Technical Report CS 79007-R

ONE-DIMENSIONAL EDGE DETECTION AND REPRESENTATION

by

Roger W. Ehrich

Department of Computer Science
Virginia Polytechnic Institute and
State University
Blacksburg, Virginia 24061

September, 1979

Abstract

One reason no single edge detector has been found that satisfies the requirements of a wide range of image analysis applications is that existing edge detectors often produce unjustifiable interpretations of edge regions without sufficient contextual or external information. One-dimensional edge semantics are examined in detail, and a one-dimensional edge detector is proposed whose output is not just a local edge point assertion but a data structure. This data structure contains alternative interpretations of edge crossections as global gray level transition regions of variable widths. The detector is parameter-free, and a set of descriptive edge features is proposed.

This work was supported by the National Science Foundation under Grant ENG 7712105.

1 - Introduction

Many different edge models have been proposed as a basis for edge detection algerithms, and yet, it is uncommon to find thorough discussions of these models and their suitability for the types of images to which they are applied. The need for different edge detectors for different images may be an indication that our understanding of edge detection is still inadequate. One source of difficulty may be due to attempts to use gray level edge detectors to locate boundaries between textured regions as well. As investigators at MIT noted many years ago [1], edge semantics are extremely complicated, and it is understandable that much effort has been devoted to defining simplified models that are sufficiently realistic to produce good experimental results.

The main point of this paper is to show that edges in general are so complicated that it is beyond the capability of a single edge detector, based upon local measurements, to interpret an edge. An alternative edge detector is proposed that analyzes in detail all the intensity variations across an edge and produces for each edge slice a relational data structure called a relational tree. Each relational tree contains the alternative interpretations of an edge slice. While the trees of most edges are trivial, the general tree structure is necessary for more intelligent global edge constructors to make sense out of local events. The approach in this paper is one-dimensional, and there are three important reasons for this. First, the most important interpretations of an edge are a consequence of its behavior in a direction normal to its spatial orientation; discussion of two-dimensional behavior complicates what is already an enormously intricate structure. Second, the representation of an edge by the relational trees of its crossectional slices is both simple and useful. Third, many existing edge detectors are not, in fact, as two-dimensional as they first appear.

The most common technique for producing an edge image from a gray level image is to apply to the image a particular edge detector and then to do some sort of locally adaptive thresholding on the detector output. Milgram [2] gives an excellent discussion of the pitfalls of thresholding techniques and has proposed a boundary growing technique that produces region boundaries by examining the derivative image at multiple thresholds. In the technique described here there is no need to preselect thresholds since the various edge interpretations produced by the edge detector are self-determining. Second, it is argued that only first order difference approximations to the first derivative are viable for accurate, high resolution edge detection. Indeed, differences of 4 x 4 averages would not be suitable for the class of images with which we have been working.

Many edge detectors in common use, though two-dimensional in concept and design, have implementations that are nearly one-dimensional. An edge detector is said to be <u>directionally separable</u> if its output is a function $f(e_1, \ldots, e_n)$ of the responses $e_1, \ldots, e_n$ of a set of directionally selective detectors that test edge hypotheses in directions $\theta_1, \ldots, \theta_n$. An edge detector is <u>one-dimensional</u> if its output is computed from a one-dimensional intensity profile across an image, and it is called <u>quasi one-dimensional</u> if its output is computed from a linear combination of parallel intensity profiles. To give a few examples, the Hueckel operator [3] is an example of a non-separable edge detector, whereas the Roberts [4], Sobel [5], Kirsch [6], and Rosenfeld difference of averages [7] operators are examples of directionally separable edge detectors. For these four, the component detectors are one-dimensional in the case of the Roberts Cross, quasi one-dimensional in the case of the Rosenfeld and Sobel operators, and two-dimensional for the Kirsch.

Because many of the commonly used edge detectors are directionally separable and constructed from components that are at least quasi one-dimensional, it is the purpose of this paper to reexamine in more detail the semantics of one-dimensional edge profiles. There are a number of consequences of this view of the edge detection problem that are quite different from established views. On the other hand, it will be possible to see quite clearly why such a simple detector as the Roberts Cross performs as well as it does and why non-maximum suppression is so successful in thinning edges. Also, it should become clear why for any moderately complicated scene, no particular local edge detector with given parameters can do an adequate job for all images.

One of the most important assumptions in this paper is that the images under consideration are totally noise free and that the extreme complexity of a picture function is due entirely to tonal, reflectivity, illumination, and curvature variations in the three-dimensional scene being analyzed. It is, for example, quite clear that sensor noise is not the source of the great difficulty of general computer vision or robotics. In the same spirit, if there are dust specks on the light table used for digitizing an industrial scene, it is not to be the function of the edge detector to remove those specks. Most problems in these fields are consequences of the complexity of the world rather than the quality of the sensor system.

