COMPUTER SCIENCE TECHNICAL REPORT SERIES

UNIVERSITY OF MARYLAND

COLLEGE PARK, MARYLAND
20742

(1)

TR-508
DAAG-53-76C-0138

February 1977

SEGMENTATION OF FLIR IMAGES BY PIXEL CLASSIFICATION

Durga P. Panda
Computer Science Center
University of Maryland
College Park, MD 20742

ABSTRACT

Image segmentation can be treated as a point-wise classification problem. This classification may be done by measuring a set of features at each point and defining a decision surface in the feature space. This report presents some experiments in segmenting FLIR images by using the gray level and the edge value at each point as features.

The support of the U. S. Army Night Vision Laboratory under Contract DAAG-53-76C-0138 (ARPA Order 3206) is gratefully acknowledged, as is the help of Mrs. Shelly Rowe.

1. Introduction

Two earlier reports [1 , 2] have analyzed the joint histogram of gray level and edge value of FLIR images and have suggested possible segmentation procedures based on the analysis. The analysis indicated that the histogram is trimodal, two of the modes occurring at zero edge value and the third one occurring at some higher edge value and at a gray level between those of the first two. Some of the segmentation procedures suggested in the two reports [1 , 2] are: thresholding based on the histogram of gray levels having low edge values; thresholding based on the histogram for high edge values; and valley seeking in the joint histogram. (These methods are defined below.) The present paper investigates the success of these methods in segmenting FLIR images into backgrounds and objects.

The segmentation procedure based on the histogram of points having low edge values (which will be referred to here as the "L-method") finds the valley between the modes in the histogram and uses the location of that valley as the gray level threshold for the image. The segmentation procedure based on the histogram of points having high edge values is referred to as the "H-method"; it uses as threshold the conditional mean or the conditional mode of all pixels with edge value greater than a certain percentile, p, of the maximum possible edge value. The quantity p is taken to be 95; the heuristics leading to this selection are discussed in detail in [1, 2].

Segmentation by valley seeking in the joint histogram involves finding a "bottommost" curve that separates one mode from the others. As described in [1, 2], two such curves are possible. These two curves have a few common points at low edge values and then diverge from each other as edge value increases. Thus, for a given edge value, one of the curves, which we shall call "L", has gray levels smaller than or equal to those on the other curve, which we call "R". Either of these curves may be used as a decision surface in the two-dimensional space of gray level and edge value, for classifying the image points into the object and background classes. The segmentation procedure using valleys such as the curve L or the curve R will be called the "V-method". In particular, the method using the curve R as the decision surface will be called the "VR-method", and the method based on the curve L will be called the "VL-method".

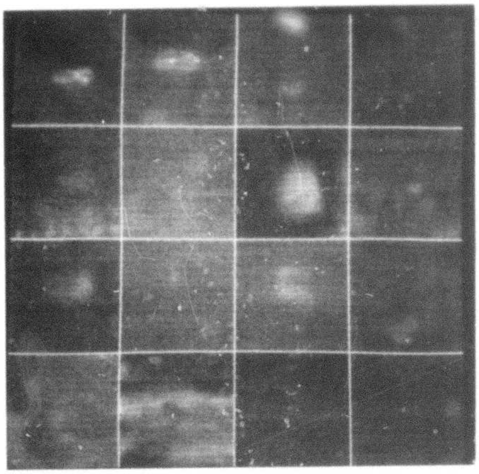
The goal of the work reported here was to investigate the usefulness of the above methods as segmentation procedures for FLIR images, and not necessarily to automate these methods. For this reason the valley selections were done manually.

2. Experiments

Sixteen windows were selected from the "NVL data base" (see [2]) as test images. The images were 64x64 in size and had grayscales of 0 to 63. Figure 1a shows the 16 images. The identifiers of these images, as given in [2], are shown in Figure 1b.

