UNDERSTANDING FEATURES, OBJECTS, AND BACKGROUNDS

Project Status Report 1 August 1981 - 31 July 1982

Ariel Rosenfeld
Larry S. Davis

Computer Vision Laboratory, Computer Science Center
University of Maryland, College Park, MD 20742

ABSTRACT

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

(a) Preprocessing and segmentation,
(b) Feature detection and texture analysis,
(c) Hierarchical representations,
(d) Matching and motion,

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 320f), monitored by the U.S. Army Night Vision and Electro-Optics Laboratory (Dr. George Jones). The Westinghouse Systems Development Division, under a subcontract, is collaborating on implementation and application aspects.

Work on the current phase of the project was initiated in April 1980. Accomplishments and publications during the period 1 April 1980 - 31 July 1981 are summarized in two earlier status reports [1-2], the first of which also appeared in the Proceedings of the April 1981 Image Understanding Workshop [3]. The present report, covering the period 1 August 1981 - 31 July 1982, is being issued separately and will also appear in the Proceedings of the September 1982 Image Understanding Workshop. For convenience, publications since February 1981 are also cited here, since they were not cited in the April 1981 Workshop Proceedings.

The project is concerned with three principal areas: segmentation techniques; context-based target detection in FLIR imagery; and analysis of time-varying imagery. Work in the first area is summarized in Section 2 (Preprocessing and segmentation) and 3 (Feature detection and texture analysis). Section 4 summarizes work on the use of hierarchical image representations ("pyramids") in both segmentation and feature detection. Three papers in these areas, dealing with a comparative study of segmentation techniques as applied to FLIR imagery [4], and with the use of pyramids for extracting compact objects from an image [5,6], also appear in the Workshop Proceedings. Work on context-based target detection is covered in a report that also appears in the Workshop Proceedings [7]; a second report on this topic is in preparation. Finally, Section 5 summarizes work on image matching and time-varying imagery analysis; one paper in this area also appears in the Workshop Proceedings [8].

2. Preprocessing and segmentation

2.1 Comparative segmentation study

A comparative study of FLIR image segmentation techniques was conducted, using a database of 51 images obtained from four different sources. The techniques compared included two- and three-class relaxation, "pyramid linking", and "superspike" (see below). The results are described in detail [4] and in a paper appearing in the Workshop Proceedings.

2.2 Constraint-based region identification

A context-based approach to region identification on FLIR imagery was developed; it uses constraint filtering techniques to identify regions as (possibly) belonging to the classes sky, smoke, ground, tank, and tree. A detailed description of the approach and examples of its use can be found in [7], which also appears in the Workshop Proceedings.

2.3 Histogram-based image smoothing

A powerful method of edge-preserving image smoothing known as "superspike" has been developed. It is based on repeatedly averaging each pixel with a subset of its neighbors, where the neighbors used are chosen on the basis of their relationships with the given pixel on the image's histogram. Specifically, we use a neighbor if its value is more probable than the pixel's, and there is no concavity on the histogram between its value and the pixel's; these conditions imply that it belongs to the same histogram peak as the pixel, and is higher up on that peak. This method can also be applied to multi-spectral imagery, using the scattergram rather than the histogram [9]. Figure 1 shows an example of this type of smoothing applied to a color image of a house, using only two bands (red and blue). The result is...
of the pyramid based on their similarity in value. This method has been generalized to multispectral imagery, where better results can be obtained using two bands than using one band at a time. The details were given in [15] (also briefly summarized in [2]).

Pyramid linking methods can also be used to extract significant edges from an image, by creating links between nearby pairs of edge segments on consecutive levels based on similarity in slope. The details of this approach were given in [16] (also briefly summarized in [2]).

A more recent application of pyramid linking is to the detection and extraction of compact objects from an image using local "spike filters" on each level of the pyramid. This method is described in detail in [5], which also appears in the Workshop Proceedings.