2 - Edge Regions

If one carefully considers a natural or industrial scene it is clear that edges are, by definition, regions of intensity change rather than locations at which image intensities change abruptly. In some applications it is necessary to know only a representative point for an edge element that can be used to form a smooth region boundary. In others, such as determining the curvature of an object, it is necessary to know the width of the edge region or possibly even the exact shape of the entire edge region. In any case, many edges, such as the one whose profile is shown

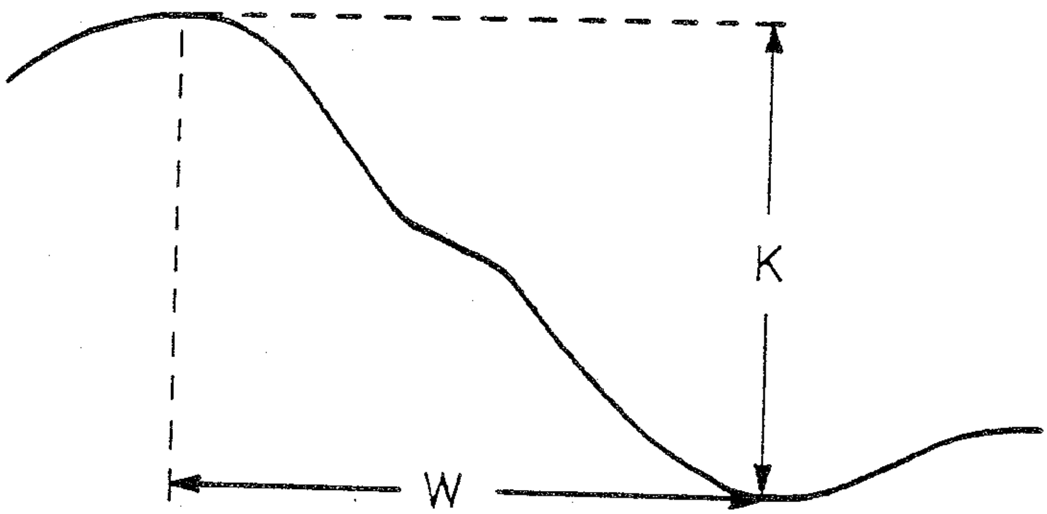

Figure 1 - Ambiguous Edge.

in Figure 1, are ambiguous. Given such a profile, one can always find an image in which the entire slope region must be interpreted globally as one large high contrast edge. In another image the very same profile may have to be interpreted on the basis of its context as two adjacent low contrast edges, separated by a narrow plateau.

Consequently, there is no way that an edge detector that produces the single edge interpretation can perform properly on the second image, and so forth. In fact, neither the thresholded Kirsch nor the thresholded Sobel operators are even capable of resolving the double edge in the center of this shadow plateau. However, this is a relatively unimportant point compared with the problem of the commit ment such a detector has to one interpretation or another. The only alternative is to construct an edge detector that is smart enough to report both interpretations so that the proper one may be chosen when more information is available. Before leaving Figure 1, notice that the edge contrast has little to do with local slope but is a global property of the entire edge region.

Unless one is analyzing very restricted scenes, it is important that an edge detector report only what is happening in an image rather than an interpretation of the data. This is an instance of Marr's Principle of Least Commit ment [8] which states roughly that whenever there is insufficient evidence for making a decision, the decision is deferred. Since the goal of edge detection is usually boundary formation or segmentation, the interpretation of edge profiles should be deferred until the time of the boundary formation process itself.

Least commit ment is the key to understanding the deficiencies of fixed size local edge operators. Most such operators are matched filters which measure the crosscorrelation between an image and a filter mask. In such an operator the edge model is embedded in the filter mask, and the detector output is a measure of the degree to which the data can be interpreted to be identical with the shape of the filter mask. Marr also recognized the problems associated with the use of a fixed size edge operator; instead he used a set of filter masks with varying sizes to determine the edge width by finding the detector that produced the largest output. Still, Marr would select one of a small number of interpretations for each edge, and to some extent he appears to violate his own principle.

The information required to be reported by an edge detector is determined in large part by the application. If one wishes to deduce the shape of an object in a scene, a functional approximation of a slope region might be appropriate. Other applications require much less information, and the edge features considered most important in this paper are the contrast and the locations of the maximum slope point, full edge region boundaries, and high derivative region boundaries.

In order to motivate the presentation of the edge detection algorithm proposed in this paper, let us consider some image data. The house scene in Figure 2 is a moderately complex image that contains both gray level edges and texture region boundaries. This image will be used for demonstrations throughout the remainder of the paper. Notice, in particular, the small circled region above the left window. A 25 point segment of a vertical profile through this region oriented from the top toward the bottom is shown in Figure 3a. The proper interpretation of this profile would be two parallel edges, separated by the light shadow region above the window. The shadow boundary is not the important edge, however, although that is obvious only because we know about shadows and the fact that they are derived rather than intrinsic parts of the scene.

