TR-626
DAAG53-76C-0138

January 1978

SOME EXPERIMENTS ON VARIABLE THRESHOLDING

Yasuo Nakagawa*
Azriel Rosenfeld
Computer Science Center
University of Maryland
College Park, MD 20742

COMPUTER SCIENCE TECHNICAL REPORT SERIES

UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND
20742

LEVEL II

See 1473 in back

TR-626
DAAG53-76C-0138

January 1978

SOME EXPERIMENTS ON VARIABLE THRESHOLDING

Yasuo Nakagawa*
Azriel Rosenfeld
Computer Science Center
University of Maryland
College Park, MD 20742

ABSTRACT

Chow and Kaneko proposed a method of variable thresholding in which an image is divided into windows; thresholds are selected for those windows that have bimodal histograms; and these thresholds are interpolated to define a variable threshold for the entire image. This method was applied to several TV images of machine parts; the results obtained appeared to be considerably better than the results of thresholding at a fixed level. An extension of the method was defined that allowed histograms to be either bimodal or trimodal; this yielded some further improvement in the results, but was also more sensitive to shadows. Finally, an adaptive quantization scheme, based on histogram peak sharpening, was applied to two of the images; the results do not seem to be as good as those obtained using variable thresholding. Some results for FLIR images of tactical targets are also presented.

The support of the U. S. Army Night Vision Laboratory under Contract DAAG53-76C-0138 (ARPA Order 3206) is gratefully acknowledged as is the help of Mrs. Shelly Rowe in preparing this paper.

---

*Permanent address: Production Engineering Research Lab.
Hitachi Ltd., 292 Yoshida-cho, Totsuka-ku,
Yokohama, Japan 244

1. Introduction

In [1] Chow and Kaneko proposed a method of variable thresholding for image segmentation. In this method, the image is divided into windows; a gray level histogram is computed for each window; and thresholds are selected for those windows that have bimodal histograms. (A bimodal histogram indicates that the gray level population in the given window is a mixture of the subpopulations; by choosing a threshold that separates the two peaks, we can discriminate between these populations.) These thresholds are then interpolated (or extrapolated) to define a variable threshold for the entire image. This method is particularly useful in cases where there is a large range of variation in gray scale from one part of the image to another, so that a single fixed threshold cannot be used for the entire image. Chow and Kaneko successfully applied this method to detect the heart region on chest x-rays.

This report describes an application of the Chow-Kaneko approach to TV images of machine parts. These images were obtained from General Motors Research Laboratory; they are part of a data base described in [2]. Some experiments were also performed using Forward Looking InfraRed (FLIR) images of tactical targets; these are described in Section 5.

The TV images used are shown in Figure 1. Parts (a) and (b) each show four nonoverlapping "connecting rods", but (b) has more uneven illumination than (a). Parts (c) and (d)

show cases involving overlap and shadows; these conditions are much more severe in (d) than in (c).

The version of Chow and Kaneko's method used is described in Section 2, and the results obtained are compared with the results of using a single fixed threshold for the entire image. The variable-threshold results are considerably better. An extension of the Chow-Kaneko approach, described in Section 3, took into consideration both bimodal and trimodal histograms, and allowed either one or two thresholds to be chosen for each window; this yielded some further improvement in the results, but was also more sensitive to shadows. Finally, an adaptive quantization scheme, based on histogram peak sharpening [3-4], was applied to the images; the results (Section 4) do not seem to be as good as those obtained using variable thresholding. Section 5 shows results for a set of FLIR images.

2. Variable thresholding based on local bimodality

The first step in the Chow/Kaneko method is to divide the image into windows. The windows used for the images of Figure 1 are indicated by grid boxes. In this case the images were 256x243 pixels each; we used windows of size 32x32 (the bottom three rows of each image were ignored), so that each image was divided into 64 windows.

The method used to determine bimodality and to select thresholds for the bimodal histograms was based on a process of Gaussian fitting. This method consisted of the following steps, which were carried out for each of the window histograms:

a) The mean and standard deviation of the histogram are computed. These are defined by

$$\mu = \frac{1}{N} \sum F(i) \cdot i$$

$$\sigma = \sqrt{\frac{1}{N} \sum F(i) \cdot (i-\mu)^2}$$

where F(i) is the histogram value for gray level i (i.e., the number of window points having gray level i), and N is the number of points in the window (= 960 in our case). [In our images, there were 32 gray levels, so that all sums were taken over the range $0 \leq i \leq 31$.] If $\sigma \leq 3$, no threshold was computed for the histogram. If $\sigma > 3$, the threshold computation was carried out as described next.

b) A least-squares fit of

$$f(i) = \frac{P_1}{\sigma_1} e^{-(i-\mu_1)^2/2\sigma_1^2} + \frac{P_2}{\sigma_2} e^{-(i-\mu_2)^2/2\sigma_2^2}$$

to the histogram F(i) is found by adjusting the parameters P_1, μ_1, σ_1, P_2, μ_2, and σ_2. This is done as follows:

b1) The histogram is smoothed by taking a local weighted average:

$$F'(i) = \frac{F(i-2)+2F(i-1)+3F(i)+2F(i+1)+F(i+2)}{9}$$

On the smoothed histogram, the deepest valley (= lowest value) is found, and is used to divide the histogram into two parts. Initial estimates of P_1, μ_1, σ_1, P_2, μ_2, σ_2 are computed on these two parts (for the original histogram F(i)) as follows:

$$N_1 = \sum_{i=0}^{v} F(i) \qquad N_2 = \sum_{i=v}^{31} F(i)$$

$$\mu_1 = \frac{1}{N_1} \sum_{i=0}^{v} F(i)\cdot i \qquad \mu_2 = \frac{1}{N_2} \sum_{i=v}^{31} F(i)\cdot i$$

$$\sigma_1 = \sqrt{\frac{1}{N_1} \sum_{i=0}^{v} F(i)\cdot(i-\mu_1)^2} \qquad \sigma_2 = \sqrt{\frac{1}{N_2} \sum_{i=v}^{31} F(i)\cdot(i-\mu_2)^2}$$

$$P_1 = \frac{N_1\sigma_1}{\sum_{i=0}^{v} e^{-(i-\mu_1)^2/2\sigma_1^2}} \qquad P_2 = \frac{N_2\sigma_2}{\sum_{i=v}^{31} e^{-(i-\mu_2)^2/2\sigma_2^2}}$$

b2) A hill-climbing method is used to minimize $\sum_{i=0}^{31} [f(i)-F(i)]^2$, using the program FMCG, which is available in the IBM System/360 Scientific Subroutine Package. In this program, the following parameters were used:

EST (estimate of the minimum function value) = 1.0

EPS (expected absolute error) = 0.1

LIMIT (maximum number of iterations) = 30

c) The resultant best-fitting f(i) is tested for bi-modality, based on the criteria

$$\mu_2-\mu_1 > 4$$

$$0.1 < \sigma_1/\sigma_2 < 1.00$$

$$\delta_{12} < 0.8$$

where δ_{12}, the valley-to-peak ratio, is defined by

$$\delta_{12} = \frac{\text{Minimum value of f in } [\mu_1,\mu_2]}{\text{Min}[f(\mu_1),f(\mu_2)]}$$

If these tests are not satisfied, no threshold is selected for that window. If they are satisfied, a threshold is selected that minimizes the probability

of misclassification for the mixture distribution f(i). This threshold t satisfies

$$\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right) t^2 + 2\left(\frac{\mu_2}{\sigma_2^2} - \frac{\mu_1}{\sigma_1^2}\right) t + \frac{\mu_1^2}{\sigma_1^2} - \frac{\mu_2^2}{\sigma_2^2} + 2 \ln \frac{P_2\sigma_1}{P_1\sigma_2} = 0.$$

After thresholds have been chosen in this way for the bimodal windows, thresholds are then defined for the other windows by a local weighted averaging process. Specifically, let T(u,v) be the threshold assigned to the window centered at (u,v), or 0 if no threshold was assigned to that window. Then for a window, say centered at (x,y), for which no threshold has yet been assigned, we compute the threshold

$$(T(x+1,y) + T(x-1,y) + T(x,y+1) + T(x,y-1)$$

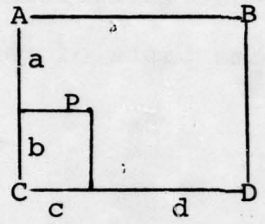
$$+ \frac{1}{\sqrt{2}}(T(x+1,y+1) + T(x+1,y-1) + T(x-1,y+1) + T(x-1,y-1))$$

provided at least one of these neighboring windows, other than a diagonal neighbor, has had a threshold assigned. We then smooth the resulting array of thresholds by local weighted averaging using the array of weights

$$\begin{array}{ccc} \frac{1}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}} \\ 1 & 2 & 1 \\ \frac{1}{\sqrt{2}} & & \frac{1}{\sqrt{2}} \end{array}$$

Finally, we assign thresholds to individual points by bilinear interpolation on the window thresholds. Let P be

surrounded by the four window centers A, B, C, D as shown below, where the thresholds for these windows are T_A, T_B, T_C, and T_D, respectively.