Pyramid linking is usually based on forced choices, where a pixel must link to one of the nearby pixels on the level above it. A "softer" approach is to use weighted links (the more similar, the stronger). This too gives rise to trees whose roots are pixels that have only negligibly weighted links to the level above them. Typically, the leaves of a tree constitute a compact, homogeneous piece of the image. The approach is described in detail in [6], which also appears in the Workshop Proceedings.

5. Matching and motion

5.1 Corner-based image matching

Some experiments on relaxation image matching, based on "corner" features extracted from the images, were described in [17] (also briefly summarized in [2]). Further experiments, in which local gray level correlation was used to resolve ambiguous cases, are described in [18].

5.2 Corner-based motion computation

By computing (approximately) the spatial and temporal derivatives of the image gray level at a given pixel, the component of the velocity of that pixel in the gradient direction can be estimated. If the pixel is at a "corner" of an object, where edges having two different directions meet, its velocity is thus completely determined. When the velocities are due to observer motion ("optical flow"), knowing them at a few points suffices to determine the translation and rotational components of the flow [19]. When an object is moving, estimates of the velocities of its corners can be "propagated" along its contours to yield a consistent estimate of object motion [20,21]. Further details of this approach, together with examples, are presented in [8], which also appears in the Workshop Proceedings.
REFERENCES
1. Understanding Features, Objects and Backgrounds, Semi-Annual Report, 1 April 1980 – 31 January 1981, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD.

2. Understanding Features, Objects and Backgrounds, Semi-Annual Report, 1 February – 31 July 1981, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD.


5. T. H. Hong and M. Shneier, Extracting compact objects using linked pyramids, TR-1123, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, November 1981 (also in the Workshop Proceedings).

6. T. H. Hong and A. Rosenfeld, Unforced image partitioning by weighted pyramid linking, TR-1137, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, January 1982 (also in the Workshop Proceedings).

7. L. Kitchen, Scene analysis using region-based constraint filtering, TR-1150, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, February 1982 (also in the Workshop Proceedings).

8. L. S. Davis, Z. Q. Wu, and H. F. Sun, Contour-based motion estimation, TR-1179, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, June 1982 (also in the Workshop Proceedings).


10. S. Dunn, L. Janos, and A. Rosenfeld, Bimean clustering, TR-1106, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, September 1981.

11. R. B. Boppana and A. Rosenfeld, Some properties of Hückel-type edge operators, TR-1178, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, June 1982.

12. Z. Q. Wu and A. Rosenfeld, Filtered projections as an aid in corner detection, TR-1078, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, July 1981.

13. M. Pietikäinen, A. Rosenfeld, and L. S. Davis, Texture classification using averages of local pattern matches, TR-1098, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, September 1981.

14. C. Y. Wang and A. Rosenfeld, Elevation texture, TR-1085, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, August 1981.

15. T. H. Hong and A. Rosenfeld, Multiband pyramid linking, TR-1025, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, March 1981.

16. T. H. Hong, M. Shneier, and A. Rosenfeld, Border extraction using linked edge pyramids, TR-1080, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, July 1981.

17. C. Y. Wang, H. F. Sun, S. Yada, and A. Rosenfeld, Some experiments in relaxation image matching using corner features, TR-1071, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, July 1981.

18. H. F. Sun, Image registration by combining feature matching and gray level correlation, TR-1091, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, August 1981.

19. K. Prazdny, Computing motions of (locally) planar surfaces from spatio-temporal changes in image brightness: a note, TR-1090, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, August 1981.

20. L. S. Davis, H. F. Sun, and Z. Q. Wu, Motion detection at corners, TR-1130, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, December 1981.

Figure 1. Multispectral "superspike". a) (Top) Red and green bands of a color image of a house. (Bottom) Scatter plot of (red, green) values, linearly (left) and logarithmically (right) scaled.

b) Results after application of "superspike"; the parts correspond to those in (a).

Figure 2. 28 texture samples. Left: grass, raffia, sand, wool. Right: three geological terrain types.

Figure 3. Four 5x5 Laws masks.

Table 1. Numbers of samples correctly classified using a single texture feature. CONX and CONY are Haralick's CON feature for displacements (1,0) and (0,1); (W)E/A is (magnitude-weighted) amount of edge per unit area.