Clearly the main function of an edge detector is detection of the presence of local slope, and Figures 3b and 3c show the first forward difference and central difference approximations to the first derivative, respectively. The plateau shows dramatically on Figure 3b and is totally missed on Figure 3c because it is only one picture element wide. Consequently it is necessary to use first forward differences if we are to resolve plateaus that are this narrow. Curiously, the edge detector with which most others have been compared over the years, the Roberts Cross, is made from two detectors that difference the intensities of diagonally adjacent picture elements. Typically one would threshold the output of the Roberts Cross at some intensity

Figure 2 - House Scene.

whose value is just above the valley between the two adjacent peaks on Figure 3b, and the two edge elements would be resolved. The problem, of course, is that there is no way of knowing in advance what threshold to use for this particular edge.

Looking again at Figure 3b it is obvious that the overall edge regions should be bounded by the intersection points of the derivative peaks and a line parallel to the horizontal axis at a height equal to the detector threshold. In order to bound the high slope region one might use the

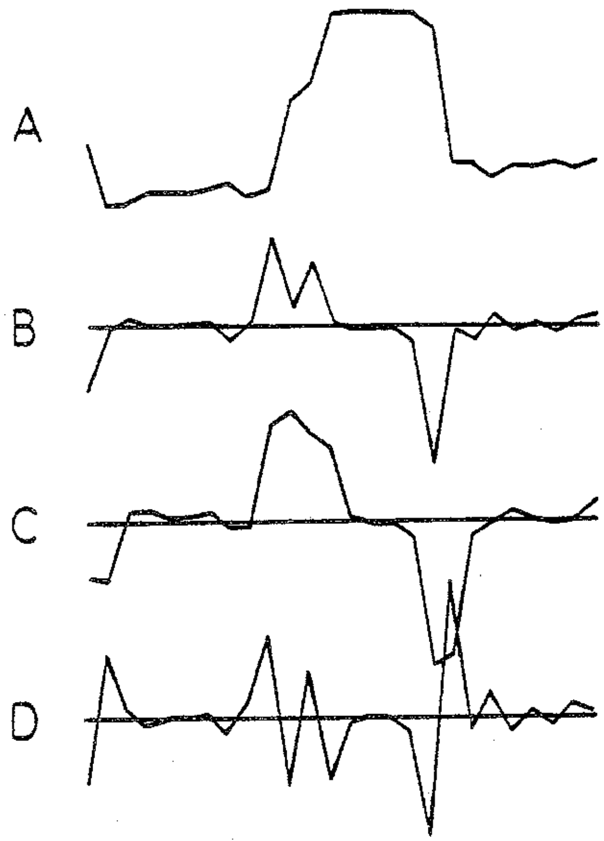

Figure 3 - Vertical profile from part of Figure 2 (a), First difference (b), Central Difference (c), Laplacian (d).

extrema of the second derivative shown in Figure 3d. Notice that there are two positive-negative peak pairs - one pair for each of the two adjacent edges. It was thought at first that the centers of the high slope regions would be excellent places to mark in order to produce thin and visually smooth edge boundaries. In all our experiments with various definitions of the high slope region, we were never able to make that idea work. The best place to mark is invariably the maximum slope point. This explains why the Roberts Cross produces such visually pleasing results, and it also justifies non-maximum suppression procedures for thinning edges. With that motivation, the next section defines precisely the edge detector that is being proposed.

3 - 1-D Edge Detector

An edge region is defined as a maximal set of consecutive points in an intensity profile whose slopes exceed a threshold T. Let D_1 be the derivative operator that differences the intensities of adjacent picture elements. If an intensity profile is given by

$$X = x_1, x_2, \ldots, x_{n-1}, x_n$$

then

$$D_1 X = x_2 - x_1, x_2 - x_1, x_3 - x_2, \ldots, x_n - x_{n-1}$$
$$= \dot{x}_1, \dot{x}_2, \ldots, \dot{x}_n$$

are the first forward differences of X. Since x_0 is unavailable for computing $\dot{x}_1$, $\dot{x}_1$ is simply set equal to $\dot{x}_2$. Then

$x_j, x_{j+1}, \ldots, x_{k-1}, x_k$ is a wide edge region

iff $\dot{x}_{j+1}, \dot{x}_{j+2}, \ldots, \dot{x}_{k-2}, \dot{x}_{k-1} \geq T$ and

$\dot{x}_j < T$ and $\dot{x}_k < T$.

The contrast of an edge is the signed difference between the extrema of X over the wide edge region.

Let

$$x_\ell = \max \{x_j, \ldots, x_k\}, \quad j \leq \ell \leq k$$

and let

$$x_m = \min \{x_j, \ldots, x_k\}, \quad j \leq m \leq k.$$

Then the contrast, K, is given by

$$K = x_{\ell} - x_m \quad \text{if } m < \ell$$
$$= x_m - x_{\ell} \quad \text{if } m > \ell$$
$$= 0 \quad \text{otherwise.}$$

The contrast is defined in this way rather than by $x_{k-1} - x_{j+1}$ so that the definition still works properly when another difference operator is substituted for D_1. The high derivative edge region is determined from the second derivative of the intensity profile X, and it is referred to as the <u>narrow edge region</u>. Let D_2 be the Laplacian operator defined by

$$D_2(x_i) = x_{i-1} - 2x_i + x_{i+1}.$$

Then

$$D_2X = x_1 - 2x_2 + x_3,\ x_1 - 2x_2 + x_3, \ldots,\ x_{n-2} - 2x_{n-1} + x_n$$
$$= \ddot{x}_1,\ \ddot{x}_2, \ldots,\ \ddot{x}_{n-1},\ \ddot{x}_n.$$

