initialized as in 1 (taking the connecting path to be a path of sample points) and then merges regions iteratively by eliminating the globally weakest boundary first. That is, on each iteration the pair of regions whose common boundary is weakest in the current image partition is merged into one region.

The boundary strenth in this stage is evaluated directly from the differences across the boundary and its geometrical structure. 28 The stopping crietrion in this case can be a threshold on the weakest boundary, that is the merger is stopped when the weakest boundary is stronger than a given threshold. This threshold is chosen very conservatively so as to stop this simple minded region grower before it produces false merges. In our experiments it turned out that the simple initialization algorithm had to be stopped quite early so that the semantic control was called with about 100 regions present, see reference 28 for more details.

The main algorithm first computes additional properties (like shape) of the regions and boundaries resulting from the initialization. It then assigns probabilities to the tentative interpretations of the regions, i.e. computes

PfR(i) is X values of measurement on R(I)"(

The boundary strength may be evaluated by two related methods: 1) the probability that the boundary is a real boundary (a boundary between different objects in our semantic world model) and 2) the change in the value of the interpretation as a result of eliminating the boundary. We will describe here the first of these which is the one currently used. The second method has some advantages and will be discussed below.

We approximate the probability of the boundary to be real as follows:

$$\frac{\{P \text{ is a boundary between INT1 and INT2} \mid \text{B's features}) \times P(R1 \text{ is INT1} \mid R1 \text{ features}) \times P(R2 \text{ is INT2} \mid R2 \text{ features})\}}{INT1 \neq INT2} \qquad [3]$$

$$\frac{\{P(B \text{ is a boundary between INT1 and INT2} \mid \text{B's features}) \times P(R1 \text{ is INT1} \mid R1 \text{ features}) \times P(R2 \text{ is INT2} \mid R2 \text{ features})\}}{INT1, INT2}$$

The next step is to search for a partition which yields a good value of [1]. This involves both forming partitions and computing the value attainable from a labelling of regions in each partition.

A lower bound on the value of an image partition is computed by actually finding a good global interpretation using a simple fast algorithm. Briefly what we are doing is to take the region of highest confidence interpretation and assign to it its most probable interpretation. This assignment allows the program to update the probabilities of adjacent regions of the newly interpreted region by considering the boundary features of the newly assigned region. Then the region of highest confidence from all un-interpreted regions is assigned, etc. This is essentially starting a depth first search of the tree of interpretations and yields a value for the partition which is the desired lower bound. Extending this search to full tree search would yield the optimal interpretation. More details on the sequential assignment process are given below.

Recall that we want to approximate the maximum possible value of the expression in [3] over all possible values of INT(i) for a given picture partition.

The assignment procedure we use to estimate the best possible assignement of INT(i) for all R(i) for a given image partition is as follows:

1) Compute for each region the ratio (based just on local measurements of the region) between the most likely interpretation and the next most likely interpretation. This ratio will be called the CONFI(REG). Let xl be such that P(R is x1 values of measurements on R) is maximized for R and let x2 be such that P{R(i) is x2 values of measurements on R(i)l is the next highest. Then

$$\text{CONFIfR(i))} - \frac{P(R(i) \text{ is xl measurements of } R(i))}{P(R(i) \text{ is x7 measurements of } R(i))}$$

2) Sort the regions by their confidence ratio.

3) Assign the region with highest confidence (the one with highest ratio) its most likely interpretation.

4) Update probabilities of different assignment to regions that were not assigned already, assuming that the last assignment is true. Let the region assigned most recently be R(l) and its interpretation be INT(l). Now if R(i) has boundary B(i,i) with R(l), then for any interpretation x of R(i) in evaluating equation 1 above, there will be a term of the form

P{R(i) is x values of measurement on R(i)]

from the first product and one of the form

P[B(l,i) *is boundary between* INT(l) *and* x B(l,i)'s measurements' values)

from the second product. Therefore a better approximation of the probability of R(i) being x, assuming that R(l) is INT(l), is

$$\text{Pnew(R(i) is x)} = \text{Pold(R(i) is x) x P(b(l,i) is between INT(l) and x B(i,j)'s features)}$$

Thus we use the new information to find a more accurate probability for the different possible assignments for R(i), by counting the newly interpreted region R(l).

We do that updating to all possible interpretations for all adjacent regions of R(l).

5) Compute the new confidence ratio and rrsort the regions by the new confidence ratios.