Then the threshold for P is taken to be

$$t_P = \frac{1}{(a+b)(c+d)}[bdT_A+bcT_B+daT_C+cat_D]$$

If P is not surrounded by four window centers (i.e., it lies near the border of the image), its threshold is taken to be that at the nearest window center.

Figure 2 shows the window histograms for the four pictures in Figure 1. [All of these were scaled so the tallest peak has a fixed height; the vertical lines provide an indication of the scale used on each histogram (each line represents 50 pixels).] Figure 3 shows the two-Gaussian approximations to the histograms for those winodws that were judged to be bimodal. The point thresholds obtained by bilinear interpolation are displayed in Figure 4. (The horizontal lines in Figure 2 showed the thresholds at the window centers, superimposed on the window histograms.) Figure 5 shows the results of applying these point thresholds to the pictures in

Figure 1. For comparison, the results of using a fixed threshold (chosen by fitting two Gaussians to the histograms of the entire images) are shown in Figure 6. They are much worse in the first three cases, but (surprisingly) somewhat better in the fourth case, in some parts of the image.

3. Variable thresholding based on local bi- or trimodality

An extension of the Chow-Kaneko approach was also implemented in which the local histograms were tested for both bimodality and trimodality. The latter test involves least-squares fitting of

$$g(i) = \frac{P_3}{\sigma_3} e^{-(i-\mu_3)^2/2\sigma_3^2} + \frac{P_4}{\sigma_4} e^{-(i-\mu_4)^2/2\sigma_4^2} + \frac{P_5}{\sigma_5} e^{-(i-\mu_5)/2\sigma_5^2}$$

to the histogram F(i). The initial values of the nine parameters (P's, μ's, and σ's) are obtained by dividing the smoothed histogram at its deepest and second-deepest valleys. The hill-climbing iteration process was the same as described in Section 2.

The bimodality test ((c) in Section 2) was applied to the two-Gaussian fits, and the same test was also applied to each pair of Gaussians ((3,4) and (4,5)) in the three-Gaussian fits. If both the two- and three-Gaussian fits satisfied these tests for a given window, the three-Gaussian fit was used.

Let t_{12} be the minimum-error threshold for the two-Gaussian fit, and let t_{34} and t_{45} be the minimum-error thresholds for the pairs of Gaussians (3,4) and (4,5). Let $\bar{t}_{34}$ and $\bar{t}_{45}$ be the averages of these latter thresholds in those neighboring windows for which three-Gaussian fits were used. For each window at which the two-Gaussian fit was chosen, a decision was made as to whether the resulting threshold t_{12}

should correspond to the higher or lower threshold of a three-Gaussian fit. Specifically, if $|\bar{t}_{45}-t_{12}| > |\bar{t}_{34}-t_{12}|$ and $|\bar{t}_{34}-t_{12}| < 4$, t_{12} was regarded as a t_{34} threshold; if $|\bar{t}_{34}-t_{12}| \geq |\bar{t}_{45}-t_{12}|$ and $|\bar{t}_{45}-t_{12}| < 4$, t_{12} was regarded as a t_{45} threshold.

Thresholds were then interpolated for the remaining windows (at which neither two- nor three-Gaussian fits were obtained), and for the individual image points, exactly as in Section 2.

Figure 7 shows the two- and three-Gaussian approximations (whichever was chosen in each case) to the window histograms. The t_{34} and t_{45} interpolated point thresholds are shown in Figures 8 and 9, and the window histograms with the two thresholds used at the window centers superimposed are shown in Figure 10. Figure 11 shows the thresholded images, with the three gray level ranges displayed as black, gray, and white. These seem somewhat more useful than the two-Gaussian thresholded images of Figure 5, but (in the third and fourth parts) they are seen to be quite sensitive to shadows.

4. Comparison with iterative histogram modification

A possible alternative to variable thresholding might be to use an adaptive requantization scheme which identifies peaks on a histogram and replaces each peak by a spike (i.e., it maps all gray levels belonging to the given peak into a single gray level). Such a scheme is described in [3, 4]; it iteratively enhances maxima on a (smoothed) histogram until the major peaks become spikes. By increasing the amount of smoothing used, one can reduce the number of spikes that result, since fewer peaks will be regarded as "major". This approach does not require the number of peaks to be specified (e.g., 2 or 3, as in the present study); the shape of the histogram determines the number of peaks that are detected at a given degree of smoothing.

Figure 12 illustrates the application of this method to two of the pictures (a and c) in Figure 1, for various amounts (W) of smoothing. [The histograms of the original pictures, and the resulting spike histogram for each value of W, are shown in Figure 13.] For the first picture, the results for two- and three-spike histograms are reasonable, but are not as good as those obtained in Section 2 and 3. For the second picture, the three-spike result is somewhat affected by shadows and suffers from loss of internal detail.

It appears that adaptive quantization is not as good as variable thresholding for these types of pictures. This may be because the adaptive quantization approach is best suited

for histograms that are mixtures of sharply unimodal distributions; in an unevenly illuminated picture, on the other hand, the gray level distribution from a uniformly reflective region is broad and flat-topped, rather than sharply peaked.

5. Results for FLIR images

Figure 14 shows a set of four Forward Looking InfraRed (FLIR) images containing warm objects on a cooler background. These pictures were divided into 32x32 pixel windows (as indicated by the grids), and were processed essentially as in Section 2, with the following exceptions.

a) The criterion used for picking windows with large standard deviations was $\sigma \geq 2.4$, rather than $\sigma \geq 3$; even with this smaller threshold, very few windows were chosen.

b) The initial estimates of the parameters in the two-Gaussian approximations were chosen on the original (unsmoothed) window histograms; these histograms were judged to be relatively smooth, so that further smoothing did not seem to be necessary.

c) The bimodality criteria used were

$$\mu_2 - \mu_1 > 3;\ 0.1 \quad \sigma_1/\sigma_2 < 10.0 \quad ;\ \delta_{12} < 1.0$$

In spite of these more tolerant criteria on μ and δ, very few windows were judged to be bimodal.

Figure 15 shows the window histograms for the four FLIR images; Figure 16 shows the two-Gaussian approximations to the histograms for those (few) windows that were judged bimodal. The interpolated point thresholds for these pictures are not shown, because they have almost no variability . (The horizontal lines superimposed on the histograms in Figure 15

show the thresholds at the centers of the windows.) The results of applying these thresholds to the pictures are shown in Figure 17. For comparison, the results of using a fixed threshold for each picture are shown in Figure 18. This threshold was chosen using the (gray level, edge value) clustering scheme described in [5]. The variable-thresholding results seem to be less noisy for three out of the four pictures, but are noisier for picture (b).

The three-Gaussian approximation method was also applied to the FLIR images, but almost no windows were found to be trimodal. The results of this method are therefore not shown here.

6. Concluding remarks

Chow and Kaneko's method of variable thresholding based on local bimodality detection, which was developed for use in angiography, appears to be useful in other applications as well. An extension to allow for the possibility of trimodality yielded relatively little improvement.

References

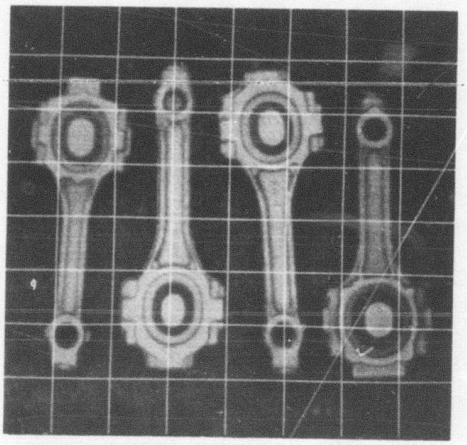
1. C. K. Chow and T. Kaneko, Automatic boundary detection of the left ventricle from cineangiograms, Computers in Biomedical Research 5, 1972, 388-410.

2. M. L. Baird, A computer vision data base for the "industrial bin of parts" problem, Research Publication GMR-2502, Research Laboratories, General Motors Corp., Warren, MI, August 1977.

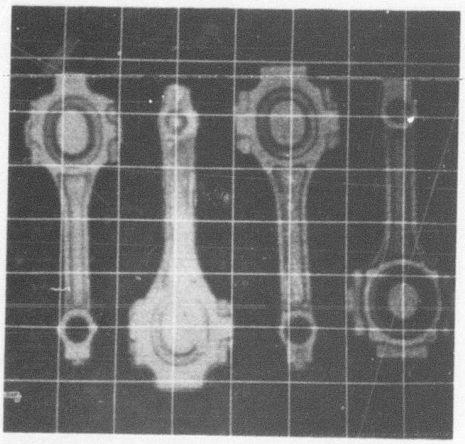
3. A. Rosenfeld and L. S. Davis, Iterative histogram modification, IEEE Trans. Systems, Man, and Cybernetics, to appear. (University of Maryland, Computer Science Center, Technical Report 519, April 1977.)

4. S. Peleg, Iterative histogram modification, 2, University of Maryland, Computer Science Center, Technical Report 606, November 1977.