Once again since x_0 and x_{n+1} are unavailable, $\ddot{x}_1$ is set equal to $\ddot{x}_2$ and $\ddot{x}_n$ is set equal to $\ddot{x}_{n-1}$. The extrema of D_2X are the points where the maximal changes are occurring in the slope of X, and as discussed previously, these points seem natural for bounding the narrow edge region. Within the wide edge region, D_2X frequently has more than two extrema, and there needs to be a consistent way to select the proper ones. The situation is shown in Figure 3 if T=1 so that the wide edge region consists of the combination of the two subpeaks. To do this, the maximum slope point is determined, and then D_2X is searched for extrema on either side of the maximum slope point. Since it may happen that D_1

X has several maximum slope points within the wide edge region, the maximum slope point, s, is defined as follows.

Let $\dot{x}_i = \max\{\dot{x}_j, \ldots, \dot{x}_k\}$, $j \leq i \leq k$, if $K>0$
and $\dot{x}_i = \min\{\dot{x}_j, \ldots, \dot{x}_k\}$, $j \leq i \leq k$, if $K<0$
Then $s = (\ell+m)/2$ where

$\ell = \{p \mid \dot{x}_p \in \{\dot{x}_j, \ldots, \dot{x}_k\}, \dot{x}_p = \dot{x}_i$, and p is minimum$\}$
and $m = \{p \mid \dot{x}_p \in \{\dot{x}_j, \ldots, \dot{x}_k\}, \dot{x}_p = \dot{x}_i$, and p is maximum$\}$.

Thus s is determined by finding the extreme value of $D_1 X$ within the wide edge region and searching from both ends of the wide edge region for the first points that achieve that value. Next the narrow edge region is located by searching in either direction from the maximum slope point, s, for the first values of the second derivative that achieve the proper extreme values. First,

let $\ddot{x}_\ell = \max\{\ddot{x}_j, \ldots, \ddot{x}_s\}$
and $\ddot{x}_m = \min\{\ddot{x}_s, \ldots, \ddot{x}_k\}$ if $K>0$
or $\ddot{x}_\ell = \min\{\ddot{x}_j, \ldots, \ddot{x}_s\}$
and $\ddot{x}_m = \max\{\ddot{x}_s, \ldots, \ddot{x}_k\}$ if $K<0$.
Then the narrow edge region is given by

$x_r, \ldots, x_t$ where

$r = \{p \mid \ddot{x}_p \in \{\ddot{x}_j, \ldots, \ddot{x}_s\}, \ddot{x}_p = \ddot{x}_\ell$, and p is maximum$\}$
and $t = \{p \mid \ddot{x}_p \in \{\ddot{x}_s, \ldots, \ddot{x}_k\}, \ddot{x}_p = \ddot{x}_\ell$, and p is minimum$\}$.

At this point it still appears that there is a threshold, T, that must be selected before the algorithm above is run. This is a clear disadvantage since it fixes the detector's interpretation of a low slope region and causes exactly the problems we have sought to avoid. In the next section it is shown how T can be eliminated by allowing the edge detector to generate alternative hypotheses about the edges.

4 - Edge Hypotheses and Representations

Consider again the ambiguous edge of Figure 3a which has been redrawn together with the first derivative in Figure 4. As long as the edge profile is monotonic with polarity P it is necessary to form an edge hypothesis for the region over which the profile is monotonic. If there are two adjacent monotonic regions of polarity P separated by a monotonic region with polarity -P, then one must hypothesize three edge regions and require a more intelligent edge analyzer to determine whether the central edge region was caused by an undesired disturbance or by an anticipated and significant part of the scene.

The derivative in Figure 4b shows that the monotonic edge region in Figure 4a has substantial structure. With threshold T as shown, only one edge region would be reported by the edge detector, and with T set at T', two separate edge regions would be reported. The alternative of fixing

