COMPUTER SCIENCE TECHNICAL REPORT SERIES

UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND
20742

1

UNDERSTANDING FEATURES, OBJECTS AND BACKGROUNDS

Semi-Annual Report

1 April 1980-31 January 1981

Contract DAAG-53-76C-0138

(DARPA Order 3206)

Computer Science Center
University of Maryland
College Park, MD 20742

ABSTRACT

Current activities on the project are reviewed under the following headings:

1) Segmentation

2) Local feature detection

3) Feature linking

4) Hierarchical representation

The help of Sherry Palmer in preparing this report is gratefully acknowledged.

1. Introduction

This project is concerned with the study of advanced techniques for the analysis of reconnaissance imagery. It is being conducted under Contract DAAG-53-76-C-0138 (DARPA Order 3206), monitored by the U.S. Army Night Vision and Electro-Optics Laboratory, Ft. Belvoir, VA (Dr. George Jones). The Westinghouse Systems Development Division, under a subcontract, is collaborating on implementation and application aspects.

The previous phase of the project, entitled "Image Understanding Using Overlays", was concluded during the past reporting period. Accomplishments under this phase are summarized in a Final Report dated May 1980 [1], which also contains a bibliography of all reports and papers produced during this period.

The current phase of the project is concerned with three principal areas: (a) comparative analysis of segmentation techniques applied to FLIR imagery; (b) development of an inference-based approach to target detection on FLIR imagery; and (c) optical flow analysis of time-varying imagery. Work in area (b) is in progress and will be described in forthcoming technical reports. Area (a) has emphasized methods based on hierarchical ("pyramid") image representations, some of which are reviewed in this report and in individual Technical Reports [2,3]. Other individual reports [4,5] deal with some of the work done in area (c). In addition, the project is preparing software contributions to the DARPA/DMA Image Understanding Testbed; the first of these will be a general-purpose software package for implementing relaxation processes at the pixel level.

This report reviews activities on the project during the period April 1980-January 1981. This work is covered under the headings of segmentation; local feature detection; feature linking; and hierarchical representation. The work is summarized only briefly, since it is covered in greater detail in individual technical reports and Image Understanding Workshop papers.

2. Segmentation

2.1 Color pixel classification

When pixels in a black-and-white image are classified by thresholding their gray levels, gradient magnitude information can be used in various ways as an aid in threshold selection. In particular, a histogram of the gray levels of pixels whose gradient magnitudes are low has sharper peaks and deeper valleys than the histogram of the entire image, since the low-gradient pixels tend to come from the interiors of regions, not from region border zones; it is easier to choose useful thresholds (at valley bottoms) from this improved histogram. Analogously, when pixels in a color or multispectral image are classified on the basis of their spectral signatures, the color gradient magnitude can be used as an aid in defining decision surfaces that separate clusters of pixels having like signatures. In fact, a scatterplot of the signatures of pixels whose color gradient magnitudes are low has more clearly separated clusters than the scatter plot of the entire image, for the same reason as in the grayscale case. This phenomenon is illustrated in Figure 1. For further details and additional examples, see [6].

2.2 Mosaicking

When aerial photographs are combined into a photomosaic, seams are often apparent between the parts. These seams are caused by gray level differences due to the different conditions under which the parts were recorded. A relaxation method has been developed that generates a gray level correction function such that, when this function is subtracted from the mosaic, the seams are eliminated, but the details of the photographs are not affected. The algorithm does not assume any specific types of gray level differences among the parts, nor does it require the existence of overlaps between the parts, and it can be used for arbitrary numbers of parts; but it does have the drawback that if a seam coincides with an edge between two regions, that edge will be eliminated. The algorithm constructs a seam-eliminating function which, when subtracted from the mosaic, causes the gray levels at pairs of adjacent points on opposite sides of a seam to become equal, and which otherwise is as smooth as possible. An example of mosaic seam elimination using this algorithm is shown in Figure 2. Other examples, and further details, can be found in [7].

3. Local Feature Detection

3.1 Higher-order edge detectors

One way to define edge detectors for digital images is to fit a polynomial surface to a neighborhood of each pixel, and take the magnitude of the gradient of that surface as an estimate of edgeness. The polynomial fitting process is usually carried out for symmetric neighborhoods, using polynomials of degree 1 or 2. Using least squares fitting by orthonormalization and Richardson extrapolation, one can calculate such edge estimates for other classes of neighborhoods, and for higher-order polynomial models [8].