Figure 2 shows the gray level histograms of the images in Figure 1. Most of these histograms do not possess strong bimodality. Classically, images can be segmented using thresholds located at valley bottoms on their histograms. In locating valleys on the histograms of Figure 2, consideration was given to the fact that the object points are much brighter than the background points and they are a small fraction of the total number of points in the image. Thus, a valley near the middle of the grayscale range occupied by the histogram was given less weight, or discarded as spurious, compared to a valley much less prominent but occurring near the light end of the grayscale range. The results of segmenting the images of Figure 1 using valleys found in this way on the histograms in Figure 2 are shown in Figure 3. As this figure shows, only a few of the objects are extracted satisfactorily.

Figure 4 shows the joint (gray level, edge value) histograms of the same images. The horizontal direction represents gray level and the vertical direction represents edge value. The brightness at each point in a joint histogram represents the number of pixels having the corresponding gray level and edge value. The edge operator used is the



(a)

6T	15T	34T	57T
3R	26R	47R	58R
21A	34A	48A	57A
2N	20N	38N	56N

(b)

Figure 1. The 16 test images.

(a) The images.
(b) The image names. The suffixes T, R, A, and N indicate that the object in the image is a tank, a truck, an APC, or a "non-target", respectively.

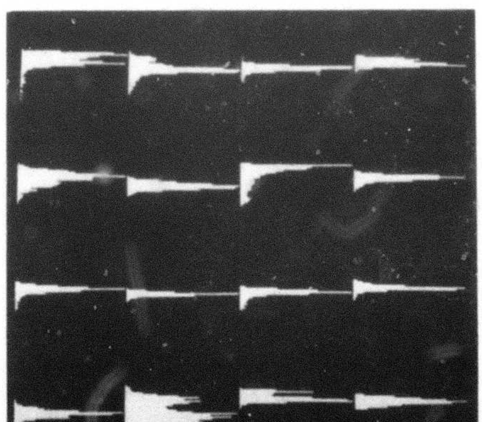


Figure 2. The gray level histograms of the test images.

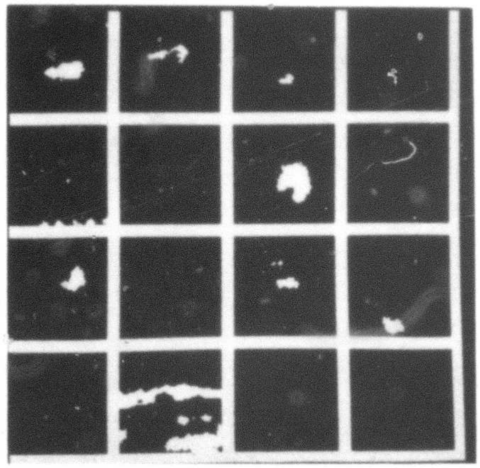


Figure 3. The image segments extracted by finding valleys in the gray level histograms of Figure 2.

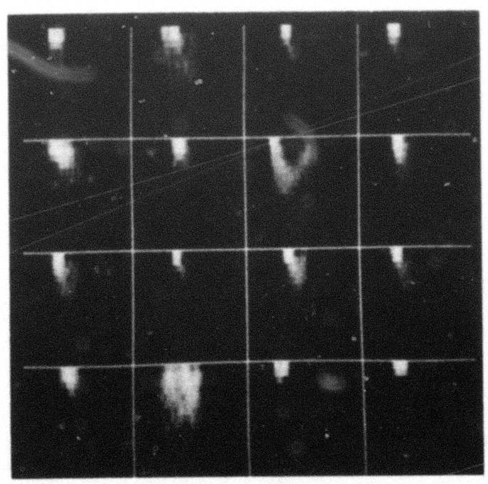


Figure 4. The two-dimensional histograms of the test images.

4x4 DIFF operator of Hayes and Rosenfeld [3]. The general structure of these joint histograms indicates that for low edge values there is a clearly distinct valley (dark region) separating a strong mode, corresponding to the background region, from a weak mode, corresponding to the object region.