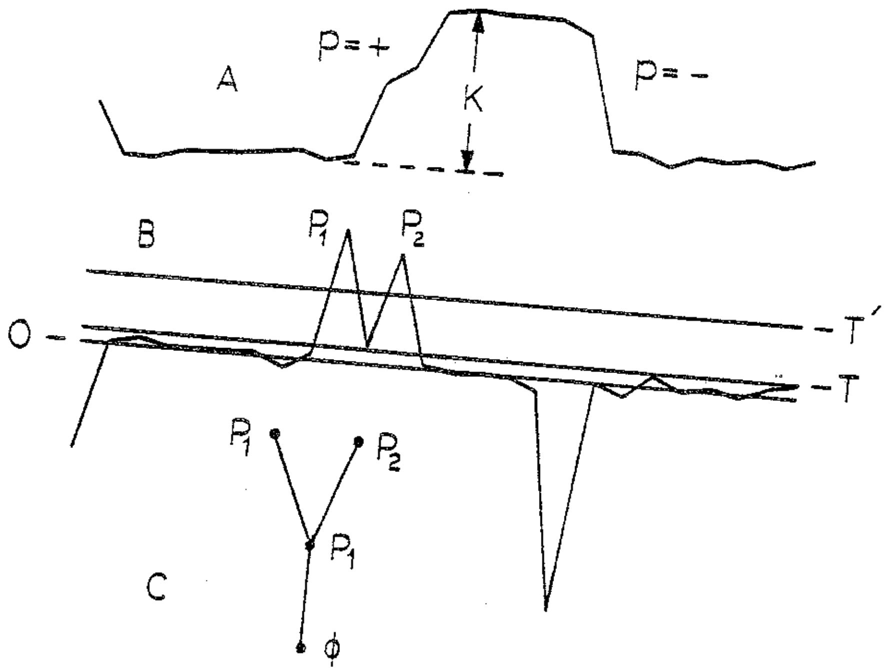

Figure 4 - Profile (a) and Derivative (b) from Figure 3 and Relational tree of edge (c).

the detector threshold at some high value so that multiple hypotheses are always produced is not viable because then high contrast, low slope edges would become undetectable. Therefore it is necessary to vary the threshold T to produce all possible hypotheses, and the tree data structure that is used to store these edge hypotheses is called a <u>relational tree</u>.

The relational tree of an edge can be easily motivated with the help of the example in Figure 4. Suppose T is set to some large value and then lowered in discrete steps to T=0. First peak P_1 will exceed the threshold and then P_2

will as well. As soon as T is lowered below the level of the valley V_1, peaks P_1 and P_2 merge to form a superpeak that is labeled by P_1 since it is the higher or dominant peak of the two that have merged. The merger of small peaks to form larger and larger superpeaks is represented naturally by a tree such as the one shown in Figure 4c, and this tree is called a relational tree because it represents relations among relations. Each valley induces a node in the relational tree, and each node is labeled with the dominant descendant peak. In general, relational trees are computed by a multi-stack algorithm, and the full development is given in [9]. The principal difference is that in this application a separate relational tree is computed from the first derivative for each monotonic portion of the profiles of a scene.

Each node of a relational tree consists of a pointer to a list of descriptive attributes for the peak it represents. In this case the attribute list would contain the following information (See Figure 1).

1 - Coordinates of the peak, giving maximum slope point and value of maximum slope.

2 - Horizontal coordinates of the wide edge region, W, defined with T at the valley that induced the peak.

3 - Contrast, K, for the peak, measured over the wide edge region.

4 - Horizontal coordinates of the narrow edge region.

The following properties of the relational trees of edges show why they are so useful in representing alternative hypotheses.

__Property 1__ The attribute list for each edge hypothesis contains exactly the right descriptive information about edge attributes as discussed earlier.

__Property 2__ As a relational tree is traversed upward from the root, one encounters finer and finer interpretations of an edge region.

__Property 3__ If one prunes a relational tree by removing all descendants of various nodes, the frontier of the pruned tree always represents adjacent, non-overlapping edge hypotheses. The description is complete in that all parts of the wide edge region of the root node are covered by the frontier nodes.

__Property 4__ Since most edge regions are very narrow, the relational trees of edges are usually very simple.

__Property 5__ The edge hypotheses represented by the nodes of the relational tree are determined by the depths of the valleys of the derivative, hence by the slopes connecting the high slope regions of the edge. For example, in Figure 5, there is no hypothesis P_2 P_3 because the slope between them is so low. Even if P_3 were higher than P_1, the structure of the relational tree would be the same. This property ensures that the hypotheses represented in the relational tree are also the best ones for structural reasons. On the other hand, there is enough information in the relational tree to form a compound hypothesis like P_2P_3, although the contrast will be slightly underestimated if the contrast is computed from information stored elsewhere in

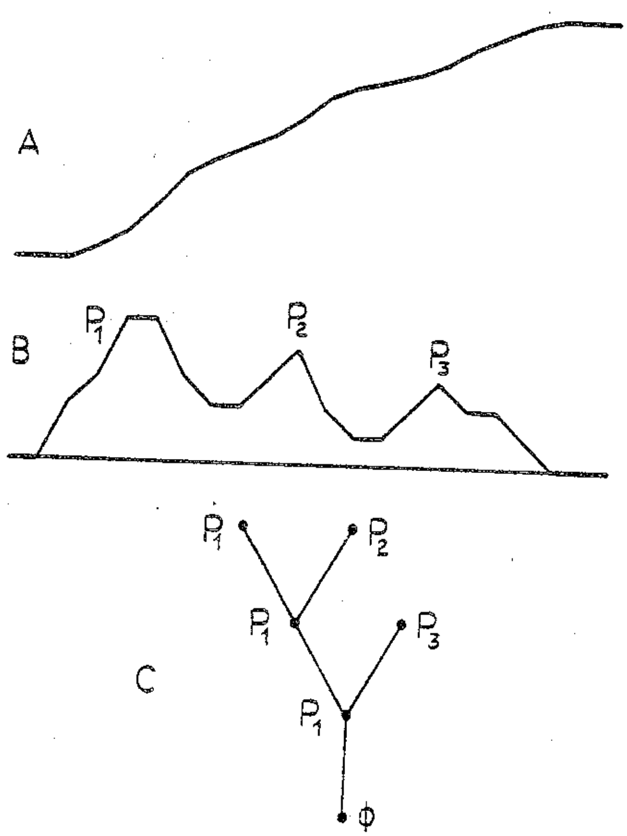

Figure 5 - Part of an intensity profile (a), First derivative (b), and Relational tree (c).

the tree.