6) If any region *is* still unassigned go to step 3 else exit.

This process of assigning interpretations iteratively provides a good guess on the possible best interpretation, but it does not guarantee the total maximization of our product. We can extend the current algorithm into a full tree search (undoing some assignments and trying alternative ones) to get the best interpretation. This will be a depth first search in the tree of all possible assignments, where each node Mill stand for the assignment of a meaning to a region. For efficiency purposes we can use various pruning techniques to reduce the search effort required to secure optimality,^° Our current algorithm is the portion of the tree search up to the point where we get to the first terminal node (first global assignment). One should also note that the same sequential assignment and extension into tree search can be applied to an extended first order world model, where we allow relations between any two regions (not necessarily adjacent) if we continue to assume independence. The only difference is that we will have to update the probabilities and confidence of all regions not just those regions adjacent to the newly interpreted region.

We have the option to use this assignment procedure as a region grower by taking all pairB of adjacent regions that were assigned the same meaning and merging them. To avoid false merging we consider all regions which were assigned meaning with a low confidence level not mergable into other regions. This approach may be extended by adding it to the meaning assignment algorithm as a new step 3,5. If any adjacent region of the newly interpreted region is already assigned a meaning and it is identical with the meaning of the newly interpreted region, then merge the two together. From that point on the unified region will be considered in updating probabilities of other, not yet interpreted regions.

We can use the two extensions (merging on the run, and full tree search) together. This will generate a very reliable meaning assignment concurrent with a region growing procedure which has backup capabilities. It will be however, relatively slow.

Upper Sound

The upper bound is computed by relaxing the consistency constraint. This condition means that a boundary between two regions of known interpretation has to be counted in as a boundary between those two interpretations. We relax this condition by hreaking the product (l) into local sub-products and finding the best local interpretation for the terms involved in this subproduct. We take the best possible value for each sub-product separately, and, multiplying them, obtain in an upper bound on the value of the best global interpretation. One such relaxation is to consider all regions and boundaries independently and to assign for each the best possible interpretation considering only its own properties. The product of all these probabilities is an upper bound on the value of equation 1. It is this sort of estimate which would be used to approximate the single step improvement in the second method of boundary evaluation mentioned above. An exact computation of the change in interpretation value would be too time consuming. We do not yet know which boundary strength computation will be better.

Given the lower and upper bounds of the value of the best possible interpretation for a given picture partition, a variety of graph searching techniques can be applied to find a suitable interpretation and to pick out the best partitions observed for more detailed investigation using the full semantic knowledge.

^. LEARNING

One of the basic problems with any recognition system is the development of sharp classification capabilities for objects (in our case interpreting regions and boundaries). Our case is especially complicated since we need to recognize portions of objects, spurious boundaries and to overcome partial occlusions. To make the game of developing these capabilities easier we developed an interactive learning system. Its main task is to carry out bookkeeping jobs, to estimate probabilities, and to point out pitfalls and options for improvement in the classification scheme.

The non-parametric approximation of the probability density function works as follows: Given a set of measure functions on some class of objects, we break the space of all possible combinations of values of those measurements into cells (not necessarily cartesian), trying to get an effective classification. That is, given that the values of the measurements of an object fall into some cell, we want that often the probability estimate of the real meaning of the object to be high. Given a fixed partition of the measurement vector space into cells we want to learn the probabilities of different interpretations of objects whose measurements' values fall into a cell. This is done by keeping, for each cell and for each possible interpretation, the count of how many times in the past the value of the measurements of objects of that interpretation fell into this cell. The probability of that interpretation is just the number of times the measurements of an object of that interpretation fell into the cell, divided by the total number of objects which fell into that cell.

This brings us to the second learning system which would try to create a cell structure with as few cells as possible while attaining a good classification among the possible interpretations. For this purpose we sould utilize an augmented classification tree whose leaves correspond to the cells. The augmented tree also allows representation of the fact that two measurements are independent.