As further application of this approach, edge detectors can be defined based on least squares surface fitting in which the surface is a step edge superimposed on a low-order polynomial function. This makes it possible to "filter" optimal step-based operator responses so as to discriminate against noise responses, by rejecting responses for which the fit is poor, without discriminating against low-contrast edges (which is unavoidable if thresholding is used for noise suppression). An example of such edge "filtering" is shown in Figure 3. For other examples, and further details, see [9].

3.2 Edge evaluation

A method of evaluating edge detector output has been developed, based on the local good form of the detected edges. It combines two desirable qualities of well-formed edges -- good continuation and thinness. The measure has the expected behavior for known input edges as a function of their blur and noise. It yields results generally similar to those obtained with measures based on discrepancy of the detected edges from their known ideal positions, but it has the advantage of not requiring ideal positions to be known. It can be used as an aid to threshold selection in edge detection (pick the threshold that maximizes the measure), as a basis for comparing the performances of different detectors, and as a measure of the effectiveness of various types of preprocessing operations facilitating edge detection. This method is described in detail in a separate Technical Report, where examples of its performance are also given [10].

4. Feature Linking

4.1 Edge segment linking

A system of programs that links edge segments based on both gray level and geometric criteria has been developed and applied to the detection of buildings and roads on aerial photographs. Preliminary results using these programs were described in [11]; a more detailed description, and numerous additional results, are presented in [12]. Further work along these lines led to the development of figures of merit for linking compatible segments (i.e., segments that could be consecutive sides of an object) and antiparallel segments (i.e., segments that could be opposite sides). For compatible pairs, the figure of merit is based on the geometrical configuration of the segments, the similarity of the gray levels on their "object" sides, and the similarity between their object sides and the line joining their endpoints. For antiparallel pairs, it is based on the homogeneity of gray level between the edges and the amount of overlap between them. These figures of merit have highly bimodal histograms, making it quite easy to decide which pairs of segments should be linked, as illustrated in Figures 4 and 5. They should be useful in the design of relaxation-like schemes for classifying edge segments. For further details, and many additional results, see [13,14].

4.2 Reconstruction from gray-weighted medial axes

A method of defining a "min-max medial axis transformation" (MMMAT) for grayscale images, based on iterated local MIN and MAX operations, was described in a previous report [1]. This transformation associates with each pixel a vector of gray level increments, and exact reconstruction of the image is possible from these vectors. Moreover, good approximations to the image can be reconstructed using only the strongest components of the strongest few vectors. A few illustrations of this were given in an earlier report; further details and additional examples can be found in [15].

5. Hierarchical Representation

Extensive work has been done on this project on the use of pyramid and quadtree structures for image representation and processing. The work done in this area through March 1980 was summarized in [16]. In this section we briefly summarize developments in this area during the past reporting period.

5.1 Quadtree-to-raster conversion

An algorithm for converting quadtree representations of binary images to row-by-row (e.g., run-length) representations was described and partially analyzed in an earlier report. More recently, a comparative study and complete analysis of four such algorithms has been conducted [17]. The simplest algorithm is a straightforward top-down approach that visits each run in a row in succession starting at the root of the tree; the other algorithms proceed in a manner akin to an in-order tree traversal. The analysis shows under what circumstances each algorithm is preferable. They have all been shown to have execution times proportional to the sum of the heights of the blocks comprising the image.

5.2 Quadtree-based image smoothing

Two methods for smoothing an image using quadtree approximations to the image have been developed. One uses the sizes of the leaves in the quadtree to determine neighborhood sizes over which to apply the smoothing. The other method maps each image gray level i into the gray level j into which i most frequently maps when we replace the level of each pixel by the level of the quadtree leaf to which it belongs. Results obtained using these methods, as well as a local histogram peak sharpening method, are shown in Figure 6. The second quadtree-based method seems to give the best results. Additional examples, and detailed descriptions of the methods, can be found in [18].

5.3 Edge pyramids and quadtrees

An edge (or curve) pyramid is a sequence of successively lower-resolution versions of an image, each containing a summary of the edge information in its predecessor. This summary includes the average edge magnitude and direction in each "block" of the higher-resolution image, together with an intercept in that block and a measure of the error in the direction estimate. An edge quadtree, analogously, is a variable-resolution representation of the edge or curve information in the given image, constructed by recursively splitting the image into quadrants based on magnitude, direction, intercept, and error information. Advantages of these representations include their registration with the original image, their ability to represent many edges or curves in a single tree structure, and their ability to perform many operations on the represented data efficiency. A detailed description of these representations, together with examples, can be found in a separate Technical Report [2].