In the preceding development, relational trees were used to represent only alternative hypotheses about edges that were monotonic slope regions. Figure 6 shows a situation commonly found in low slope shading edges. In Figure 6a there are three monotonic regions, and three relational trees would be generated since the slope of the center region has opposite polarity. If the wide edge region for the center segment was sufficiently narrow and the slope sufficiently low, a relational tree for the combined regions could easily be generated as illustrated. While it seems more consistent to require the subsequent edge constructor to form compound hypotheses from non-monotonic segments, the example shows how easy it is to have the edge detector do the work directly. Notice, however, that no explicit record of the center slope region is stored in the tree.

5 - Experimental Results

The following results were computed from the house scene in Figure 2. In viewing these results it is necessary to keep in mind both the sensitivity of the edge detector and the vast amount of detailed information it produces. The edge constructor, to be described in another paper [10], produces linked edges and a cleaned image that is as good as any result we have have seen for this image. We have superimposed the results for both vertical and horizontal processing and have plotted only the high slope point of each edge. Wherever edges appear thickened, it is a

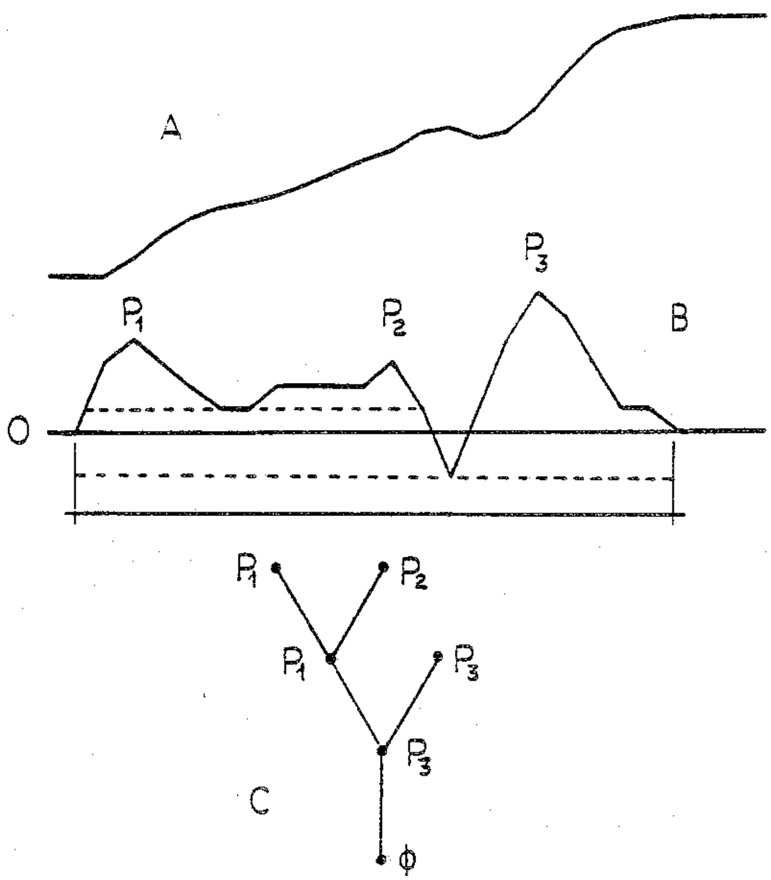

Figure 6 - Part of an intensity profile (a), First derivative (b), and Relational tree for merged edge regions.

consequence of detection of adjacent edges of alternating polarity.

Earlier in the paper it was mentioned that the high slope point was the best location for an edge. Several unsuccessful attempts were made to define a narrow edge region that would be centered on a good location point. In one experiment the wide edge region was searched from both ends for extrema of the second derivative, and the edge location point was taken to be the center of this region. Curiously, the high slope point was invariably within this region, but rarely at its center. The significant point is that we were unable to improve on the use of the high slope point as an edge locator.