Currently this tree is generated interactively. To generate a sub-optimal classification tree automatically the System would keep a whole history list containing objects observed in the past, their properties and their real meaning. Based on this history the system could try to order the application of measurements so as to get good and cheap classification, creating as few as possible cells (leaves), while still keeping the good classification probability high. It also has the ability to point out cells that are not sufficiently discriminating so that they may be worked on interactively or automatically (primarily by breaking each such cell into finer subcells, so that for each eubcell the classification is more reliable). A detailed description of this learning and classification system is given in reference 28. Techniques for organizing the classification tree so as to get near optimal sequential classification is described in Slagle and Lee, where game (a-b) type tree Search is utilized in creating the decision tree.[23]

Such learning techniques are common to many patforn recognition and sequential decision problems. A vast amount of research, both theoretical and experimental, has been done in this area. Reference 10 is a good description of the theory and Reference 6 is a good introduction to various applicable techniques. It is interesting to comment that a learning scheme similar to the first (emphasizing correlations) was developed by Arthur Samuel. Attempts are now being made to apply this learning scheme to speech segmentation and recognition. 25

It is interesting to compare our technique with the nearest neighbor classification which is investigated in various-papers. 29 This principle is to take for a new unknown occurence of an object, the interpretation of the object observed in the past whose features of the new object. There are two deficiencies in this approach. First, only rarely is there an obvious metric on the space of values of measurements, and hence only rarely is it clear exactly how to measure distance in the features of two objects. Secondly, it is very hard to search for the nearest object observed in the past (unless we are in one dimension) since we have to compute the distance from many examples observed in the past to get the minimal distance. An effective way of reducing the search time will call for breaking the space into cells the way we do. That is, locating first the cell into which the measurements of the new object fall and then searching only among known objects whose measurements fail into that cell for the nearest one, ignoring objects which fall into other cells. Thirdly, the answer returned is just one possible interpretation and not a list of different possible interpretations with various probabilities. Extending the nearest neighbor principle to find the n- nearest objects and computing the probabilities of different interpretations based on them will make the computation even less efficient because of search time and will force even more reliance on space partitioning than the method we currently use.

4. RESULTS AND CONCLUSIONS

As a first experiment, the program was applied to the problem of segmenting images which might be seen driving on a road in the vicinity of the laboratory. The analysis was simplified by assuming the camera was in an upright position. There were six possible labels for a region: sky, road, roadside vegetation, car, shadow of car and tree. The non-terminal nodes in the classification tree are calls on integer valued functions. Some properties and the number of possible values for each are: light intensity (4) color hue(8)l color saturation *(4),* size (5, logarithmic), vertical position *(4),* horizontal position (4), position on edge of image (4) and some crude shape descriptors.

With six types of region, we get 18 (6x6/?) types of boundary. Some boundary properties used, with the number of values for each are: relative size (6 logarithmic), relative intensity (6, logarithmic), relative color (3 green relation x 3 red x 3 blue), boundary shape and orientation (21 classes), relative position *(4* right extremes x 4 left extremes x 4 above extremes 4 'i below extremes), boundary length *(5)* and position of boundary in frame (5x5). The algorithms for computing the various properties and the discrimination in each were chosen intuitively.

Initial values for the counts in cells were also set Intuitively, and the learning routine was used interactively to refine them (approximating the probabilities and breaking cells to finer cells if desired for better classification).

The program was able to segment the scenes correctly using only region properties except that it had difficulty isolating the image on the car on the road. Since a car can be of any color, the program either needs to make use of boundary relations, (e.g. the car is on the road) or perhaps shape discrimination should be made more sophisticated.

A region growing algorithm based on absolute properties would not work in these scenes mainly for the following reasons: 1) The trees and sky generated very many regions that were more varied in the values of their measurements than any other thing in the picture. 2) The sides of the road are patches of brown, green and yellow. 3) Strong shadows appear frequently on the road.

The second domain to which the system was applied was left ventricular angiograms (x-ray images of the left ventricular made visible by injection of a radio-opaque dye). These angiograms are useful for various cardiologic applications since they allow observation

of myocardial movement. The semantics used for this application described the heart interior, chest cavity background and the dark frame border. No color was available here, and as a result light intensity, position and shape was the major recognition tools. In addition the non-semantic region grower had to stop at a relatively early stage because of noise and lack of high contrast border. The number of regions on termination of the non-semantic region grower was two hundred. It is encouraging that the adjustment to the second domain was very easy. We hope that in the future a general and rich library of feature extracting routines with the capability of working on many models will be achieved.

Shown below are illustrations of the results of experiments. All pictures are taken from a graphics terminal with gray level capabilities. There are six bits available per image point. Five are used for displaying the original picture, while the high order bit is used for the overlay of displaying the boundary lines.

This is a preliminary version of a general system for utilizing decision theory in scene analysis. There are a number of ideas from both areas that have yet t" be tried and many experiments yet to be run. However, there are already some additional considerations which should be mentioned.