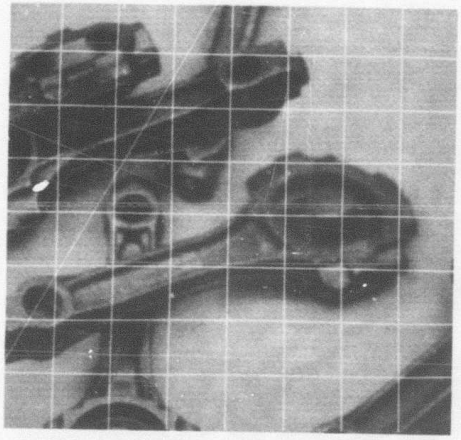
5. D. L. Milgram and M. Herman, Clustering edge values for threshold selection, University of Maryland Computer Science Center, Technical Report 617, December 1977.



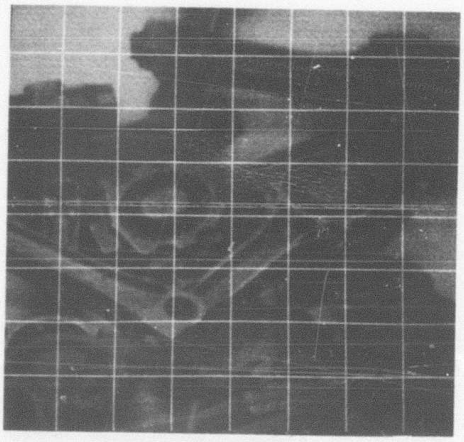
(a)



(b)

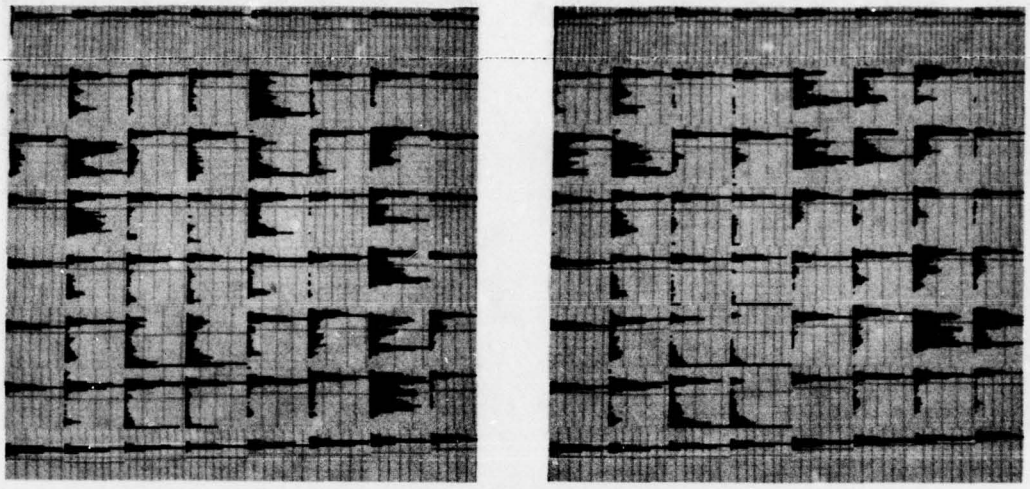


(c)

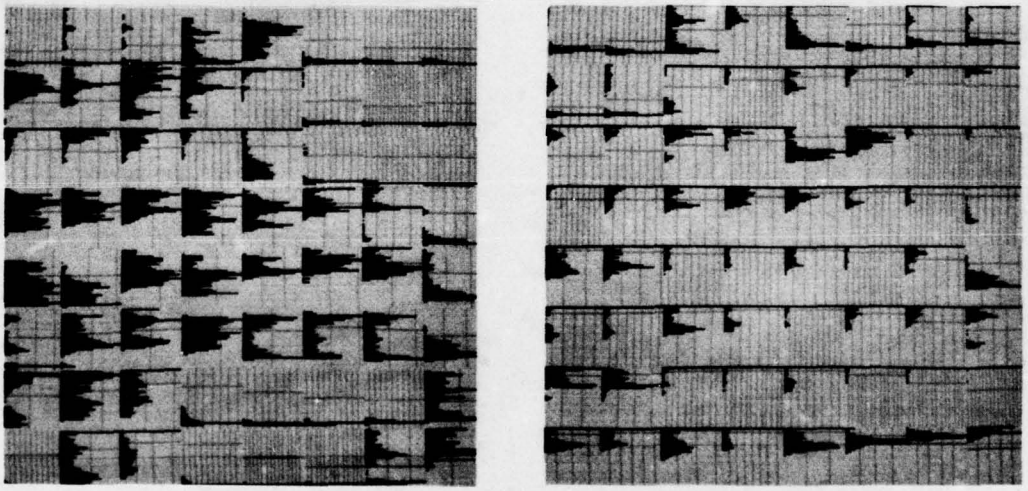


(d)

Figure 1. Original pictures, with grids superimposed showing windows over which local histograms were computed.



(a) (b)



(c) (d)

Figure 2. Window histograms for the pictures in Figure 1.

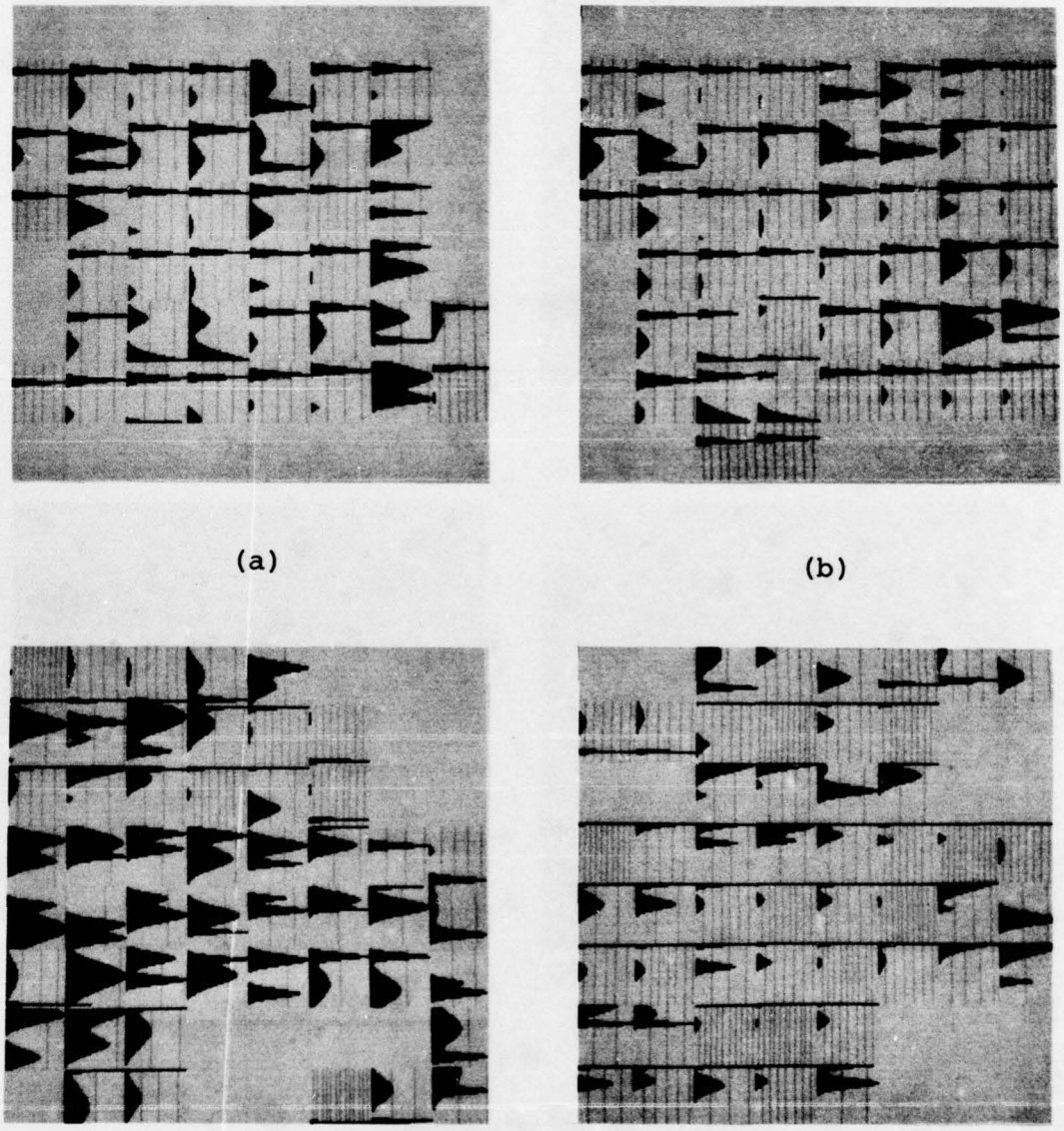


Figure 3. Two-Gaussian approximations to those window histograms in Figure 2 that were judged to be bimodal.