First, let us compare the results of substituting a central difference operator for the forward difference operator, D_f. Figure 7 shows all edge points with contrast K>30 and T=1. While at first glance the central difference detector appears to produce a clearer image, this is because many significant edge points have been missed, such as the multiple edges on the roof line and on several of the windows. In Figure 8 are the detector responses with K>30

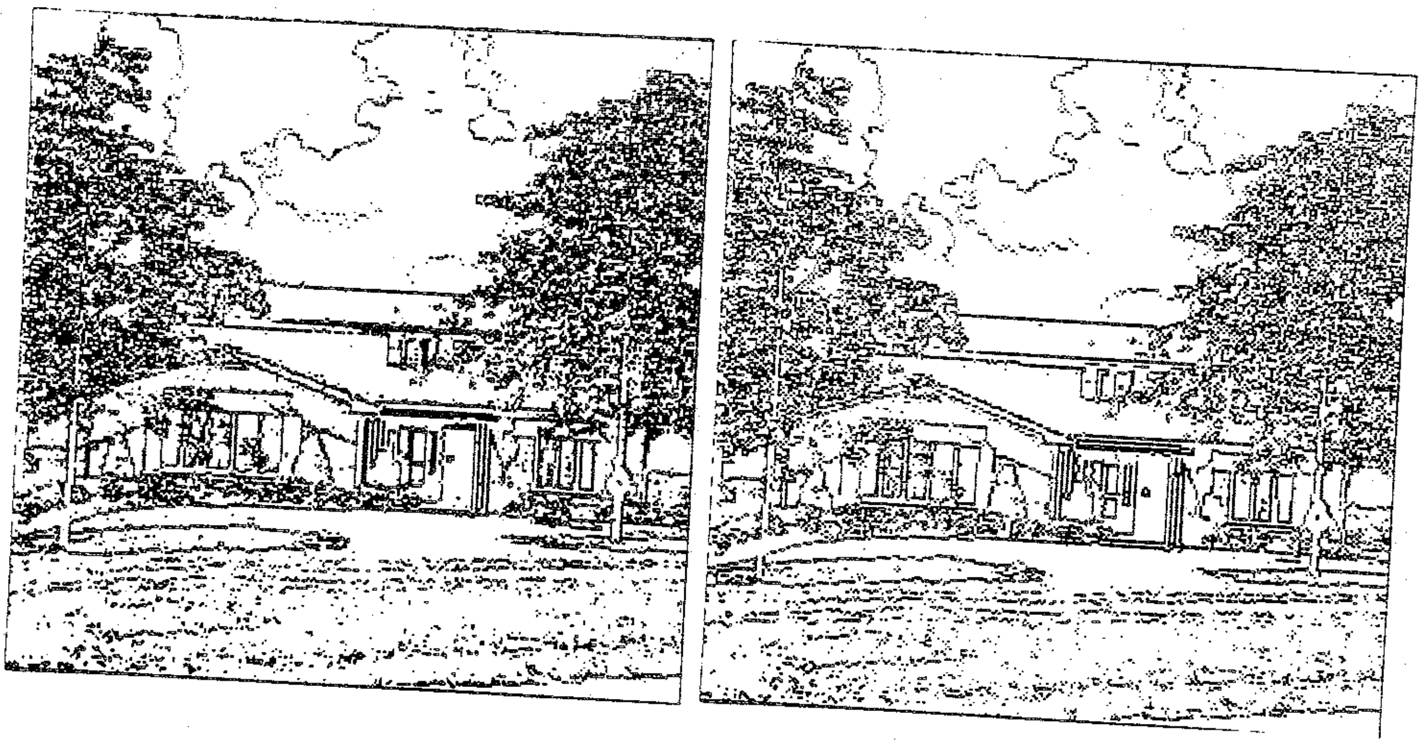

Figure 7 - K>30, T=1, Forward differences (a), and Central differences (b).