5.4 Pyramid linking

When an image is smoothed using small blocks or neighborhoods, the results may be somewhat unreliable due to the effects of noise on small samples. When larger blocks are used, the samples become more reliable, but they are more likely to be mixed, since a large block will often not be contained in a single region of the image. A compromise approach is to use several block sizes, representing versions of the image at several resolutions, and to carry out the smoothing by means of a cooperative process based on links between blocks of adjacent sizes. These links define "block trees" which segment the image into regions, not necessarily connected, over which smoothing takes place. The basic "pyramid linking" scheme was described in an earlier report. Further experiments with this scheme have led to some improvements over the original method, based on better ways of initializing the process and measuring the link merit. A detailed description of these experiments and their results can be found in a separate Technical Report [3]. It has also been found that forced-choice linking of blocks to larger blocks is not necessary; one can use weighted links, recomputing the weights at each iteration, and it turns out that the weights converge to 0's and 1's as the process stabilizes. Generalizations of this approach to image features other than gray level, including color signatures and textural properties, have also been investigated and will be described in future reports.

References

1. A. Rosenfeld, Image Understanding Using Overlays, Final Report, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, May 1980.

2. M. Shneier, Two Hierarchical Linear Feature Representations: Edge Pyramids and Edge Quadtrees, TR-961, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, October 1980; also in Proc. DARPA Image Understanding Workshop, April 1981.

3. T. Silberberg, S. Peleg, and A. Rosenfeld, Multi-Resolution Pixel Linking for Image Smoothing and Segmentation, TR-977, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, November 1980; also in Proc. DARPA Image Understanding Workshop, April 1981.

4. K. Prazdny, Relative Depth and Local Surface Orientation from Image Motions, TR-996, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, January 1981; also in Proc. DARPA Image Understanding Workshop, April 1981.

5. K. Prazdny, Determining the Instantaneous Direction of Motion from Optical Flow Generated by a Curvilinearly Moving Observer, TR-1009, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, February 1981; also in Proc. DARPA Image Understanding Workshop, April 1981.

6. A. Broder and A. Rosenfeld, Gradient Magnitude as an Aid in Color Pixel Classification, TR-906, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, June 1980.

7. S. Peleg, Elimination of Seams from Photomosaics, TR-895, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, April 1980.

8. D.G. Morgenthaler, Feature Detectors Based on Higher Order Polynomials, TR-896, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, April 1980.

9. D.G. Morgenthaler, A New Hybrid Edge Detector, TR-897, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, April 1980.

10. L. Kitchen and A. Rosenfeld, Edge Evaluation Using Local Edge Coherence, TR-981, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, December 1980; also in Proc. DARPA Image Understanding Workshop, April 1981.

11. M. Tavakoli, Toward the Recognition of Cultural Features, Proc. DARPA Image Understanding Workshop, April 1980, 33-57.

12. M. Tavakoli and A. Rosenfeld, Toward the Recognition of Buildings and Roads on Aerial Photographs, TR-913, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, July 1980.

13. M. Tavakoli and A. Rosenfeld, A Method for Linking Pairs of Compatible Linear Features, TR-930, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, August 1980.

14. M. Tavakoli and A. Rosenfeld, A Method for Finding Pairs of Anti-Parallel Linear Features, TR-943, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, September 1980.

15. S. Wang, A.Y. Wu, and A. Rosenfeld, Image Approximation from Grayscale "Medial Axes", TR-900, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, May 1980.

16. A. Rosenfeld, Some Uses of Pyramids in Image Processing and Segmentation, Proc. DARPA Image Understanding Workshop, April 1980, 112-120.

17. H. Samet, Algorithms for the Conversion of Quadtrees in Rasters, TR-979, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, November 1980.

18. S. Ranade and M. Shneier, Using Quadtrees to Smooth Images, TR-894, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, April 1980.