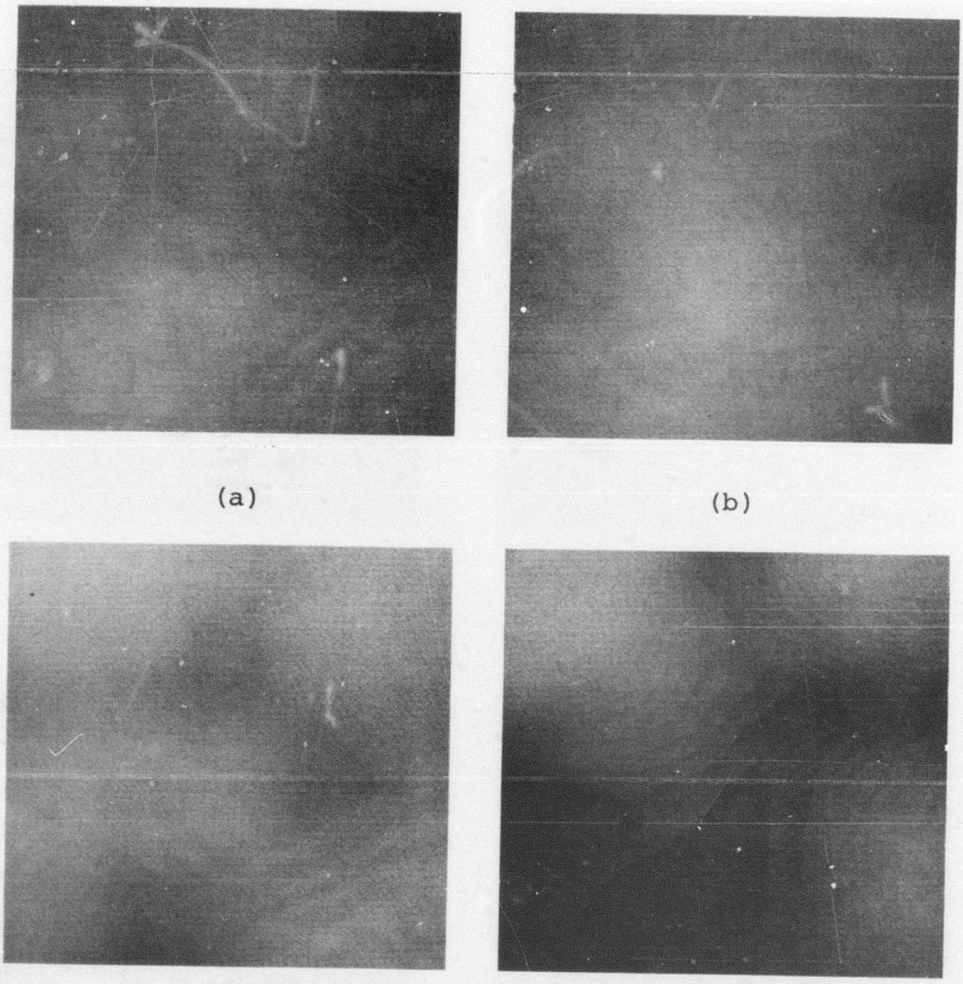
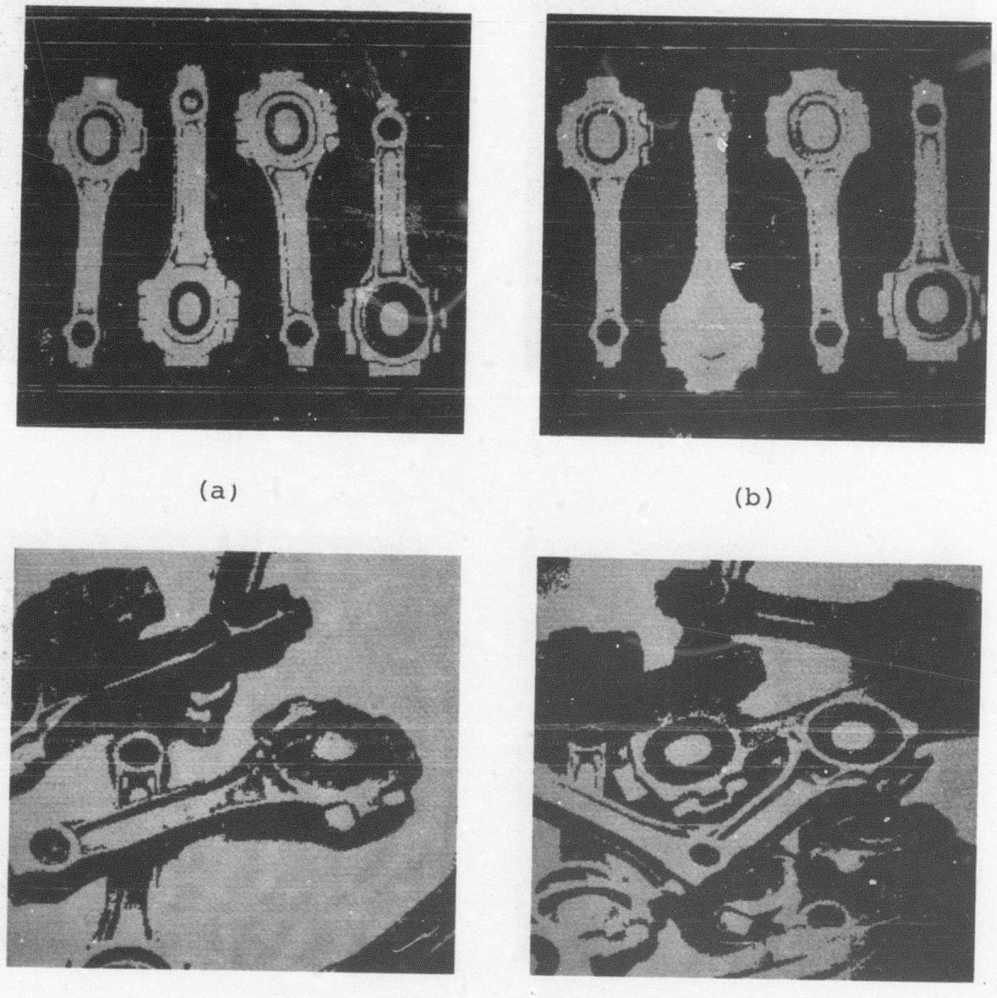


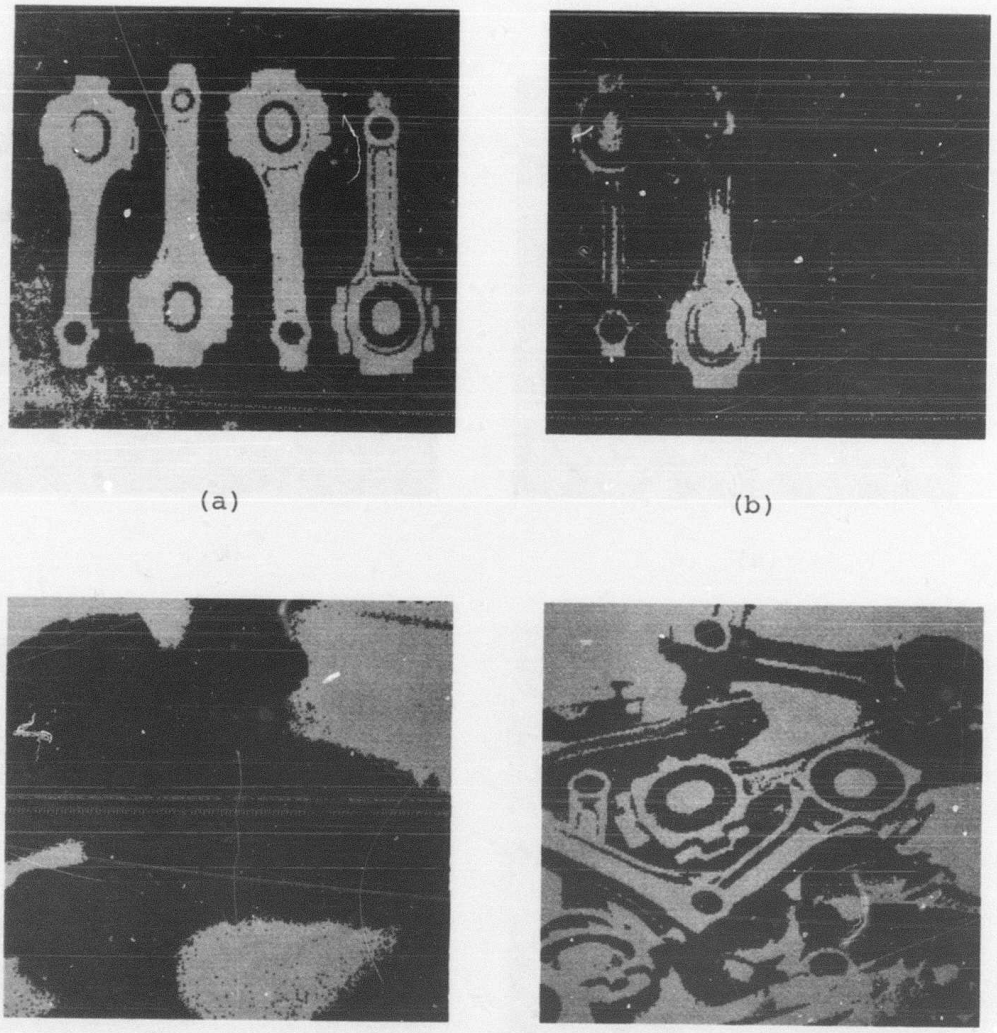
Figure 4. Point thresholds, obtained by interpolation on the window thresholds, for the four pictures.