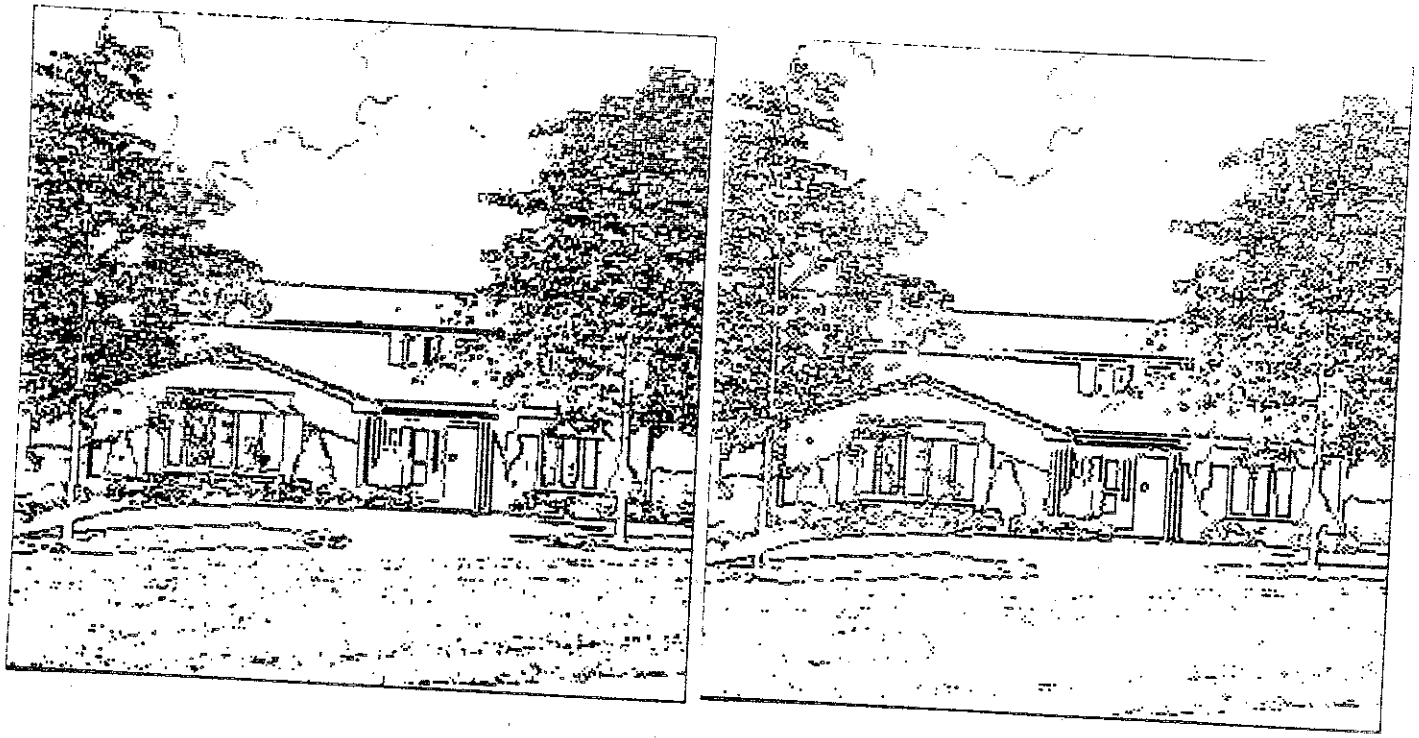

Figure 8 - K>30, T=20, Forward differences (a), and Central differences (b).

using T=20. Again, many edges have been missed by the central difference detector, although edge location seems to be slightly better. In the area above the left hand windows the shadow boundary cannot be completely detected by the central difference detector. A very careful comparison of Figures 7 and 8 reveals that although far fewer edges were detected with T=20 than with T=0, in many cases edges appear in Figure 8 where they were absent in Figure 7. These are a consequence of the mapping of the alternative hypotheses. Finally, notice that the texture boundaries between sky and clouds are not well defined. It is firmly believed that special techniques should be used to detect such boundaries and that general edge detectors should not be used here.

6 - Conclusions

It is felt that the edge detector described in this paper is more suitable for many applications in image analysis than many currently in use because it is more sensitive, more accurate, and because it does not make irrevocably incorrect interpretations of edges. The proof, of course, is in whether it is possible to build an edge constructor that makes use of the advantages of the detector. What we have had in mind from the beginning is a contextual edge linker that functions by scan line matching, and that is the subject of a companion paper [10].

References

1. Horn, B.K.P., "The Binford-Horn linefinder," Artificial Intelligence Memo 285, MIT AI Laboratory, Cambridge, March 1973.

2. Milgram, D.L., "Region Extraction Using Convergent Evidence," Computer Graphics and Image Processing 11, January 1979, pp. 1-12.

3. Hueckel, M.H., "An operator which locates edges in digital pictures," J. ACM 18, January 1971, pp. 113-125.

4. Roberts, L.G., "Machine perception of three-dimensional solids," in Optical and Electrooptical Information Processing, J.T. Tippett et.al., eds., MIT Press, Cambridge, 1965, pp. 159-197.

5. Duda, R.O. and P.E. Hart, Pattern Classification and Scene Analysis, Wiley, New York, 1973, pp. 271-272.

6. Kirsch, R., "Computer Determination of the Constituent Structure of Biological Images," Computers and Biomedical Research 4, June 1971, pp. 315-328.

7. Rosenfeld, A. and M. Thurston, "Edge and Curve Detection for Visual Scene Analysis," IEEE Trans. Comp. 20, May 1971, pp. 562-569.

8. Marr, D., "Early Processing of Visual Information," Phil. Trans. Royal Soc. London 275 on Biological Sciences, 19 October 1976, pp. 483-524.

9. Ehrich, R.W. and J.P. Foith, "Representation of Random Waveforms by Relational Trees," IEEE Trans. Comp. 25, July 1976, pp.725-736.

10. Ehrich, R.W. and F.H. Schroeder, "Contextual Boundary Formation by Scan Line Matching," Technical Report CS79009-R, CS Department, Virginia Polytechnic Institute and State University, September 1979.

List of Captions

Figure 1 - Ambiguous edge.

Figure 2 - House scene.

Figure 3 - Vertical profile from part of Figure 2 (a), First difference (b), Central difference (c), Laplacian (d).

Figure 4 - Profile (a) and Derivative (b) from Figure 3 and Relational tree of edge (c).

Figure 5 - Part of an intensity profile (a), First derivative (b), and Relational tree (c).

Figure 6 - Part of an intensity profile (a), First derivative (b), and Relational tree for merged edge regions (c).

Figure 7 - K>30, T=1, Forward differences (a), and Central differences (b).

Figure 8 - K>30, T=20, Forward differences (a), and Central differences (b).