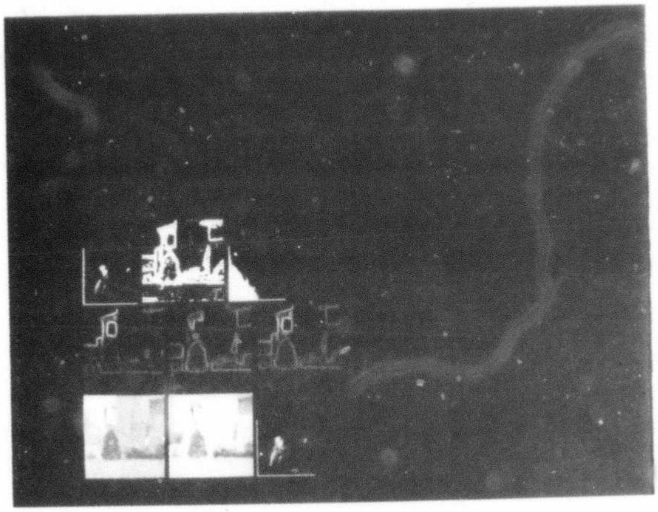
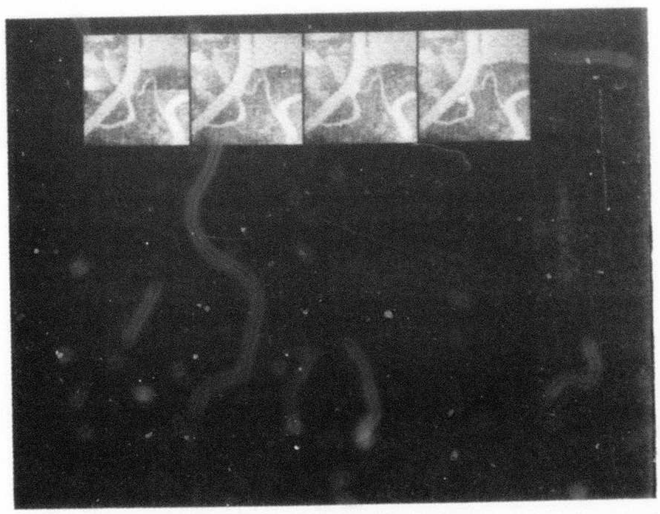


Figure 1. Scatterplot enhancement by suppression of high-gradient pixels.

key to parts:

g	h	i
d	e	f
a	b	c

a-b) Two bands
c) Scatterplot of (a) vs. (b), log scaled
d-e) Edge responses in the two bands (RMS Roberts operator)
f) Color edge response: RMS of (d) and (e)
g) Enhanced scatterplot, log scaled; pixels with edge responses > 2 have been suppressed
h) Mask showing suppressed pixels
i) Histogram of edge responses, log scaled



(a) (b) (c) (d)

Figure 2. Mosaic seam elimination by relaxation: (a) Original (b-d) After 50,100, and 200 iterations.

a)

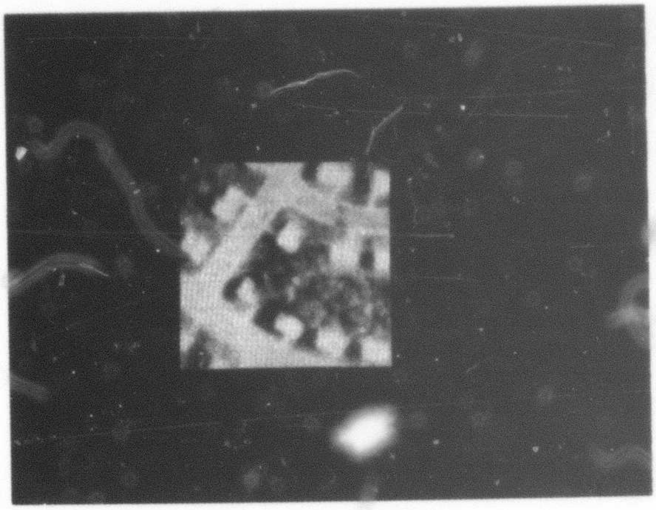
b)

c)

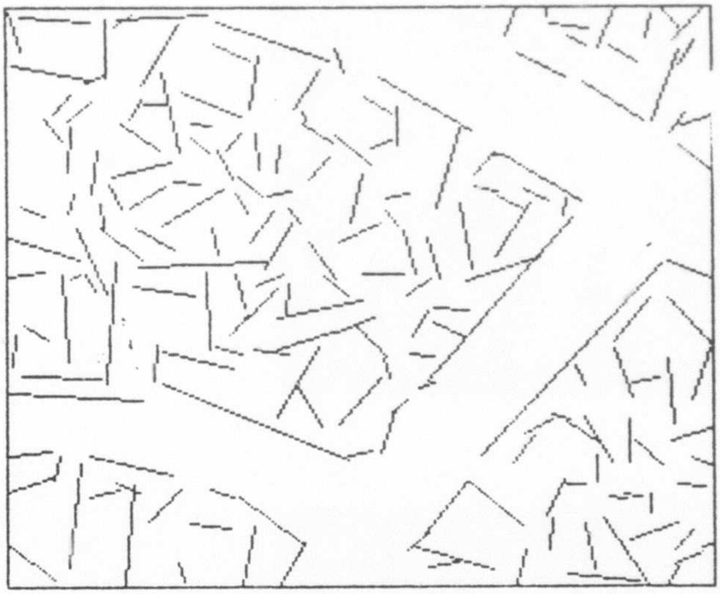
Figure 3. Suppression of low-fit edges.