Figure 5 shows the results of segmentation by the L-method, in which the gray level at the valley of the histogram for edge value zero (a common point of the curves L or R) is chosen as a threshold. This method is more successful at extracting the objects, but images with faint objects have very small extracted segments. In image 26R the histogram for zero edge value had no valley, so that the L-method yielded no extracted segment for this image.

Figure 6 shows the result of segmenting the test images by the H-method, as was done in [2]. For many of the test images the objects are well segmented by this method. However, for the images with extremely faint objects the output of the H-method is very noisy. The most undesirable results occur for the two images 38N and 56N, where even though the images contain no object, the H-method classifies some regions as objects. This is due to the fact that the threshold found by the H-method is always within the gray-scale occupied by the image and hence it will always yield some segmented regions in the image regardless of whether or not the image contains an object.

Figure 7a shows the test images segmented by the VR-method, while Figure 7b shows similar results for the VL-method. It may be pointed out here that the V-method of

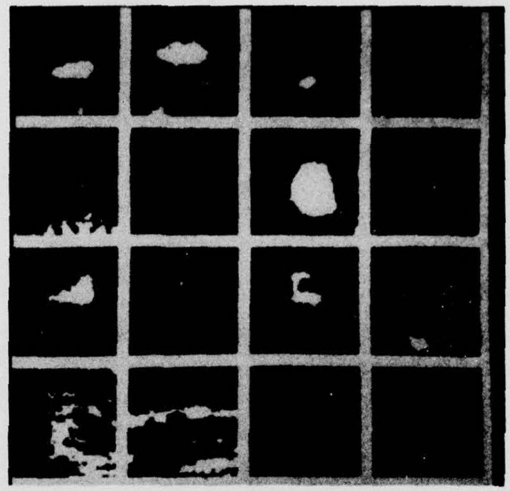


Figure 5. The test images segmented by the L-method.

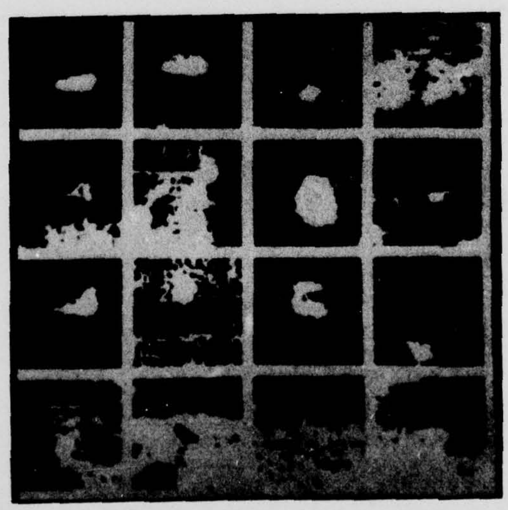
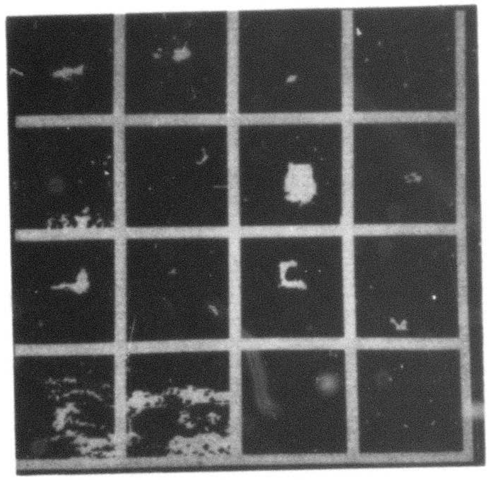
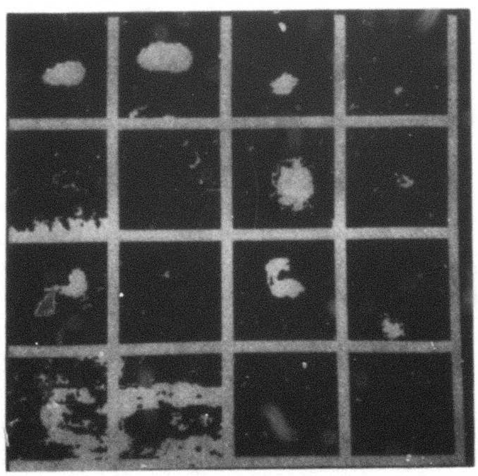


Figure 6. The test images segmented by the H-method.