(a) (b)

(c) (d)

Figure 5. Results of applying the thresholds shown in Figure 4 to the pictures of Figure 1.



(a) (b)

(c) (d)

Figure 6. Results of applying fixed thresholds (obtained by fitting two Gaussians to the histograms of the entire pictures) to the pictures of Figure 1.

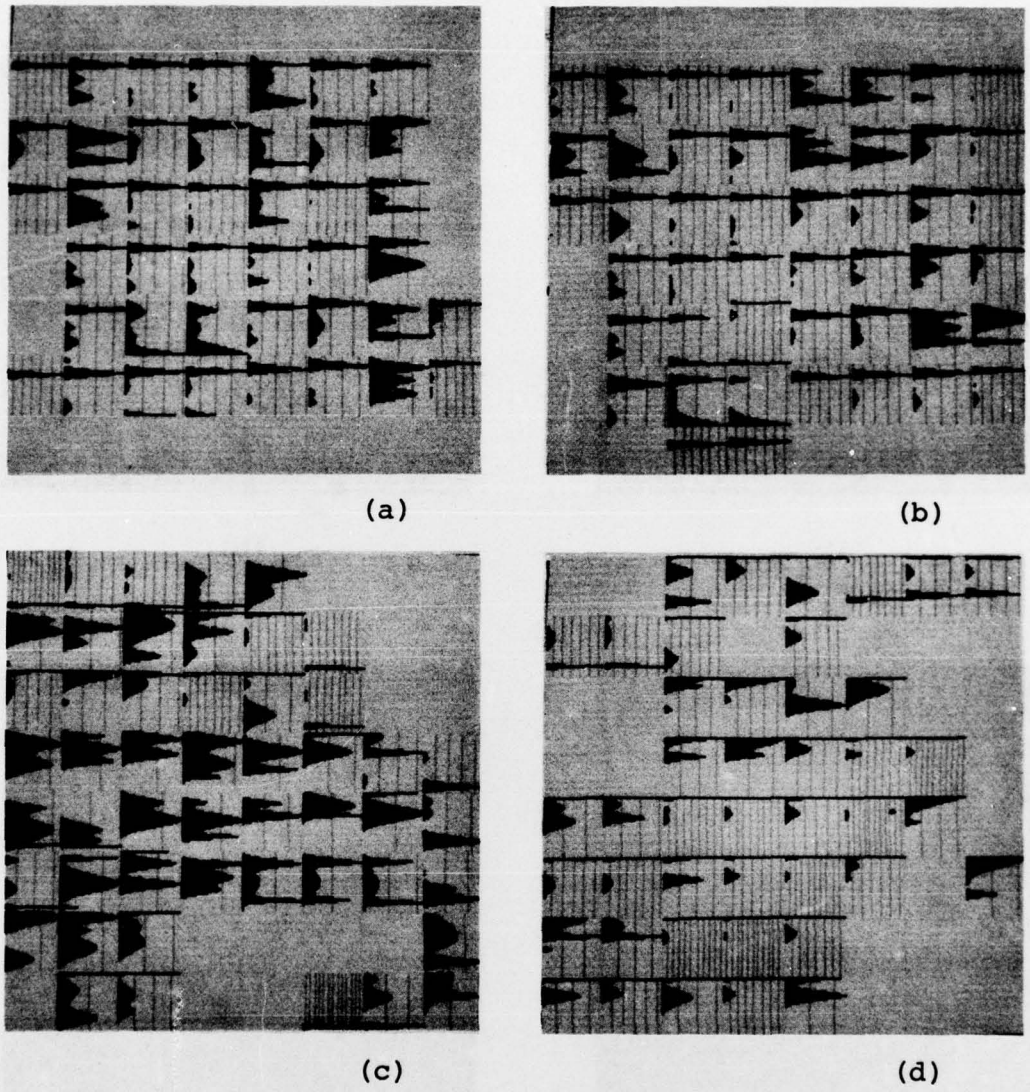
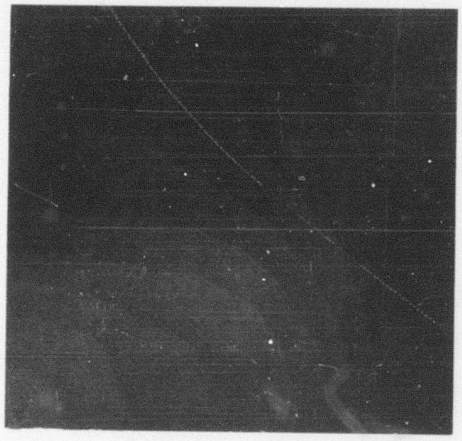
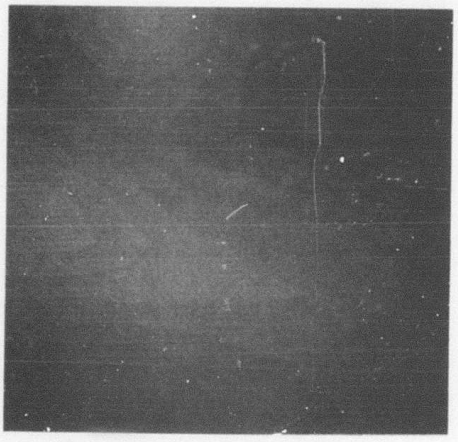


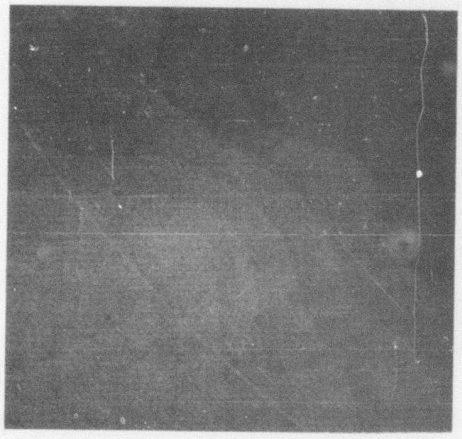
Figure 7. Two- and three-Gaussian approximations to those window histograms in Figure 2 that were judged to be bi- and tri-modal.



(a)



(b)



(c)

(d)

Figure 8. Darker thresholds obtained for the four pictures by interpolating on the window thresholds.

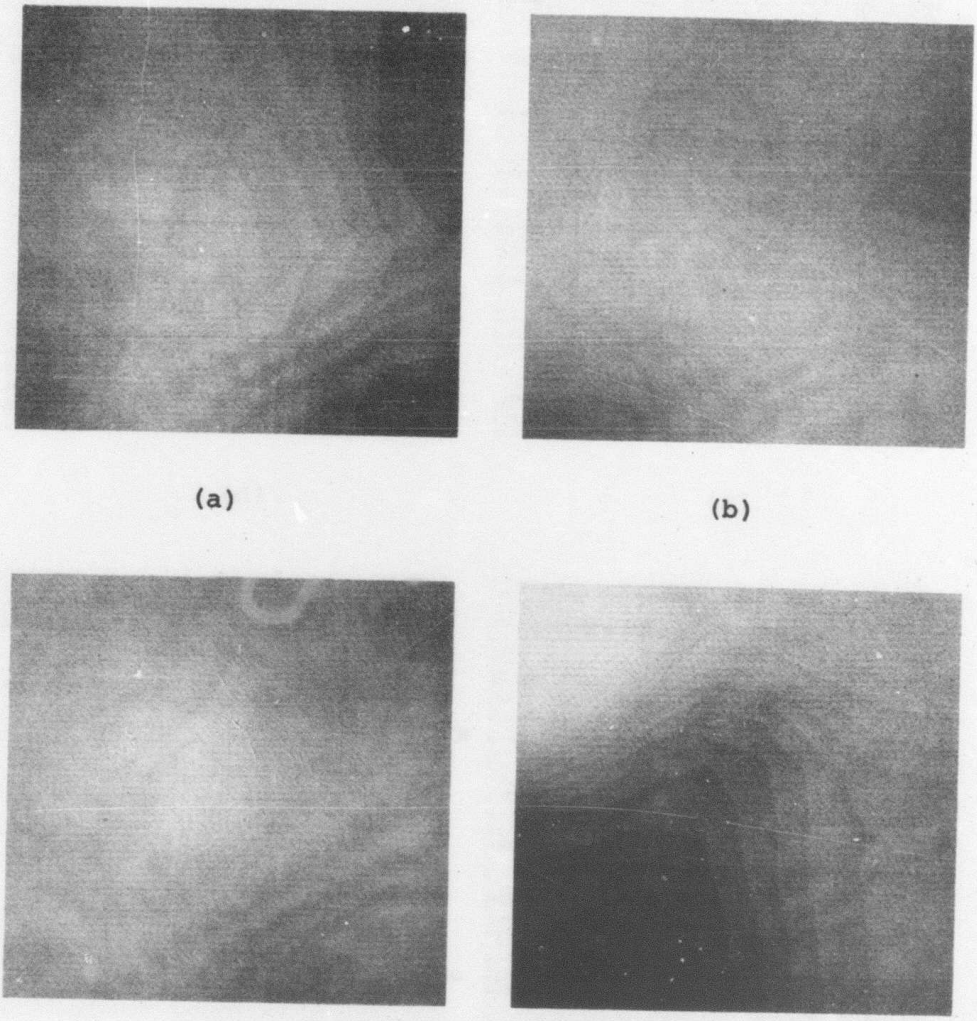


Figure 9. Lighter thresholds obtained for the four pictures by interpolating on the window thresholds.