a) Original (skull cross-section)

b-c) High-fit edges (two-thresholds). Note that the lower threshold brings out the low-contrast edges without affecting the noise edges outside the skull.



(a) Suburban scene



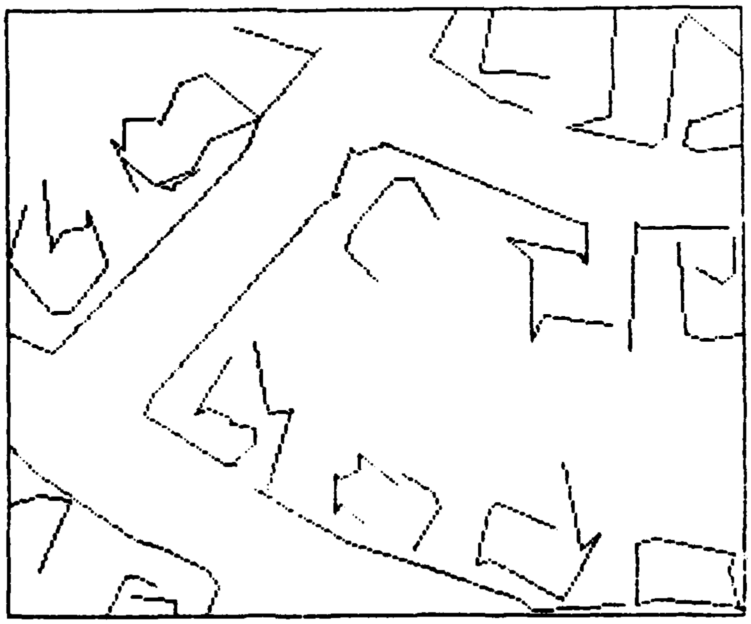
(b) Edge segments extracted from (a)

Figure 4. Compatibility merit for edge segments.

```
0----------10----------20----------30----------40----------50-----------100 --

0--*****************************************************************
.1-****
.2-******
.3-**
.4--******
.5-*****
.6-**
.7-***
.8-***
.9-******
1.0-***
1.1-
1.2-**
1.3-
1.4-**
1.5-
1.6-
1.7--**
1.8-***
1.9-**
2.0-**
2.1-*******
2.2-**
2.3-***
2.4-*******
2.5-*******
2.6--***********
2.7-**************
2.8-********
2.9--********************
```

(c) Compatibility merits for pairs of segments in (b)



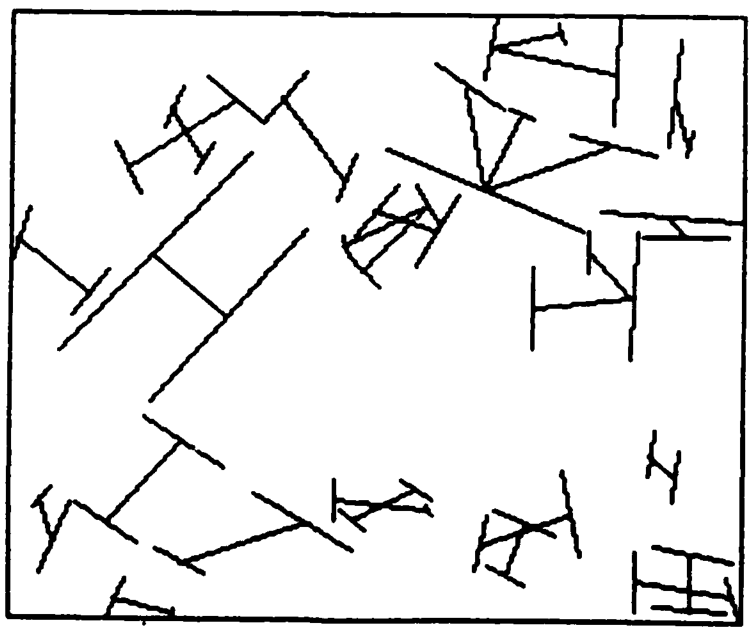
(d) Pairs of segments having compatibilities ≥ 2