The most restrictive assumption in the current program is assumption that the interpretation of a region depends only upon adjacent regions. There are ways of selectively relaxing this rule so that occluded objects can be understood without having each region depend on all others. In fact, the entire approach will stand or fall on the question of whether there is sufficient independence to allow for good performance without prohibitive calculation cost.

The choice of local measurements around each point is, of course, another crucial factor. The idea of relatively coarse sampling allow us to apply more operators, including ones like Hueckel's, or 14(texture finders which inherently involve many paints. There is the additional important potential for variable density sampling, possibly using planning in the manner of Kelly.15

A more difficult task would be to effectively incorporate 3-D constraints, as done so successfully for blocks by, i.e., Falk and Waltz (this would call for the addition of vertex properties).7,27 There are many possible refinements to the learning procedure, especially on the question of what measurements are important.

BIBLIOGRAPHY

1. Barrow, H.G., Ambler, A.P. and Burstal, R.M. "Some techniques for recognition structures in pictures" Department of Machine Intelligence and Perception, University of Edinburgh, 1972.

2. Barrow, H.G. and Popplestone, R.J. "Relational Description in Picture Processing" Machine .Intelligence. 6, 377-596, 1971.

3- Brice, C.R. and Fennema, C.L. "Scene Analysis Using Regions" Journal of Artificial Intelligence, 1, 205-226, 1970.

4. Chernoff, H. and Moses, L. "Elementary Decision Theory" Wiley Publishing Company, N.Y., 1959-

5. De-Groot, M.H. "Optimal Statistical Decisions" McGraw-Hill, N.Y., 1970.

6. Duda, R.D. and Hart, P.E. "Pattern Classification and Scene Analysis" Wiley, 1973-

7. Falk, G. "Computer Interpretation of Imperfect Line Data as a Three Dimensional Scene" Artificial Intelligence Memo 139, Stanford University, I970.

8. Ferguson, T.S. 'Mathematical Statistics, A Decision Analysis Approach" Academic Press, 1967•

9. Feldman, J.A. "A Way of Looking at Things" Dept. of Computer Science, Hebrew University, Jerusalem, January 1971-

10. Fu, K.S. "Sequential Methods in Pattern Recognition and Machine Learning" Academic Press, N.Y., 1968.

11. Griffith, A.K. 'Mathematical Models for Automatic Line Detection" JACM. 62-80, 1973-

12. Guzman, A. "Computer Recognition of Visual Object In a Three-Dimensional Scene" M.I.T., 1969-

13 Harlow, C.A. and Eisenbeis, S.E. "The Analysis of Radiographic Images" Technical Report, Dept. of Electrical Engineering, University of Missouri-Columbia.

14. Hueckel, M.H. "An Operator Which Locate Edges in Digitize Pictures" Artificial Intelligence Memo 10^, Stanford University, I969.

IS Kelly, M.D. "Visual Identification of People by Computer" Artificial Intelligence Memo 50. Stanford University, I969.

16. Nilsson, N.J. "Problem Solving Methods In Artificial Intelligence". McGraw-Hill.

17. Freparata, Franco P. and Ray, S.R. "An Approach to Artificial Nonsymoblic Cognition" Information Science, *k,* 65-86, 1972.

18. Pingle, K.K. and Tenenbaum, J,M. "An Accomodation Edge Follower", IJCAI-2.

19. Quam, L.H. "Computer Comparison of Pictures" Stanford Artificial Intelligence Memo *kh* , Stanford University, 1968.

20. Raiffa, H. "Dicision Analysis - Introductory Lectures on Choice Under Uncertainty" Addison Wesley, July 1970.

21. Rosenfeld, A. "Picture Processing by Computer" Academic Press, N.Y. I969.

22. Samuel, A.L. "Some Studies in Machine Learning Using the Game of Checkers II - Recent Progress" IBM Journal of Research and Development, 11, 601-617, 1967.