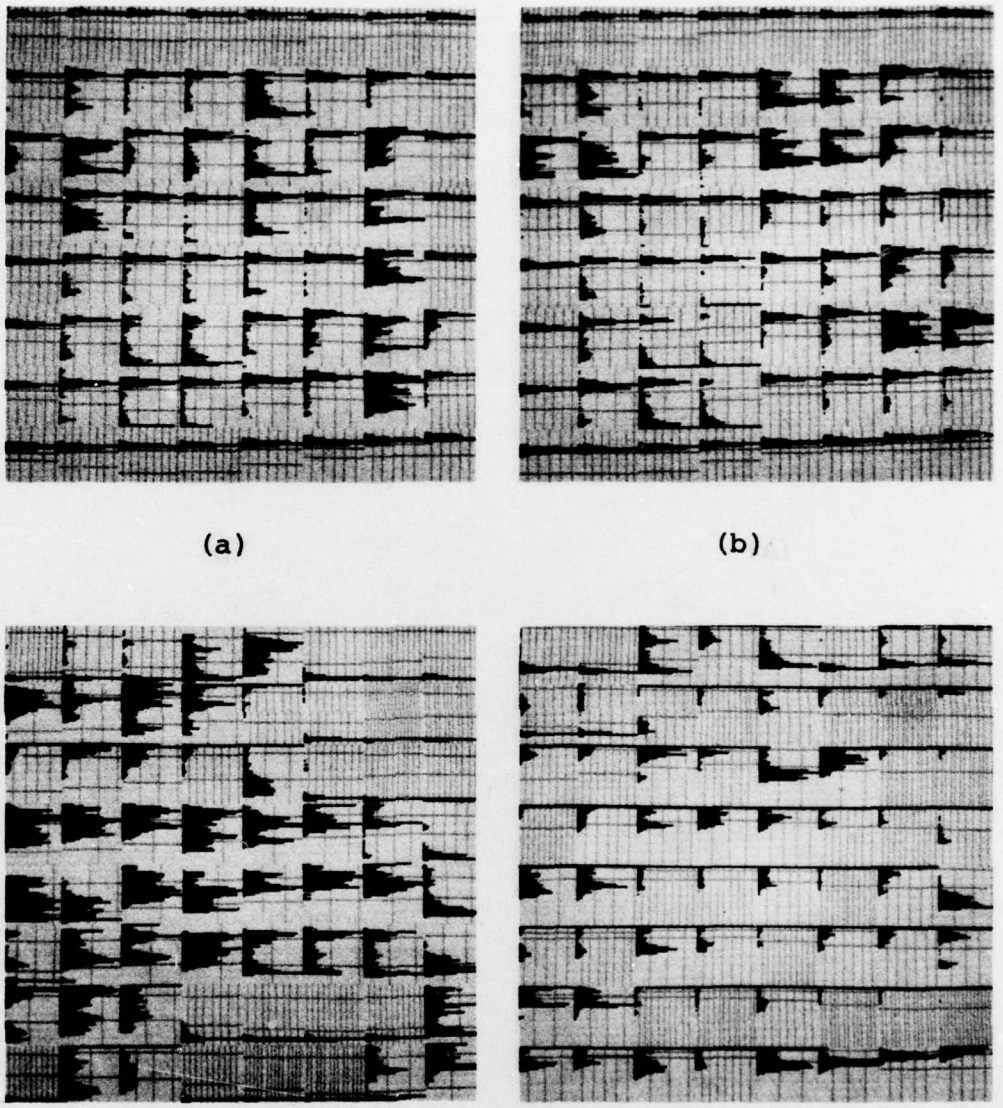
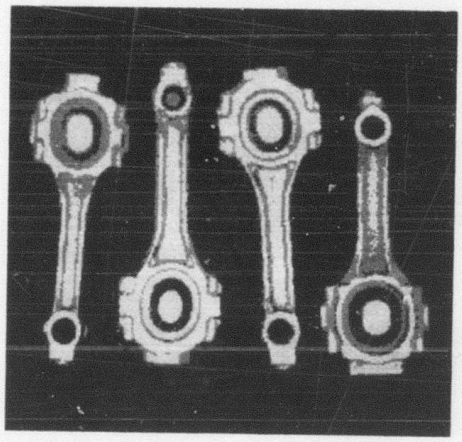
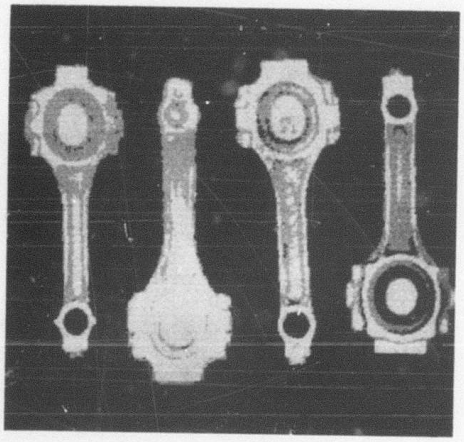


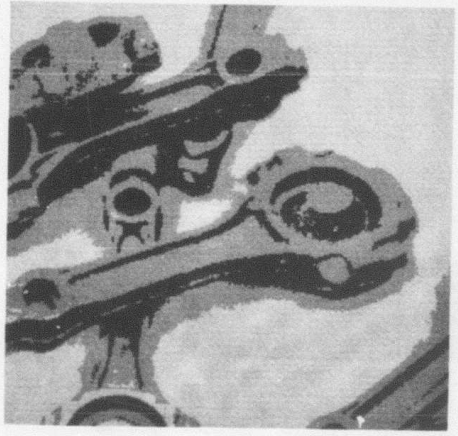
Figure 10. Window histograms with both thresholds superimposed.



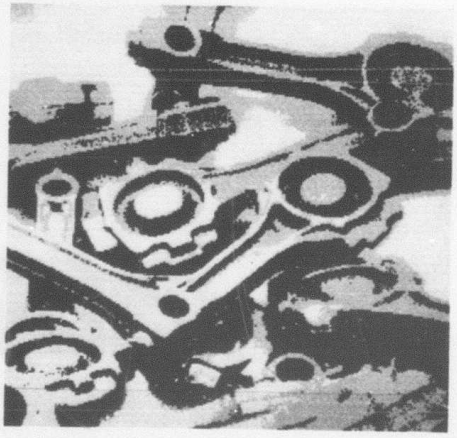
(a)



(b)



(c)



(d)

Figure 11. Three-level pictures obtained by applying the thresholds shown in Figures 8 and 9 to the pictures of Figure 1.

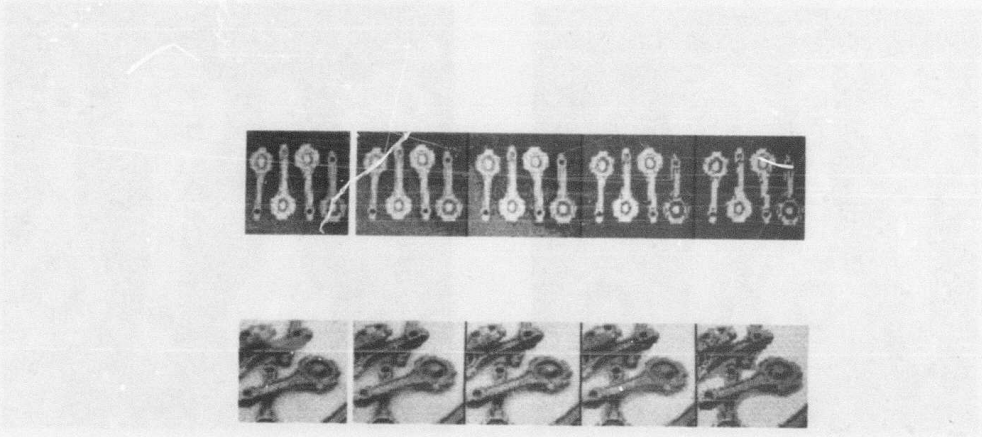


Figure 12. Results of adaptive quantization, using various amounts of smoothing, for parts (a) and (c) of Figure 1.