Figure 4, cont'd.

```
0---------10---------20---------30---------40

0-*************************************
.1-***
.2-***
.3-****
.4-**
.5-**
.6-
.7-
.8-
.9-**
1.0-
1.1-**
1.2-
1.3-**
1.4-
1.5-**
1.6-***
1.7-**
1.8-
1.9-**
2.0-
2.1-*****
2.2-**
2.3-**
2.4-**
2.5-
2.6-
2.7-****
2.8-
2.9-**
```

(a) Antiparallelness merits for the segments in Figure 4b.



(b) Pairs of segments having antiparallelness merit ≥ 2.

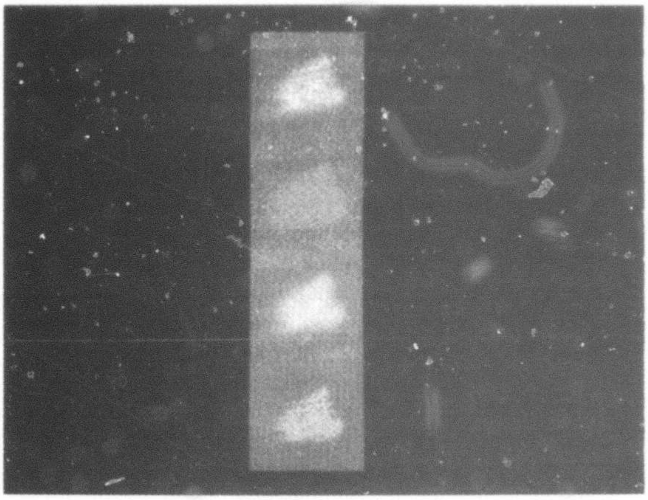
Figure 5. Antiparallelness merit for edge segments

1)

2)

3)

4)



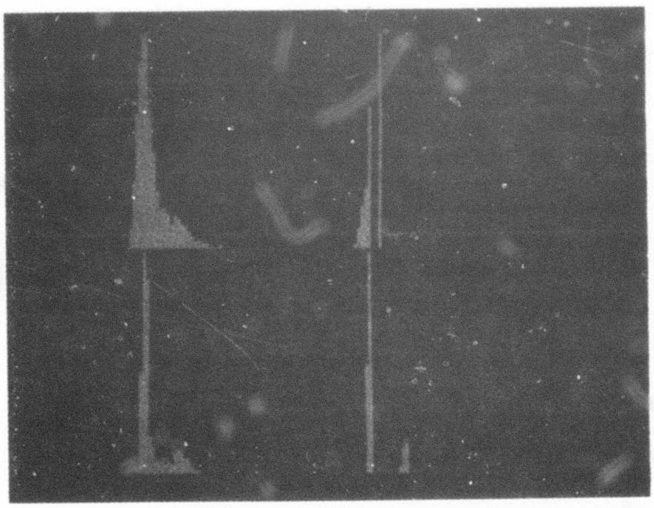
a1) Input image
2) Results of histogram sharpening
3) Results of variable-neighborhood smoothing
4) Results of smoothing using most frequent leaf value

1)

2)

3)

4)



b) Histograms of the images in (a)

Figure 6. Quadtree-based image smoothing

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)

UNDERSTANDING FEATURES, OBJECTS AND BACK-GROUNDS

5. TYPE OF REPORT & PERIOD COVERED

Semi-Annual Report
1 April 1980-31 January '81

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)

8. CONTRACT OR GRANT NUMBER(s)

(15) DAAG-53-76C-0138

DARPA Order-3206

9. PERFORMING ORGANIZATION NAME AND ADDRESS

Computer Science Center
University of Maryland
College Park, MD 20742

403018

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

(12) 42

11. CONTROLLING OFFICE NAME AND ADDRESS

12. REPORT DATE

13. NUMBER OF PAGES

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)

15. SECURITY CLASS. (of this report)

Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Image processing
Image understanding
Pattern recognition
Artifical intelligence

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Current activities on the project are reviewed under the following headings:

1. Segmentation
2. Local feature detection
3. Feature linking
4. Hierarchical representation