23. Slagle, J.R. and Lee, R.T.C. "Application of Game Tree Searching Techniques to Sequential Pattern Recognition" CACM, IO3-IIO, Feb. 1971-

24. Thomas, A. "Color Perception" Stanford University 1973-

25. Thosar, R.B. and Samuel, A.L. "Some Preliminary Experiments in Speech Recognition Using Signa-

ture Table Learning" Artificial Intelligence Memo, Stanford University, (to appear),

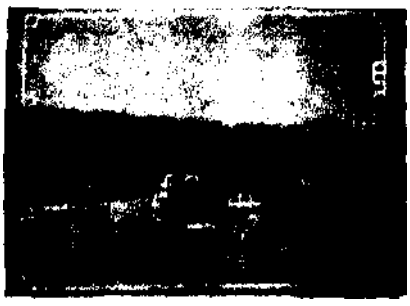
26. Vicens, P. "Aspects of Speech Recognition by Computer" Artificial Intelligence Memo 85, Stanford University.

27. Waltz, D.L. "Generating Semantic Description from Drawing of Scenes with Shadows" M.I.T. TR-27I.

28. Yaklmovsky, Y. "Picture Processing, Use of Probabilistic Semantics for Region Growing and Recognition" Stanford University. (Forthcoming dissertation).

29. Cover, T.M. and Hart, P.E. 'fewest Neighbor Pattern Classification" IEEE Trans. Information Theory, Vol. 11-13, 21-27, 1967-

(A-I) The output of the segmentation base on path connectivity when it is stopped by the default stopping criterion. The result ie image segmented into a few hundred regions.



(A-2) The effect of reducing the number of regions to 40 using path connectivity algorithm (using more liberal threshold than our current stopping threshold).

(A-3) The output of the region grower which melts weakest boundary first with non-semantics boundary strength evaluation. This is the result of stopping with the default stopping cri terion.

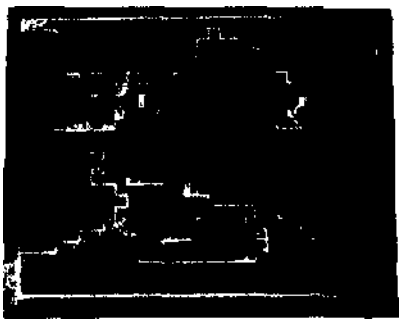
(A-4) Result of merging regions down *to 38* regions using weakest boundary first algorithm and non-semantics boundary strength evaluation. Note that the top of the car ie melted into the roadside vegetation.

(A-5J Result of attempt to reduce the number of regions to 28 without using semantics (melting weakest boundary first non-semantic boundary strength evaluation}.

(A-G) Output of region grower based on semantics. (Melting weakest bounQary first where boundary strength is computed using the semantic world model).

(B-2) Output of the non-semantic weakest boundary melted first regi On grower.

(A-7) F inal grouping of regions based on the interpretat ion assigned to them by the world model. Regions whose meaning was assigned with confidence less than 10 are not tnergable. They occur usually on the real boundary between two regions.

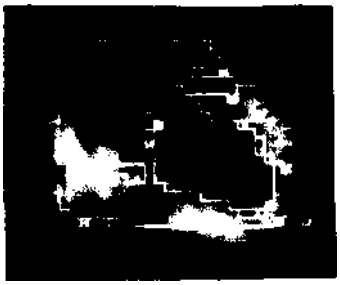


(B-3) Output of the semantic based region grower

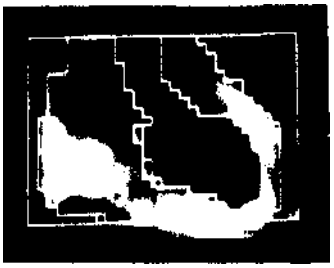
(B-l) Original picture.

(B-4) Result of grouping regions by their assigned meaning. Taking only regions which were assigned meaning w i th confi dence over 18 to be niergable.

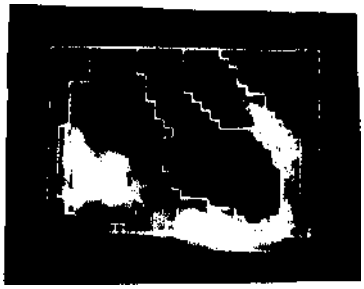
(B-5! Grouping regions by their assigned meaning, all regions considered mergable.



(F-I) Left ventricular angiogram. Output of the non-semantics weakest boundary first region grower. The stopping criterion is to stop when the merger gets down to two hundred regions.



(F-2-3-4) Iterations of semantic region grower. The region grower used is grouping of all adjacent regions which are assigned the same meaning by the sequential assignment procedure, before the first assignment with low confidence level occurs. On each iteration the confidence threshold is lowered.



(F-S) Final output . The heart interior is the dark center, around it is the chest cavity and on the two sides there is the dark frame border.