(a)



(b)

Figure 7. The test images segmented by the VR- and the VL-methods.

a) The VR-method.
b) The VL-method.

segmentation is based on the concept that an image may contain pixels belonging to three different classes, the background, the object, and the object boundary (see [2]). Pixels belonging to the object boundary class are expected to surround, in the image domain, the pixels belonging to the object class; and are expected to have higher edge values than those of the object or the background pixels, in general. Since the curve R separates the object pixels from both the background pixels and boundary pixels, using the curve R as the decision surface will exclude the boundary points from the segmented image. Comparison of Figure 6 with Figure 7a indicates that this is indeed true in general. However, for the images containing very small and faint objects, such as 57T, 58R, and 34A, the VR-method yields relatively noise-free segments as compared to the H-method. Also impressive is the result of the VR-method for the last two non-target images, 38N and 56N, where the extracted segments are empty. The two-dimensional histograms of these images display no valley and hence, in contrast with the H-method, the extracted segments are empty. Conversely, the curve L separates the background pixels from both the objects and the object boundaries. Thus using the curve L as the decision surface will include in the extracted segments the boundary points that the VR-method excluded. As Figure 7b shows, the extracted segments are larger in the case of the VL-method than they are in the case of the VR-method.

Figure 8 shows the result of using a hybrid of the

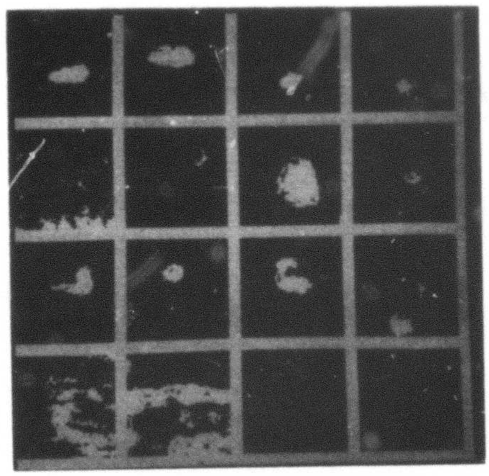


Figure 8. The test images segmented by the VH-method.

VR-method, the VL-method, and the H-method as the segmentation procedure. This method classifies the pixels using a decision surface constructed as follows. For a given edge value, beginning with the edge value zero, if the two curves L and R have a common gray level then that (gray level, edge value) pair is selected as a point on the decision surface. As the edge value is increased, the two curves will begin to depart from each other at some point. For this and all higher edge values, the threshold used is the same as the threshold of the H-method. In other words, for low edge values the points on the decision surface are chosen by the V-method, and for higher edge values the points are chosen by the H-method. Some of the extracted segments that were very small in the VL-method are relatively large in the VH-method.

An alternative to the VH-method would be to classify the pixels by the straight line S joining the threshold due to the H-method at the 95th percentile edge value with the threshold due to the L-method at zero edge value. Figure 9 shows the results of segmenting the FLIR images by this S-method. This method does not follow the actual valley bottom for the low edge values, and the results are therefore somewhat inferior to those obtained using the VH-method. The VH-method yields the best result of all the methods tested in this report.