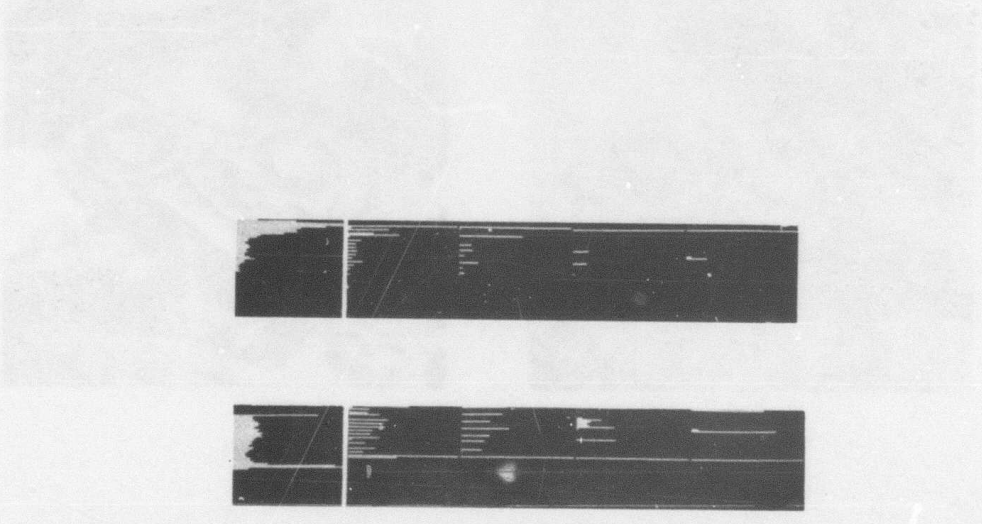


Figure 13. Histograms of the pictures in Figure 12.

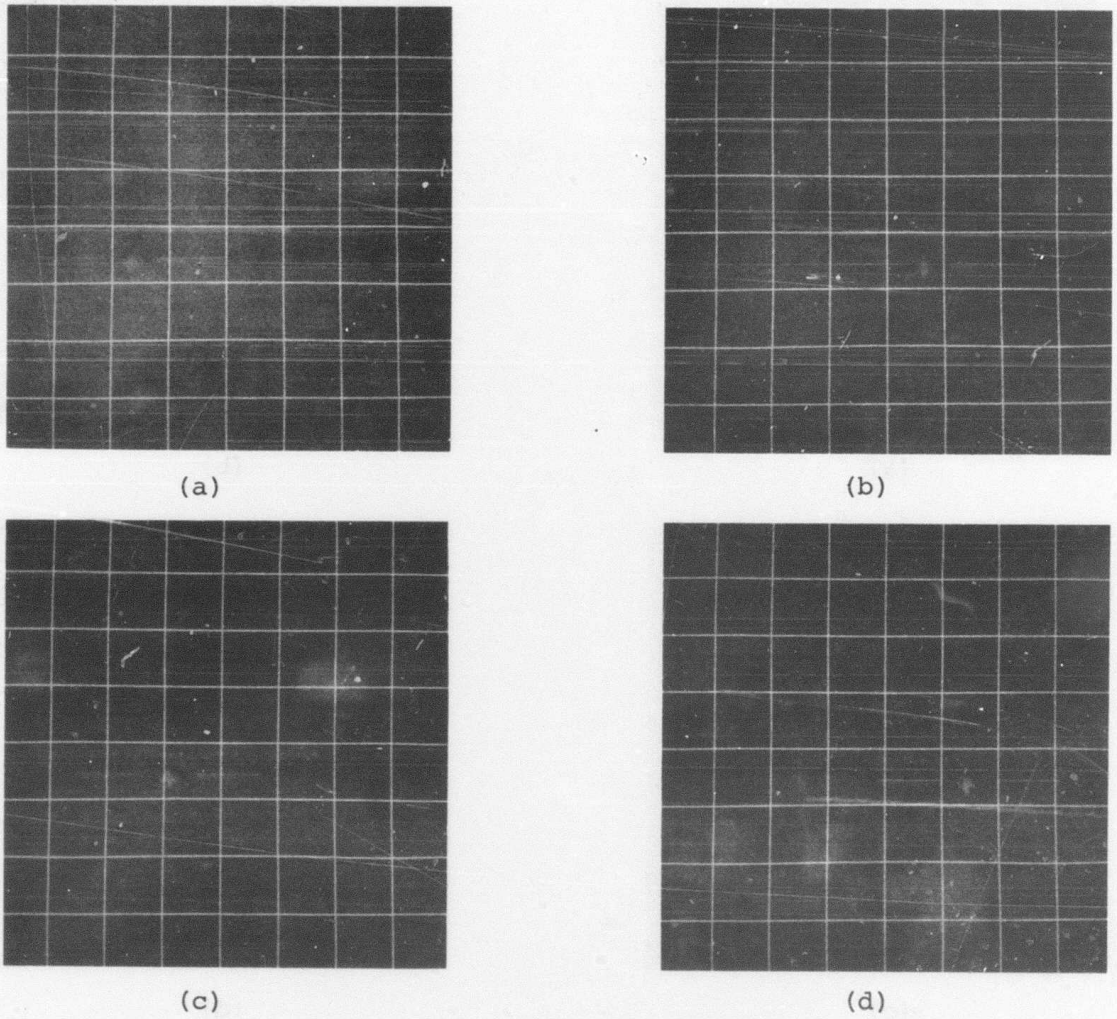
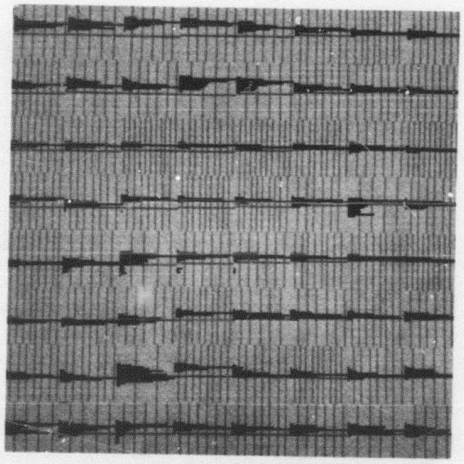
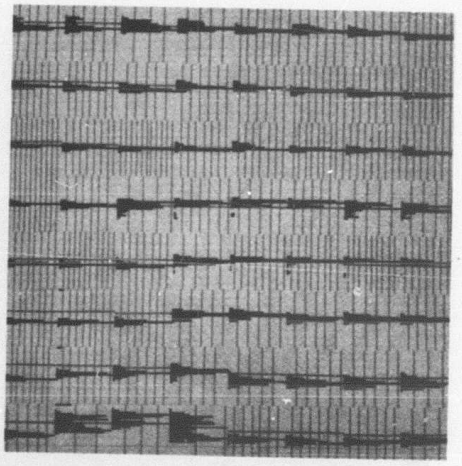


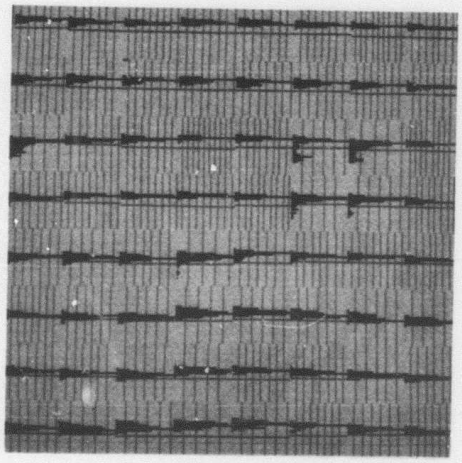
Figure 14. Four FLIR images.



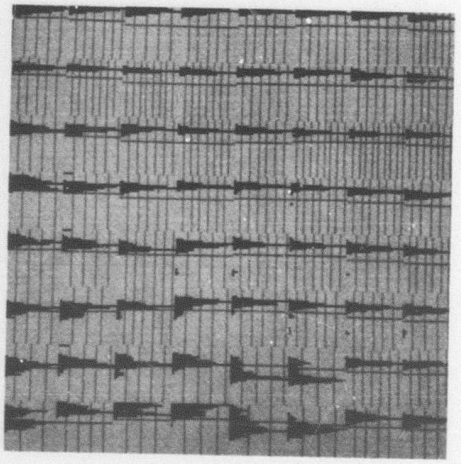
(a)



(b)

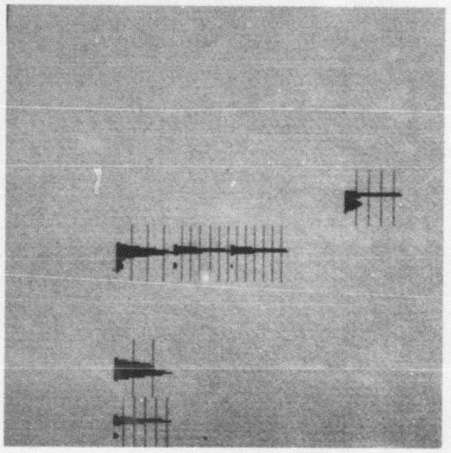


(c)

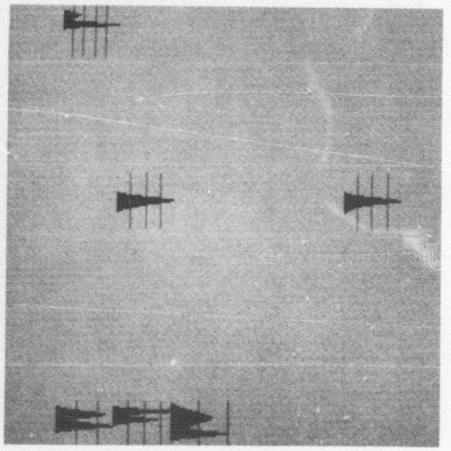


(d)

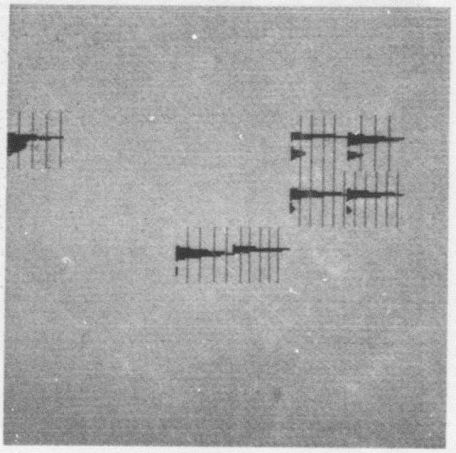
Figure 15. Window histograms for the pictures in Figure 14.