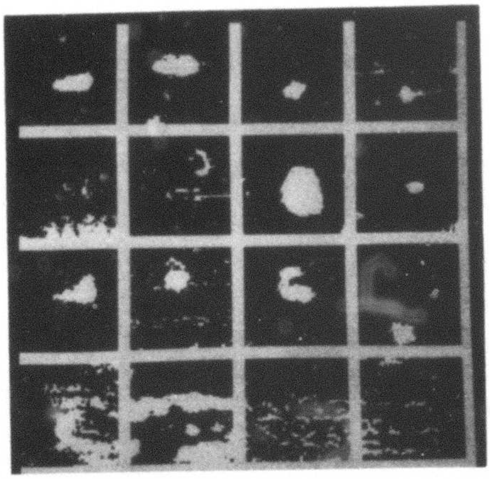


Figure 9. The test images segmented by the S-method.

3. Discussion and Conclusions

It is evident that the two-dimensional histogram enables us to extract better objects from FLIR images than those extracted using the one-dimensional gray level histogram. Several different decision surfaces can be used in the two-dimensional feature space, to give varied degrees of success in segmentation. Among all the methods considered, the VH-method seems to give the best results (Figure 8). A heuristic explanation for this is the following. The background and the object pixels away from the object boundary have low edge values, and the histogram for these pixels seems to have a distinct bimodality. Thus the valley between the two modes successfully classifies such pixels into the background class and the object class. The pixels near the boundary, however, have higher edge values and do not have this bimodality. Since some of these points are from the object class and some from the background class, the mean value of such points may be expected to classify the pixels successfully. The VH-method of classification is effectively just that -- the low edge value points on the decision curve are at the valleys of the corresponding histograms, and the high edge value points on the curve are at the mean of the high edge value pixels.

The two-dimensional histograms sometimes resemble truncated or "folded over" mixtures of two multi-variate normal distributions with unequal covariances. The folding over, which occurs at the edge value zero, may be due to the fact that the edge value at each pixel is defined as an absolute value of certain differences measured at that pixel. It is conceivable that if somehow appropriate signs, positive or negative, were incorporated into the edge value at each pixel, the resultant distribution would be an "unfolded" mixture of two multi-variate normal distributions with unequal covariances. In such a case the maximum-likelihood decision surface is quadratic. Unfortunately, how to incorporate the appropriate sign into the edge value at a pixel is not obvious at present.

While the use of the edge value as an additional feature has certainly improved the results of pixel classification, it is obvious that the edge value is not the only feature that can be used for this purpose. It is conceivable that there exist other local properties that perform as well as or better than the edge value. Further studies of this approach to image segmentation would be desirable.

References

1. Algorithms and hardware technology for image recognition, Quarterly DARPA Report for the period 1 May - 31 July, 1976, Computer Science Center, University of Maryland, College Park, MD.

2. Algorithms and hardware technology for image recognition, Semi-Annual Report for the period 1 May - 31 October, 1976, Computer Science Center, University of Maryland, College Park, MD.

3. K. C. Hayes, Jr. and A. Rosenfeld, Efficient edge detectors and applications, Tech. Rep. No. 207, Computer Science Center, University of Maryland, College Park, MD, November 1972.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
SEGMENTATION OF FLIR IMAGES BY PIXEL CLASSIFICATION.

5. TYPE OF REPORT & PERIOD COVERED
Technical rept.

6. PERFORMING ORG. REPORT NUMBER
TR-508

7. AUTHOR(s)
Durga P. Panda

8. CONTRACT OR GRANT NUMBER(s)
DAAG-53-76C-0138 ARPA Order-3206

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Computer Science Ctr.
Univ. of Maryland
College Park, MD 20742

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

11. CONTROLLING OFFICE NAME AND ADDRESS
U. S. Army Night Vision Lab.
Ft. Belvoir, VA 22060

12. REPORT DATE
Feb. 1977

13. NUMBER OF PAGES
12 18p.

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

403 018

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Image Segmentation
Pattern Classification
FLIR Imagery

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
Image segmentation can be treated as a point-wise classification problem. This classification may be done by measuring a set of features at each point and defining a decision surface in the feature space. This report presents some experiments in segmenting FLIR images by using the gray level and the edge value at each point as features.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)