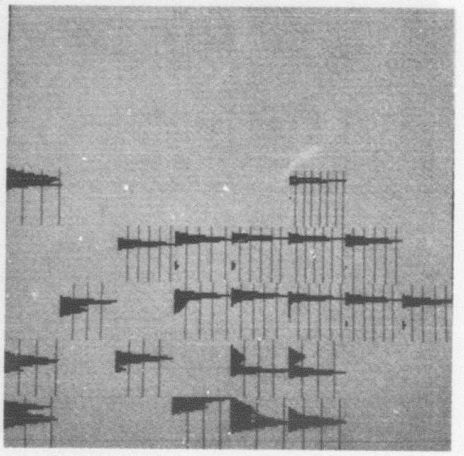
(a)



(b)

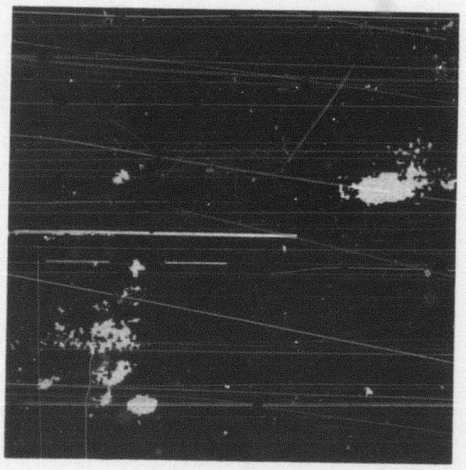


(c)

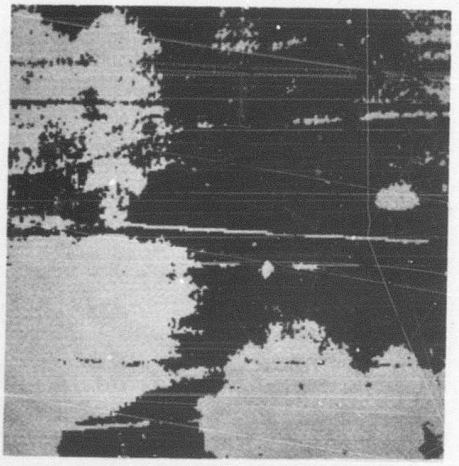


(d)

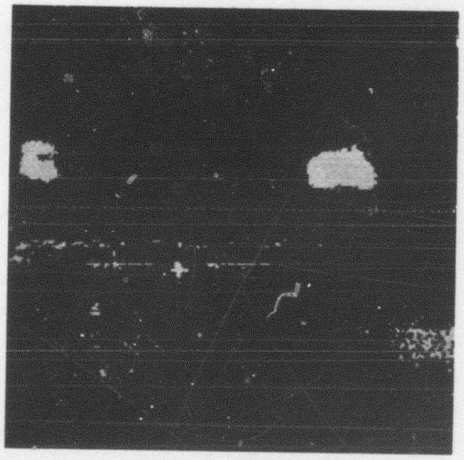
Figure 16. Two-Gaussian approximations to those window histograms in Figure 15 that were judged to be bimodal.



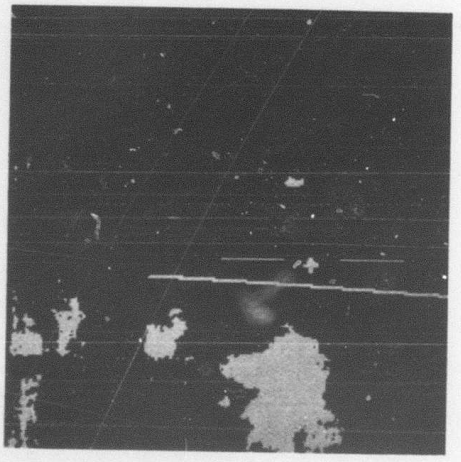
(a)



(b)

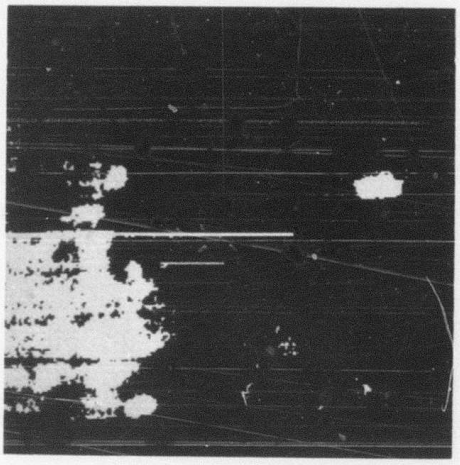


(c)

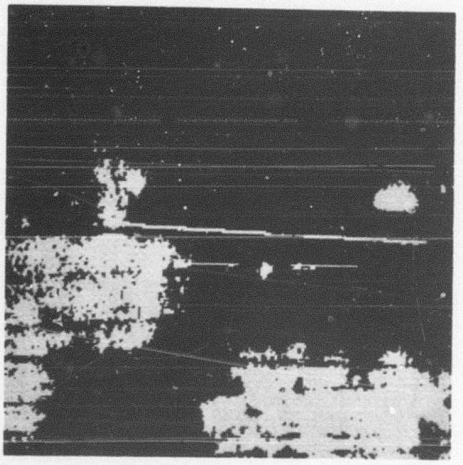


(d)

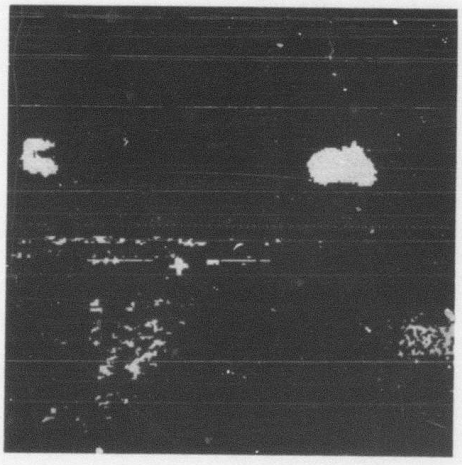
Figure 17. Results of applying the interpolated point thresholds to the pictures of Figure 14.



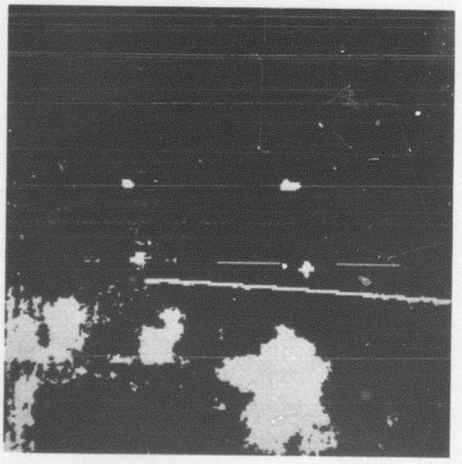
(a)



(b)



(c)



(d)

Figure 18. Results of applying fixed thresholds to the pictures of Figure 14.

UNCLASSIFIED

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 (9)
4. TITLE *(and Subtitle)* (6) SOME EXPERIMENTS ON VARIABLE THRESHOLDING	5. TYPE OF REPORT & PERIOD COVERED Technical Rept.
	6. PERFORMING ORG. REPORT NUMBER (14) TR-626
7. AUTHOR(s) (10) Yasuo Nakagawa and Azriel Rosenfeld	8. CONTRACT OR GRANT NUMBER(s) (15) DAAG53-76-C-0138, ARPA Order-3206
9. PERFORMING ORGANIZATION NAME AND ADDRESS Computer Science Ctr. Univ. of Maryland College Pk., MD 20742	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS (12) 37 p.	(11) 12. REPORT DATE January 1978 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
Pattern recognition
Segmentation
Thresholding
Histograms

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

Chow and Kaneko proposed a method of variable thresholding in which an image is divided into windows; thresholds are selected for those windows that have bimodal histograms; and these thresholds are interpolated to define a variable threshold for the entire image. This method was applied to several TV images of machine parts; the results obtained appeared to be considerably better than the results of thresholding at a fixed level. An extension of the method was

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

403 018

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

Abstract:

defined that allowed histograms to be either bimodal or trimodal; this yielded some further improvement in the results, but was also more sensitive to shadows. Finally, an adaptive quantization scheme, based on histogram peak sharpening, was applied to two of the images; the results do not seem to be as good as those obtained using variable thresholding. Some results for FLIR images of tactical targets are also presented.